

Algorithms of Oppression

How Search Engines Reinforce Racism

Safiya Umoja Noble



NEW YORK UNIVERSITY PRESS
New York

Introduction



The Power of Algorithms

This book is about the power of algorithms in the age of neoliberalism and the ways those digital decisions reinforce oppressive social relationships and enact new modes of racial profiling, which I have termed *technological redlining*. By making visible the ways that capital, race, and gender are factors in creating unequal conditions, I am bringing light to various forms of technological redlining that are on the rise. The near-ubiquitous use of algorithmically driven software, both visible and invisible to everyday people, demands a closer inspection of what values are prioritized in such automated decision-making systems. Typically, the practice of redlining has been most often used in real estate and banking circles, creating and deepening inequalities by race, such that, for example, people of color are more likely to pay higher interest rates or premiums just because they are Black or Latino, especially if they live in low-income neighborhoods. On the Internet and in our everyday uses of technology, discrimination is also embedded in computer code and, increasingly, in artificial intelligence technologies that we are reliant on, by choice or not. I believe that artificial intelligence will become a major human rights issue in the twenty-first century. We are only beginning to understand the long-term consequences of these decision-making tools in both masking and deepening social inequality. This book is just the start of trying to make these consequences visible. There will be many more, by myself and others, who will try to make sense of the consequences of automated decision making through algorithms in society.

Part of the challenge of understanding algorithmic oppression is to understand that mathematical formulations to drive automated decisions are made by human beings. While we often think of terms such as “big data” and “algorithms” as being benign, neutral, or objective, they are anything but. The people who make these decisions hold all types of

values, many of which openly promote racism, sexism, and false notions of meritocracy, which is well documented in studies of Silicon Valley and other tech corridors.

For example, in the midst of a federal investigation of Google's alleged persistent wage gap, where women are systematically paid less than men in the company's workforce, an "antidiversity" manifesto authored by James Damore went viral in August 2017,¹ supported by many Google employees, arguing that women are psychologically inferior and incapable of being as good at software engineering as men, among other patently false and sexist assertions. As this book was moving into press, many Google executives and employees were actively rebuking the assertions of this engineer, who reportedly works on Google search infrastructure. Legal cases have been filed, boycotts of Google from the political far right in the United States have been invoked, and calls for greater expressed commitments to gender and racial equity at Google and in Silicon Valley writ large are under way. What this antidiversity screed has underscored for me as I write this book is that some of the very people who are developing search algorithms and architecture are willing to promote sexist and racist attitudes openly at work and beyond, while we are supposed to believe that these same employees are developing "neutral" or "objective" decision-making tools. Human beings are developing the digital platforms we use, and as I present evidence of the recklessness and lack of regard that is often shown to women and people of color in some of the output of these systems, it will become increasingly difficult for technology companies to separate their systematic and inequitable employment practices, and the far-right ideological bents of some of their employees, from the products they make for the public.

My goal in this book is to further an exploration into some of these digital sense-making processes and how they have come to be so fundamental to the classification and organization of information and at what cost. As a result, this book is largely concerned with examining the commercial co-optation of Black identities, experiences, and communities in the largest and most powerful technology companies to date, namely, Google. I closely read a few distinct cases of algorithmic oppression for the depth of their social meaning to raise a public discussion of the broader implications of how privately managed, black-boxed information-sorting tools have become essential to many data-driven

decisions. I want us to have broader public conversations about the implications of the artificial intelligentsia for people who are already systematically marginalized and oppressed. I will also provide evidence and argue, ultimately, that large technology monopolies such as Google need to be broken up and regulated, because their consolidated power and cultural influence make competition largely impossible. This monopoly in the information sector is a threat to democracy, as is currently coming to the fore as we make sense of information flows through digital media such as Google and Facebook in the wake of the 2016 United States presidential election.

I situate my work against the backdrop of a twelve-year professional career in multicultural marketing and advertising, where I was invested in building corporate brands and selling products to African Americans and Latinos (before I became a university professor). Back then, I believed, like many urban marketing professionals, that companies must pay attention to the needs of people of color and demonstrate respect for consumers by offering services to communities of color, just as is done for most everyone else. After all, to be responsive and responsible to marginalized consumers was to create more market opportunity. I spent an equal amount of time doing risk management and public relations to insulate companies from any adverse risk to sales that they might experience from inadvertent or deliberate snubs to consumers of color who might perceive a brand as racist or insensitive. Protecting my former clients from enacting racial and gender insensitivity and helping them bolster their brands by creating deep emotional and psychological attachments to their products among communities of color was my professional concern for many years, which made an experience I had in fall 2010 deeply impactful. In just a few minutes while searching on the web, I experienced the perfect storm of insult and injury that I could not turn away from. While Googling things on the Internet that might be interesting to my stepdaughter and nieces, I was overtaken by the results. My search on the keywords "black girls" yielded HotBlackPussy.com as the first hit.

Hit indeed.

Since that time, I have spent innumerable hours teaching and researching all the ways in which it could be that Google could completely fail when it came to providing reliable or credible information about

► Sugary Black Pussy .com-Black girls in a hardcore action galeries
sugaryblackpussy.com/
 (black pussy and hairy black pussy,black sex,black booty,black ass,black teen pussy,big black ass,black porn star,hot black girl) ...

Figure I.1. First search result on keywords “black girls,” September 2011.

women and people of color yet experience seemingly no repercussions whatsoever. Two years after this incident, I collected searches again, only to find similar results, as documented in figure I.1.

In 2012, I wrote an article for *Bitch* magazine about how women and feminism are marginalized in search results. By August 2012, Panda (an update to Google’s search algorithm) had been released, and pornography was no longer the first series of results for “black girls”; but other girls and women of color, such as Latinas and Asians, were still pornified. By August of that year, the algorithm changed, and porn was suppressed in the case of a search on “black girls.” I often wonder what kind of pressures account for the changing of search results over time. It is impossible to know when and what influences proprietary algorithmic design, other than that human beings are designing them and that they are not up for public discussion, except as we engage in critique and protest.

This book was born to highlight cases of such algorithmically driven data failures that are specific to people of color and women and to underscore the structural ways that racism and sexism are fundamental to what I have coined *algorithmic oppression*. I am writing in the spirit of other critical women of color, such as Latoya Peterson, cofounder of the blog *Racialicious*, who has opined that racism is the fundamental application program interface (API) of the Internet. Peterson has argued that anti-Blackness is the foundation on which all racism toward other groups is predicated. Racism is a standard protocol for organizing behavior on the web. As she has said, so perfectly, “The idea of a n*gger API makes me think of a racism API, which is one of our core arguments all along—oppression operates in the same formats, runs the same scripts over and over. It is tweaked to be context specific, but it’s all the same source code. And the key to its undoing is recognizing how many of us are ensnared in these same basic patterns and modifying our

own actions.”² Peterson’s allegation is consistent with what many people feel about the hostility of the web toward people of color, particularly in its anti-Blackness, which any perusal of YouTube comments or other message boards will serve up. On one level, the everyday racism and commentary on the web is an abhorrent thing in itself, which has been detailed by others; but it is entirely different with the corporate platform vis-à-vis an algorithmically crafted web search that offers up racism and sexism as the first results. This process reflects a corporate logic of either willful neglect or a profit imperative that makes money from racism and sexism. This inquiry is the basis of this book.

In the following pages, I discuss how “hot,” “sugary,” or any other kind of “black pussy” can surface as the primary representation of Black girls and women on the first page of a Google search, and I suggest that something other than the best, most credible, or most reliable information output is driving Google. Of course, Google Search is an advertising company, not a reliable information company. At the very least, we must ask when we find these kinds of results, Is this the best information? For whom? We must ask ourselves who the intended audience is for a variety of things we find, and question the legitimacy of being in a “filter bubble,”³ when we do not want racism and sexism, yet they still find their way to us. The implications of algorithmic decision making of this sort extend to other types of queries in Google and other digital media platforms, and they are the beginning of a much-needed reassessment of information as a public good. We need a full-on reevaluation of the implications of our information resources being governed by corporate-controlled advertising companies. I am adding my voice to a number of scholars such as Helen Nissenbaum and Lucas Introna, Siva Vaidhyanathan, Alex Halavais, Christian Fuchs, Frank Pasquale, Kate Crawford, Tarleton Gillespie, Sarah T. Roberts, Jaron Lanier, and Elad Segev, to name a few, who are raising critiques of Google and other forms of corporate information control (including artificial intelligence) in hopes that more people will consider alternatives.

Over the years, I have concentrated my research on unveiling the many ways that African American people have been contained and constrained in classification systems, from Google’s commercial search engine to library databases. The development of this concentration was born of my research training in library and information science. I think

of these issues through the lenses of critical information studies and critical race and gender studies. As marketing and advertising have directly shaped the ways that marginalized people have come to be represented by digital records such as search results or social network activities, I have studied why it is that digital media platforms are resoundingly characterized as “neutral technologies” in the public domain and often, unfortunately, in academia. Stories of “glitches” found in systems do not suggest that the organizing logics of the web could be broken but, rather, that these are occasional one-off moments when something goes terribly wrong with near-perfect systems. With the exception of the many scholars whom I reference throughout this work and the journalists, bloggers, and whistleblowers whom I will be remiss in not naming, very few people are taking notice. We need all the voices to come to the fore and impact public policy on the most unregulated social experiment of our times: the Internet.

These data aberrations have come to light in various forms. In 2015, *U.S. News and World Report* reported that a “glitch” in Google’s algorithm led to a number of problems through auto-tagging and facial-recognition software that was apparently intended to help people search through images more successfully. The first problem for Google was that its photo application had automatically tagged African Americans as “apes” and “animals.”⁴ The second major issue reported by the *Post* was that Google Maps searches on the word “N*gger”⁵ led to a map of the White House during Obama’s presidency, a story that went viral on the Internet after the social media personality Deray McKesson tweeted it.

These incidents were consistent with the reports of Photoshopped images of a monkey’s face on the image of First Lady Michelle Obama that were circulating through Google Images search in 2009. In 2015, you could still find digital traces of the Google autosuggestions that associated Michelle Obama with apes. Protests from the White House led to Google forcing the image down the image stack, from the first page, so that it was not as visible.⁶ In each case, Google’s position is that it is not responsible for its algorithm and that problems with the results would be quickly resolved. In the *Washington Post* article about “N*gger House,” the response was consistent with other apologies by the company: “Some inappropriate results are surfacing in Google Maps that should not be, and we apologize for any offense this may have caused,”



Figure I.2. Google Images results for the keyword “gorillas,” April 7, 2016.

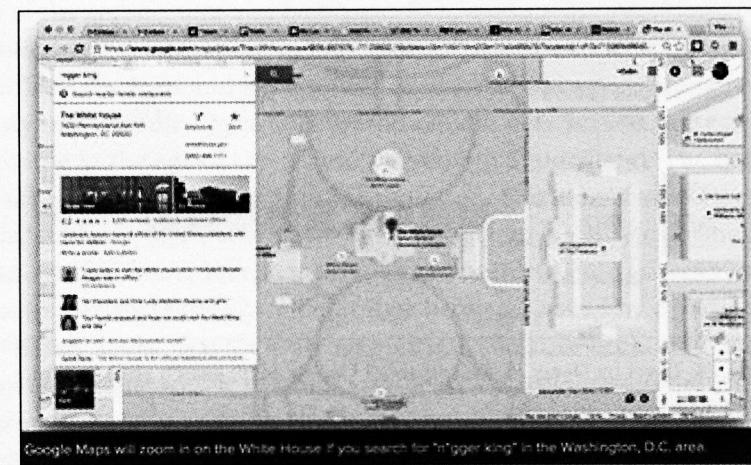


Figure I.3. Google Maps search on “N*gger House” leads to the White House, April 7, 2016.

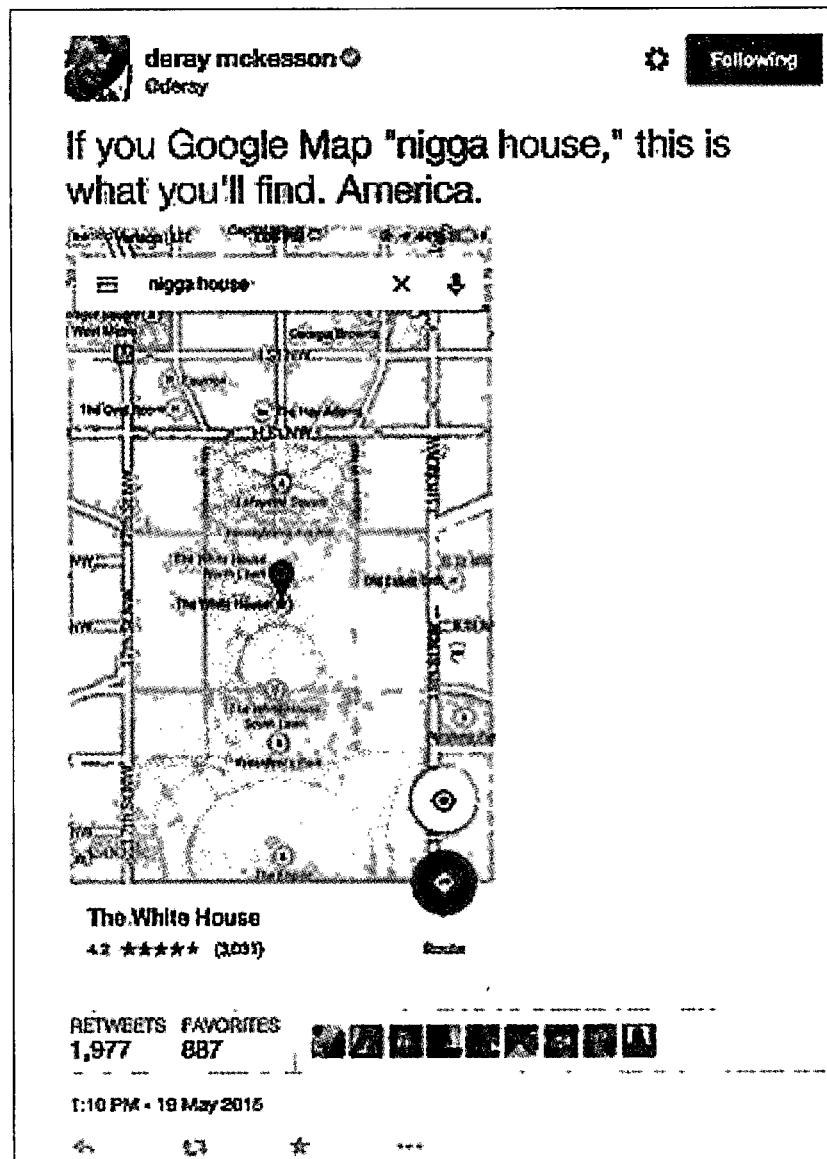


Figure I.4. Tweet by Deray McKesson about Google Maps search and the White House, 2015.

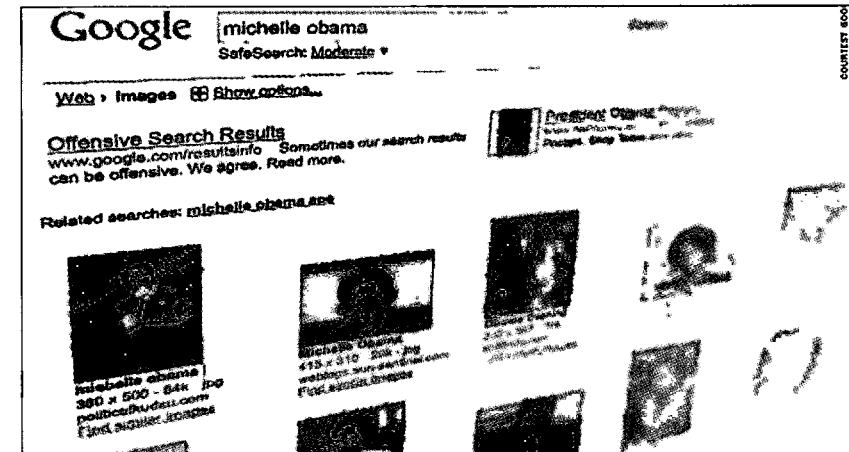


Figure I.5. Standard Google's "related" searches associates "Michelle Obama" with the term "ape."

a Google spokesperson told U.S. News in an email late Tuesday. 'Our teams are working to fix this issue quickly.'⁷

* * *

These human and machine errors are not without consequence, and there are several cases that demonstrate how racism and sexism are part of the architecture and language of technology, an issue that needs attention and remediation. In many ways, these cases that I present are specific to the lives and experiences of Black women and girls, people largely understudied by scholars, who remain ever precarious, despite our living in the age of Oprah and Beyoncé in Shondaland. The implications of such marginalization are profound. The insights about sexist or racist biases that I convey here are important because information organizations, from libraries to schools and universities to governmental agencies, are increasingly reliant on or being displaced by a variety of web-based "tools" as if there are no political, social, or economic consequences of doing so. We need to imagine new possibilities in the area of information access and knowledge generation, particularly as headlines about "racist algorithms" continue to surface in the media with limited discussion and analysis beyond the superficial.

Inevitably, a book written about algorithms or Google in the twenty-first century is out of date immediately upon printing. Technology is changing rapidly, as are technology company configurations via mergers, acquisitions, and dissolutions. Scholars working in the fields of information, communication, and technology struggle to write about specific moments in time, in an effort to crystallize a process or a phenomenon that may shift or morph into something else soon thereafter. As a scholar of information and power, I am most interested in communicating a series of processes that have happened, which provide evidence of a constellation of concerns that the public might take up as meaningful and important, particularly as technology impacts social relations and creates unintended consequences that deserve greater attention. I have been writing this book for several years, and over time, Google's algorithms have admittedly changed, such that a search for "black girls" does not yield nearly as many pornographic results now as it did in 2011. Nonetheless, new instances of racism and sexism keep appearing in news and social media, and so I use a variety of these cases to make the point that algorithmic oppression is not just a glitch in the system but, rather, is fundamental to the operating system of the web. It has direct impact on users and on our lives beyond using Internet applications. While I have spent considerable time researching Google, this book tackles a few cases of other algorithmically driven platforms to illustrate how algorithms are serving up deleterious information about people, creating and normalizing structural and systemic isolation, or practicing digital redlining, all of which reinforce oppressive social and economic relations.

While organizing this book, I have wanted to emphasize one main point: there is a missing social and human context in some types of algorithmically driven decision making, and this matters for everyone engaging with these types of technologies in everyday life. It is of particular concern for marginalized groups, those who are problematically represented in erroneous, stereotypical, or even pornographic ways in search engines and who have also struggled for nonstereotypical or nonracist and nonsexist depictions in the media and in libraries. There is a deep body of extant research on the harmful effects of stereotyping of women and people of color in the media, and I encourage

readers of this book who do not understand why the perpetuation of racist and sexist images in society is problematic to consider a deeper dive into such scholarship.

This book is organized into six chapters. In chapter 1, I explore the important theme of corporate control over public information, and I show several key Google searches. I look to see what kinds of results Google's search engine provides about various concepts, and I offer a cautionary discussion of the implications of what these results mean in historical and social contexts. I also show what Google Images offers on basic concepts such as "beauty" and various professional identities and why we should care.

In chapter 2, I discuss how Google Search reinforces stereotypes, illustrated by searches on a variety of identities that include "black girls," "Latinas," and "Asian girls." Previously, in my work published in the *Black Scholar*,⁸ I looked at the postmortem Google autosuggest searches following the death of Trayvon Martin, an African American teenager whose murder ignited the #BlackLivesMatter movement on Twitter and brought attention to the hundreds of African American children, women, and men killed by police or extrajudicial law enforcement. To add a fuller discussion to that research, I elucidate the processes involved in Google's PageRank search protocols, which range from leveraging digital footprints from people⁹ to the way advertising and marketing interests influence search results to how beneficial this is to the interests of Google as it profits from racism and sexism, particularly at the height of a media spectacle.

In chapter 3, I examine the importance of noncommercial search engines and information portals, specifically looking at the case of how a mass shooter and avowed White supremacist, Dylann Roof, allegedly used Google Search in the development of his racial attitudes, attitudes that led to his murder of nine African American AME Church members while they worshiped in their South Carolina church in the summer of 2015. The provision of false information that purports to be credible news, and the devastating consequences that can come from this kind of algorithmically driven information, is an example of why we cannot afford to outsource and privatize uncurated information on the increasingly neoliberal, privatized web. I show how important records

are to the public and explore the social importance of both remembering and forgetting, as digital media platforms thrive on never or rarely forgetting. I discuss how information online functions as a type of record, and I argue that much of this information and its harmful effects should be regulated or subject to legal protections. Furthermore, at a time when “right to be forgotten” legislation is gaining steam in the European Union, efforts to regulate the ways that technology companies hold a monopoly on public information about individuals and groups need further attention in the United States. Chapter 3 is about the future of information culture, and it underscores the ways that information is not neutral and how we can reimagine information culture in the service of eradicating social inequality.

Chapter 4 is dedicated to critiquing the field of information studies and foregrounds how these issues of public information through classification projects on the web, such as commercial search, are old problems that we must solve as a scholarly field of researchers and practitioners. I offer a brief survey of how library classification projects undergird the invention of search engines such as Google and how our field is implicated in the algorithmic process of sorting and classifying information and records. In chapter 5, I discuss the future of knowledge in the public and reference the work of library and information professionals, in particular, as important to the development and cultivation of equitable classification systems, since these are the precursors to commercial search engines. This chapter is essential history for library and information professionals, who are less likely to be trained on the politics of cataloguing and classification bias in their professional training. Chapter 6 explores public policy and why we need regulation in our information environments, particularly as they are increasingly controlled by corporations.

To conclude, I move the discussion beyond Google, to help readers think about the impact of algorithms on how people are represented in other seemingly benign business transactions. I look at the “color-blind” organizing logic of Yelp and how business owners are revolting due to loss of control over how they are represented and the impact of how the public finds them. Here, I share an interview with Kandis from New York,¹⁰ whose livelihood has been dramatically affected by public-policy changes such as the dismantling of affirmative action on

college campuses, which have hurt her local Black-hair-care business in a prestigious college town. Her story brings to light the power that algorithms have on her everyday life and leaves us with more to think about in the ecosystem of algorithmic power. The book closes with a call to recognize the importance of how algorithms are shifting social relations in many ways—more ways than this book can cover—and should be regulated with more impactful public policy in the United States than we currently have. My hope is that this book will directly impact the many kinds of algorithmic decisions that can have devastating consequences for people who are already marginalized by institutional racism and sexism, including the 99% who own so little wealth in the United States that the alarming trend of social inequality is not likely to reverse without our active resistance and intervention. Electoral politics and financial markets are just two of many of these institutional wealth-consolidation projects that are heavily influenced by algorithms and artificial intelligence. We need to cause a shift in what we take for granted in our everyday use of digital media platforms.

I consider my work a practical project, the goal of which is to eliminate social injustice and change the ways in which people are oppressed with the aid of allegedly neutral technologies. My intention in looking at these cases serves two purposes. First, we need interdisciplinary research and scholarship in information studies and library and information science that intersects with gender and women’s studies, Black/African American studies, media studies, and communications to better describe and understand how algorithmically driven platforms are situated in intersectional sociohistorical contexts and embedded within social relations. My hope is that this work will add to the voices of my many colleagues across several fields who are raising questions about the legitimacy and social consequences of algorithms and artificial intelligence. Second, now, more than ever, we need experts in the social sciences and digital humanities to engage in dialogue with activists and organizers, engineers, designers, information technologists, and public-policy makers before blunt artificial-intelligence decision making trumps nuanced human decision making. This means that we must look at how the outsourcing of information practices from the public sector facilitates privatization of what we previously thought of as the public

domain¹¹ and how corporate-controlled governments and companies subvert our ability to intervene in these practices.

We have to ask what is lost, who is harmed, and what should be forgotten with the embrace of artificial intelligence in decision making. It is of no collective social benefit to organize information resources on the web through processes that solidify inequality and marginalization—on that point I am hopeful many people will agree.

1

A Society, Searching



On October 21, 2013, the United Nations launched a campaign directed by the advertising agency Memac Ogilvy & Mather Dubai using “genuine Google searches” to bring attention to the sexist and discriminatory ways in which women are regarded and denied human rights. Christopher Hunt, art director of the campaign, said, “When we came across these searches, we were shocked by how negative they were and decided we had to do something with them.” Kareem Shubaibar, a copywriter for the campaign, described on the United Nations website what the campaign was determined to show: “The ads are shocking because they show just how far we still have to go to achieve gender equality. They are a wake up call, and we hope that the message will travel far.”¹ Over the mouths of various women of color were the autosuggestions that reflected the most popular searches that take place on Google Search. The Google Search autosuggestions featured a range of sexist ideas such as the following:

- Women cannot: drive, be bishops, be trusted, speak in church
- Women should not: have rights, vote, work, box
- Women should: stay at home, be slaves, be in the kitchen, not speak in church
- Women need to: be put in their places, know their place, be controlled, be disciplined

While the campaign employed Google Search results to make a larger point about the status of public opinion toward women, it also served, perhaps unwittingly, to underscore the incredibly powerful nature of search engine results. The campaign suggests that search is a mirror of users’ beliefs and that society still holds a variety of sexist ideas about women. What I find troubling is that the campaign also reinforces the idea that it is not the search engine that is the problem but, rather, the users of search engines who are. It suggests that what is most popular is simply what rises

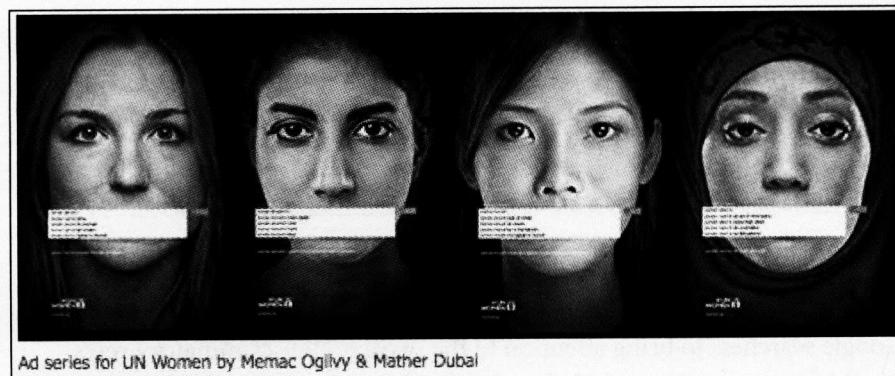


Figure 1.1. Memac Ogilvy & Mather Dubai advertising campaign for the United Nations.

to the top of the search pile. While serving as an important and disturbing critique of sexist attitudes, the campaign fails to implicate the algorithms or search engines that drive certain results to the top. This chapter moves the lens onto the search architecture itself in order to shed light on the many factors that keep sexist and racist ideas on the first page.

One limitation of looking at the implications of search is that it is constantly evolving and shifting over time. This chapter captures aspects of commercial search at a particular moment—from 2009 to 2015—but surely by the time readers engage with it, it will be a historical rather than contemporary study. Nevertheless, the goal of such an exploration of why we get troublesome search results is to help us think about whether it truly makes sense to outsource all of our knowledge needs to commercial search engines, particularly at a time when the public is increasingly reliant on search engines in lieu of libraries, librarians, teachers, researchers, and other knowledge keepers and resources.

What is even more crucial is an exploration of how people living as minority groups under the influence of a majority culture, such as people of color and sexual minorities in the United States, are often subject to the whims of the majority and other commercial influences such as advertising when trying to affect the kinds of results that search engines offer about them and their identities. If the majority rules in search engine results, then how might those who are in the minority ever be able to influence or control the way they are represented in a search engine? The same might be true of how men's desires and usage of search is able

to influence the values that surround women's identities in search engines, as the Ogilvy campaign might suggest. For these reasons, a deeper exploration into the historical and social conditions that give rise to problematic search results is in order, since rarely are they questioned and most Internet users have no idea how these ideas come to dominate search results on the first page of results in the first place.

Google Search: Racism and Sexism at the Forefront

My first encounter with racism in search came to me through an experience that pushed me, as a researcher, to explore the mechanisms—both technological and social—that could render the pornification of Black women a top search result, naturalizing Black women as sexual objects so effortlessly. This encounter was in 2009 when I was talking to a friend, André Brock at the University of Michigan, who causally mentioned one day, “You should see what happens when you Google ‘black girls.’” I did and was stunned. I assumed it to be an aberration that could potentially shift over time. I kept thinking about it. The second time came one spring morning in 2011, when I searched for activities to entertain my preteen stepdaughter and her cousins of similar age, all of whom had made a weekend visit to my home, ready for a day of hanging out that would inevitably include time on our laptops. In order to break them away from mindless TV watching and cellphone gazing, I wanted to engage them in conversations about what was important to them and on their mind, from their perspective as young women growing up in downstate Illinois, a predominantly conservative part of Middle America. I felt that there had to be some great resources for young people of color their age, if only I could locate them. I quickly turned to the computer I used for my research (I was pursuing doctoral studies at the time), but I did not let the group of girls gather around me just yet. I opened up Google to enter in search terms that would reflect their interests, demographics, and information needs, but I liked to prescreen and anticipate what could be found on the web, in order to prepare for what might be in store. What came back from that simple, seemingly innocuous search was again nothing short of shocking: with the girls just a few feet away giggling and snorting at their own jokes, I again retrieved a Google Search results page filled with porn when I looked for “black girls.” By then, I thought that my own

search history and engagement with a lot of Black feminist texts, videos, and books on my laptop would have shifted the kinds of results I would get. It had not. In intending to help the girls search for information about themselves, I had almost inadvertently exposed them to one of the most graphic and overt illustrations of what the advertisers already thought about them: Black girls were still the fodder of porn sites, dehumanizing them as commodities, as products and as objects of sexual gratification. I closed the laptop and redirected our attention to fun things we might do, such as see a movie down the street. This best information, as listed by rank in the search results, was certainly not the best information for me or for the children I love. For whom, then, was this the best information, and who decides? What were the profit and other motives driving this information to the top of the results? How had the notion of neutrality in information ranking and retrieval gone so sideways as to be perhaps one of the worst examples of racist and sexist classification of Black women in the digital age yet remain so unexamined and without public critique? That moment, I began in earnest a series of research inquiries that are central to this book.

Of course, upon reflection, I realized that I had been using the web and search tools long before the encounters I experienced just out of view of my young family members. It was just as troubling to realize that I had undoubtedly been confronted with the same type of results before but had learned, or been trained, to somehow become inured to it, to take it as a given that any search I might perform using keywords connected to my physical self and identity could return pornographic and otherwise disturbing results. Why was this the bargain into which I had tacitly entered with digital information tools? And who among us did not have to bargain in this way? As a Black woman growing up in the late twentieth century, I also knew that the presentation of Black women and girls that I discovered in my search results was not a new development of the digital age. I could see the connection between search results and tropes of African Americans that are as old and endemic to the United States as the history of the country itself. My background as a student and scholar of Black studies and Black history, combined with my doctoral studies in the political economy of digital information, aligned with my righteous indignation for Black girls everywhere. I searched on.

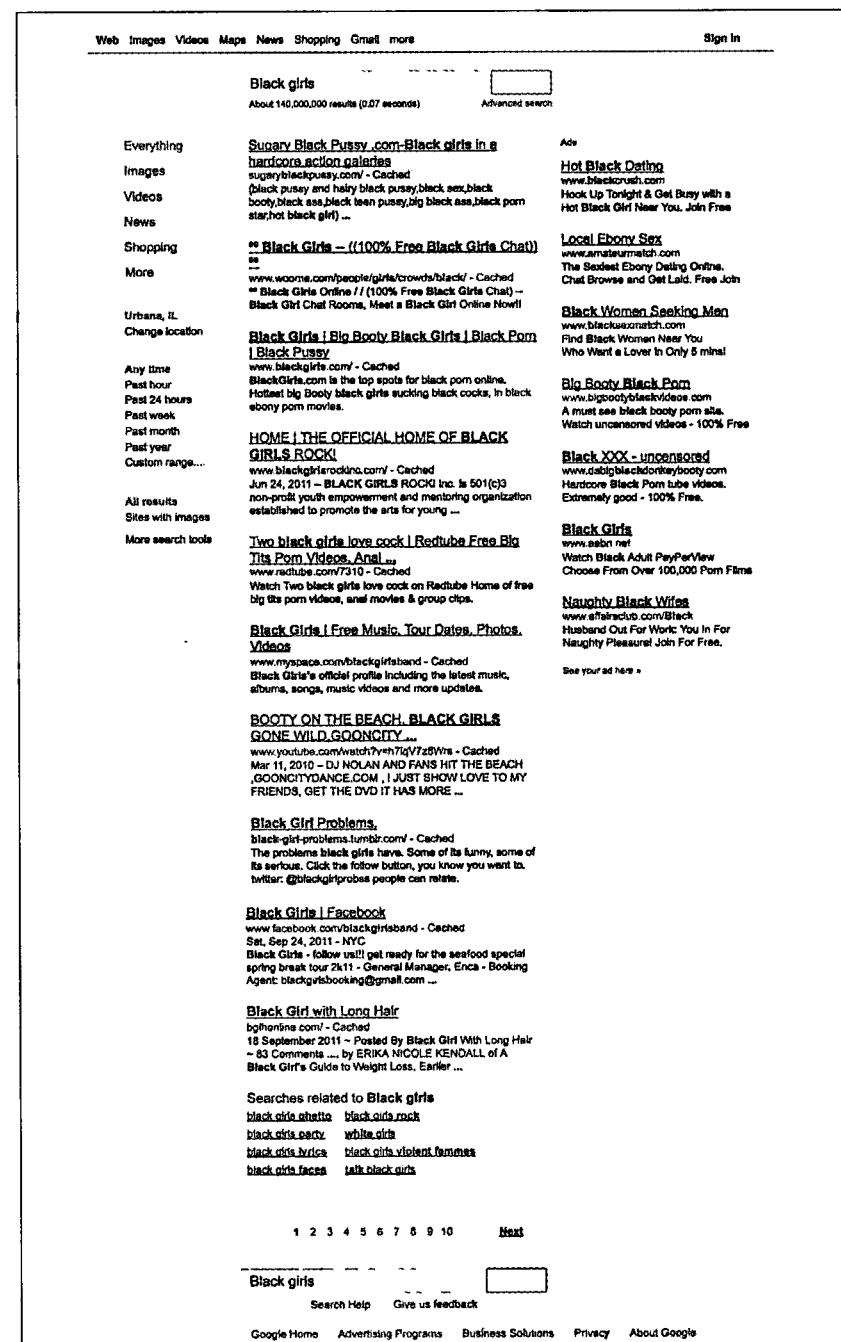


Figure 1.2. First page of search results on keywords "black girls," September 18, 2011.

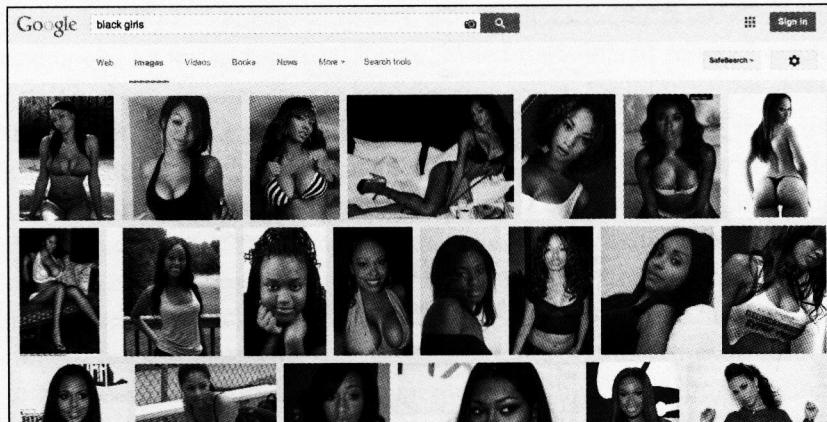


Figure 1.3. First page of image search results on keywords “black girls,” April 3, 2014.



Figure 1.4. Google autosuggest results when searching the phrase “why are black people so,” January 25, 2013.

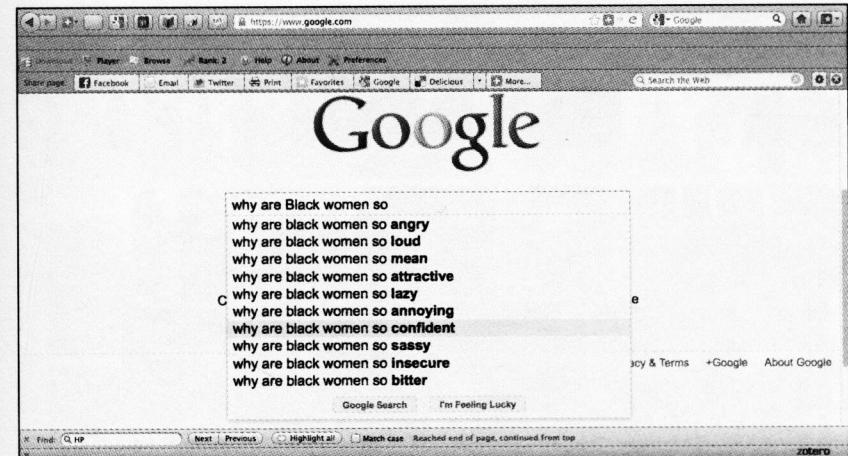


Figure 1.5. Google autosuggest results when searching the phrase “why are black women so,” January 25, 2013.

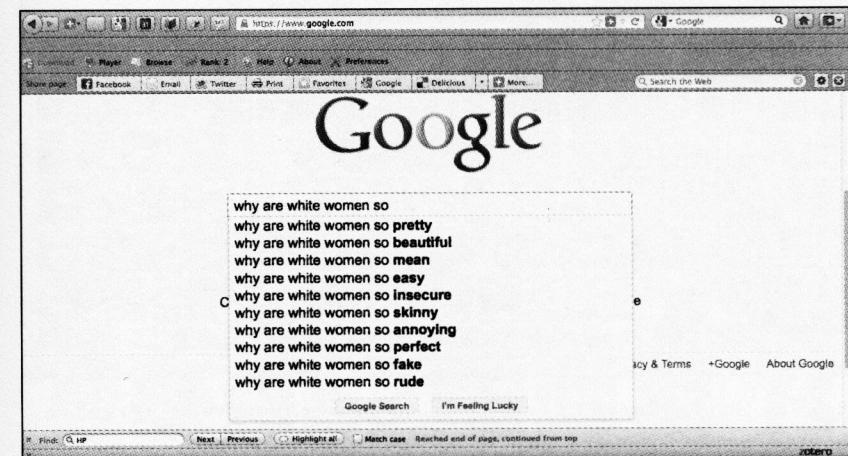


Figure 1.6. Google autosuggest results when searching the phrase “why are white women so,” January 25, 2013.

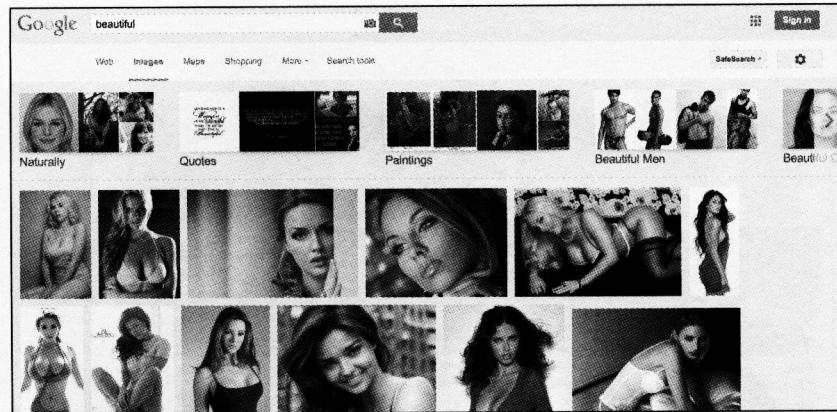


Figure 1.7. Google Images results when searching the concept “beautiful” (did not include the word “women”), December 4, 2014.



Figure 1.8. Google Images results when searching the concept “ugly” (did not include the word “women”), January 5, 2013.



Figure 1.9. Google Images results when searching the phrase “professor style” while logged in as myself, September 15, 2015.

What each of these searches represents are Google's algorithmic conceptualizations of a variety of people and ideas. Whether looking for autosuggestions or answers to various questions or looking for notions about what is beautiful or what a professor may look like (which does not account for people who look like me who are part of the professoriate—so much for “personalization”), Google's dominant narratives reflect the kinds of hegemonic frameworks and notions that are often resisted by women and people of color. Interrogating what advertising companies serve up as credible information must happen, rather than have a public instantly gratified with stereotypes in three-hundredths of a second or less.

In reality, information monopolies such as Google have the ability to prioritize web search results on the basis of a variety of topics, such as promoting their own business interests over those of competitors or smaller companies that are less profitable advertising clients than larger multinational corporations are.² In this case, the clicks of users, coupled with the commercial processes that allow paid advertising to be prioritized in search results, mean that representations of women are ranked on a search engine page in ways that underscore women's historical and contemporary lack of status in society—a direct mapping of old media traditions into new media architecture. Problematic representations and biases in classifications are not new. Critical library and information science scholars have well documented the ways in which some groups are more vulnerable than others to misrepresentation and misclassification.³ They have conducted extensive and important critiques of library cataloging systems and information organization patterns that demonstrate how women, Black people, Asian Americans, Jewish people, or the Roma, as “the other,” have all suffered from the insults of misrepresentation and derision in the Library of Congress Subject Headings (LCSH) or through the Dewey Decimal System. At the same time, other scholars underscore the myriad ways that social values around race and gender are directly reflected in technology design.⁴ Their contributions have made it possible for me to think about the ways that race and gender are embedded in Google's search engine and to have the courage to raise critiques of one of the most beloved and revered contemporary brands.

Search happens in a highly commercial environment, and a variety of processes shape what can be found; these results are then normalized

as believable and often presented as factual. The associate professor of sociology at Arizona State University and former president of the Association of Internet Researchers Alex Halavais points to the way that heavily used technological artifacts such as the search engine have become such a normative part of our experience with digital technology and computers that they socialize us into believing that these artifacts must therefore also provide access to credible, accurate information that is depoliticized and neutral:

Those assumptions are dangerously flawed; . . . unpacking the black box of the search engine is something of interest not only to technologists and marketers, but to anyone who wants to understand how we make sense of a newly networked world. Search engines have come to play a central role in corralling and controlling the ever-growing sea of information that is available to us, and yet they are trusted more readily than they ought to be. They freely provide, it seems, a sorting of the wheat from the chaff, and answer our most profound and most trivial questions. They have become an object of faith.⁵

Unlike the human-labor curation processes of the early Internet that led to the creation of online directories such as Lycos and Yahoo!, in the current Internet environment, information access has been left to the complex algorithms of machines to make selections and prioritize results for users. I agree with Halavais, and his is an important critique of search engines as a window into our own desires, which can have an impact on the values of society. Search is a symbiotic process that both informs and is informed in part by users. Halavais suggests that every user of a search engine should know how the system works, how information is collected, aggregated, and accessed. To achieve this vision, the public would have to have a high degree of computer programming literacy to engage deeply in the design and output of search.

Alternatively, I draw an analogy that one need not know the mechanism of radio transmission or television spectrum or how to build a cathode ray tube in order to critique racist or sexist depictions in song lyrics played on the radio or shown in a film or television show. Without a doubt, the public is unaware and must have significantly more algorithmic literacy. Since all of the platforms I interrogate in this book are

proprietary, even if we had algorithmic literacy, we still could not intervene in these private, corporate platforms.

To be specific, knowledge of the technical aspects of search and retrieval, in terms of critiquing the computer programming code that underlies the systems, is absolutely necessary to have a profound impact on these systems. Interventions such as Black Girls Code, an organization focused on teaching young, African American girls to program, is the kind of intervention we see building in response to the ways Black women have been locked out of Silicon Valley venture capital and broader participation. Simultaneously, it is important for the public, particularly people who are marginalized—such as women and girls and people of color—to be critical of the results that purport to represent them in the first ten to twenty results in a commercial search engine. They do not have the economic, political, and social capital to withstand the consequences of misrepresentation. If one holds a lot of power, one can withstand or buffer misrepresentation at a group level and often at the individual level. Marginalized and oppressed people are linked to the status of their group and are less likely to be afforded individual status and insulation from the experiences of the groups with which they are identified. The political nature of search demonstrates how algorithms are a fundamental invention of computer scientists who are human beings—and code is a language full of meaning and applied in varying ways to different types of information. Certainly, women and people of color could benefit tremendously from becoming programmers and building alternative search engines that are less disturbing and that reflect and prioritize a wider range of informational needs and perspectives.

There is an important and growing movement of scholars raising concerns. Helen Nissenbaum, a professor of media, culture, and communication and computer science at New York University, has written with Lucas Introna, a professor of organization, technology, and ethics at the Lancaster University Management School, about how search engines bias information toward the most powerful online. Their work was corroborated by Alejandro Diaz, who wrote his dissertation at Stanford on sociopolitical bias in Google's products. Kate Crawford and Tarleton Gillespie, two researchers at Microsoft Research New England, have written extensively about algorithmic bias, and Crawford recently

coorganized a summit with the White House and New York University for academics, industry, and activists concerned with the social impact of artificial intelligence in society. At that meeting, I participated in a working group on artificial-intelligence social inequality, where tremendous concern was raised about deep-machine-learning projects and software applications, including concern about furthering social injustice and structural racism. In attendance was the journalist Julia Angwin, one of the investigators of the breaking story about courtroom sentencing software Northpointe, used for risk assessment by judges to determine the alleged future criminality of defendants.⁶ She and her colleagues determined that this type of artificial intelligence miserably mispredicted future criminal activity and led to the overincarceration of Black defendants. Conversely, the reporters found it was much more likely to predict that White criminals would not offend again, despite the data showing that this was not at all accurate. Sitting next to me was Cathy O'Neil, a data scientist and the author of the book *Weapons of Math Destruction*, who has an insider's view of the way that math and big data are directly implicated in the financial and housing crisis of 2008 (which, incidentally, destroyed more African American wealth than any other event in the United States, save for not compensating African Americans for three hundred years of forced enslavement). Her view from Wall Street was telling:

The math-powered applications powering the data economy were based on choices made by fallible human beings. Some of these choices were no doubt made with the best intentions. Nevertheless, many of these models encoded human prejudice, misunderstanding, and bias into the software systems that increasingly managed our lives. Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists. Their verdicts, even when wrong or harmful, were beyond dispute or appeal. And they tended to punish the poor and the oppressed in our society, while making the rich richer.⁷

Our work, each of us, in our respective way, is about interrogating the many ways that data and computing have become so profoundly their own "truth" that even in the face of evidence, the public still struggles

to hold tech companies accountable for the products and errors of their ways. These errors increasingly lead to racial and gender profiling, misrepresentation, and even economic redlining.

At the core of my argument is the way in which Google biases search to its own economic interests—for its profitability and to bolster its market dominance at any expense. Many scholars are working to illuminate the ways in which users trade their privacy, personal information, and immaterial labor for “free” tools and services offered by Google (e.g., search engine, Gmail, Google Scholar, YouTube) while the company profits from data mining its users. Recent research on Google by Siva Vaidhyanathan, professor of media studies at the University of Virginia, who has written one of the most important books on Google to date, demonstrates its dominance over the information landscape and forms the basis of a central theme in this research. Frank Pasquale, a professor of law at the University of Maryland, has also forewarned of the increasing levels of control that algorithms have over the many decisions made about us, from credit to dating options, and how difficult it is to intervene in their discriminatory effects. The political economic critique of Google by Elad Segev, a senior lecturer of media and communication in the Department of Communication at Tel Aviv University, charges that we can no longer ignore the global dominance of Google and the implications of its power in furthering digital inequality, particularly as it serves as a site of fostering global economic divides.

However, what is missing from the extant work on Google is an intersectional power analysis that accounts for the ways in which marginalized people are exponentially harmed by Google. Since I began writing this book, Google’s parent company, Alphabet, has expanded its power into drone technology,⁸ military-grade robotics, fiber networks, and behavioral surveillance technologies such as Nest and Google Glass.⁹ These are just several of many entry points to thinking about the implications of artificial intelligence as a human rights issue. We need to be concerned about not only how ideas and people are represented but also the ethics of whether robots and other forms of automated decision making can end a life, as in the case of drones and automated weapons. To whom do we appeal? What bodies govern artificial intelligence, and where does the public raise issues or lodge

complaints with national and international courts? These questions have yet to be fully answered.

In the midst of Google’s expansion, Google Search is one of the most underexamined areas of consumer protection policy,¹⁰ and regulation has been far less successful in the United States than in the European Union. A key aspect of generating policy that protects the public is the accumulation of research about the impact of what an unregulated commercial information space does to vulnerable populations. I do this by taking a deep look at a snapshot of the web, at a specific moment in time, and interpreting the results against the history of race and gender in the U.S. This is only one of many angles that could be taken up, but I find it to be one of the most compelling ways to show how data is biased and perpetuates racism and sexism. The problems of big data go deeper than misrepresentation, for sure. They include decision-making protocols that favor corporate elites and the powerful, and they are implicated in global economic and social inequality. Deep machine learning, which is using algorithms to replicate human thinking, is predicated on specific values from specific kinds of people—namely, the most powerful institutions in society and those who control them. Diana Ascher,¹¹ in her dissertation on yellow journalism and cultural time orientation in the Department of Information Studies at UCLA, found there was a stark difference between headlines generated by social media managers from the *LA Times* and those provided by automated, algorithmically driven software, which generated severe backlash on Twitter. In this case, Ascher found that automated tweets in news media were more likely to be racist and misrepresentative, as in the case of police shooting victim Keith Lamont Scott of Charlotte, North Carolina, whose murder triggered nationwide protests of police brutality and excessive force.

There are many such examples. In the ensuing chapters, I continue to probe the results that are generated by Google on a variety of keyword combinations relating to racial and gender identity as a way of engaging a commonsense understanding of how power works, with the goal of changing these processes of control. By seeing and discussing these intersectional power relations, we have a significant opportunity to transform the consciousness embedded in artificial intelligence, since it is in fact, in part, a product of our own collective creation.



Figure 1.10. Automated headline generated by software and tweeted about Keith Lamont Scott, killed by police in North Carolina on September 20, 2016, as reported by the *Los Angeles Times*.

Theorizing Search: A Black Feminist Project

The impetus for my work comes from theorizing Internet search results from a Black feminist perspective; that is, I ask questions about the structure and results of web searches from the standpoint of a Black woman—a standpoint that drives me to ask different questions than have been previously posed about how Google Search works. This study builds on previous research that looks at the ways in which racialization is a salient factor in various engagements with digital technology represented in video games,¹² websites,¹³ virtual worlds,¹⁴ and digital media platforms.¹⁵ A Black feminist perspective offers an opportunity to ask questions about the quality and content of racial hierarchies and stereotyping that appear in results from commercial search engines such as Google's; it contextualizes them by decentering the dominant lenses through which results about Black women and girls are interpreted. By

doing this, I am purposefully theorizing from a feminist perspective, while addressing often-overlooked aspects of race in feminist theories of technology. The professor emeritus of science and technology at UCLA Sandra Harding suggests that there is value in identifying a feminist method and epistemology:

Feminist challenges reveal that the questions that are asked—and, even more significantly, those that are not asked—are at least as determinative of the adequacy of our total picture as are any answers that we can discover. Defining what is in need of scientific explanation only from the perspective of bourgeois, white men's experiences leads to partial and even perverse understandings of social life. One distinctive feature of feminist research is that it generates problematics from the perspective of women's experiences.¹⁶

Rather than assert that problematic or racist results are impossible to correct, in the ways that the Google disclaimer suggests,¹⁷ I believe a feminist lens, coupled with racial awareness about the intersectional aspects of identity, offers new ground and interpretations for understanding the implications of such problematic positions about the benign instrumentality of technologies. Black feminist ways of knowing, for example, can look at searches on terms such as “black girls” and bring into the foreground evidence about the historical tendencies to misrepresent Black women in the media. Of course, these misrepresentations and the use of big data to maintain and exacerbate social relationships serve a powerful role in maintaining racial and gender subjugation. It is the persistent normalization of Black people as aberrant and undeserving of human rights and dignity under the banners of public safety, technological innovation, and the emerging creative economy that I am directly challenging by showing the egregious ways that dehumanization is rendered a legitimate free-market technology project.

I am building on the work of previous scholars of commercial search engines such as Google but am asking new questions that are informed by a Black feminist lens concerned with social justice for people who are systematically oppressed. I keep my eye on complicating the notion that information assumed to be “fact” (by virtue of its legitimization at the top of the information pile) exists because racism and sexism are profitable

under our system of racialized capitalism. The ranking hierarchy that the public embraces reflects our social values that place a premium on being number one, and search-result rankings live in this de facto system of authority. Where other scholars have problematized Google Search in terms of its lack of neutrality and prioritization of its own commercial interests, my critiques aim to explicitly address racist and sexist bias in search, fueled by neoliberal technology policy over the past thirty years.

Black Feminism as Theoretical and Methodological Approach

The commodified online status of Black women's and girls' bodies deserves scholarly attention because, in this case, their bodies are defined by a technological system that does not take into account the broader social, political, and historical significance of racist and sexist representations. The very presence of Black women and girls in search results is misunderstood and clouded by dominant narratives of the authenticity and lack of bias of search engines. In essence, the social context or meaning of derogatory or problematic Black women's representations in Google's ranking is normalized by virtue of their placement, making it easier for some people to believe that what exists on the page is strictly the result of the fact that more people are looking for Black women in pornography than anything else. This is because the public believes that what rises to the top in search is either the most popular or the most credible or both.

Yet this does not explain why the word "porn" does not have to be included in keyword searches on "black girls" and other girls and women of color to bring it to the surface as the primary data point about girls and women. The political and social meaning of such output is stripped away when Black girls are explicitly sexualized in search rankings without any explanation, particularly without the addition of the words "porn" or "sex" to the keywords. This phenomenon, I argue, is replicated from offline social relations and deeply embedded in the materiality of technological output; in other words, traditional misrepresentations in old media are made real once again online and situated in an authoritative mechanism that is trusted by the public: Google. The study of Google searches as an Internet artifact is telling. Black feminist scholars have already articulated the harm of such media misrepresentations:¹⁸

gender, class, power, sexuality, and other socially constructed categories interact with one another in a matrix of social relations that create conditions of inequality or oppression.

Black feminist thought offers a useful and antiessentializing lens for understanding how both race and gender are socially constructed and mutually constituted through historical, social, political, and economic processes,¹⁹ creating interesting research questions and new analytical possibilities. As a theoretical approach, it challenges the dominant research on race and gender, which tends to universalize problems assigned to race or Blackness as "male" (or the problems of men) and organizes gender as primarily conceived through the lenses and experiences of White women, leaving Black women in a precarious and understudied position. Popular culture provides countless examples of Black female appropriation and exploitation of negative stereotypes either to assert control over the representation or at least to reap the benefits of it. The Black feminist scholar bell hooks has written extensively on the ways that neoliberal capitalism is explicitly implicated in misrepresentations and hypersexualization of Black women. hooks's work is a mandate for Black women interested in theorizing in the new media landscape, and I use it as both inspiration and a call to action for other Black women interested in engaging in critical information studies. In total, this research is informed by a host of scholars who have helped me make sense of the ways that technology ecosystems—from traditional classification systems such as library databases to new media technologies such as commercial search engines—are structuring narratives about Black women and girls. In the cases I present, I demonstrate how commercial search engines such as Google not only mediate but are mediated by a series of profit-driven imperatives that are supported by information and economic policies that underwrite the commodification of women's identities. Ultimately, this book is designed to "make it plain," as we say in the Black community, just exactly how it can be that Black women and girls continue to have their image and representations assaulted in the new media environments that are not so unfamiliar or dissimilar to old, traditional media depictions. I intend to meaningfully articulate the ways that commercialization is the source of power that drives the consumption of Black women's and girls' representative identity on the web.

While primarily offering reflection on the effects of search-engine-prioritized content, this research is at the same time intended to bring about a deeper inquiry and a series of strategies that can inform public-policy initiatives focused on connecting Black people to the Internet, in spite of the research that shows that cultural barriers, norms, and power relations alienate Black people from the web.²⁰ After just over a decade of focus on closing the digital divide,²¹ the research questions raised here are meant to provoke a discussion about “what then?” What does it mean to have every Black woman, girl, man, and boy in the United States connected to the web if the majority of them are using a search engine such as Google to access content—whether about themselves or other things—only to find results like those with which I began this introduction? The race to digitize cultural heritage and knowledge is important, but it is often mediated by a search engine for the user who does not know precisely how to find it, much the way a library patron is reliant on deep knowledge and skills of the reference librarian to navigate the vast volumes of information in the library stacks.

The Importance of Google

Google has become a ubiquitous entity that is synonymous for many everyday users with “the Internet” itself. From serving as a browser of the Internet to handling personal email or establishing Wi-Fi networks and broadband projects in municipalities across the United States, Google, unlike traditional telecommunications companies, has unprecedented access to the collection and provision of data across a variety of platforms in a highly unregulated marketplace and policy environment. We must continue to study the implications of engagement with commercial entities such as Google and what makes them so desirable to consumers, as their use is not without consequences of increased surveillance and privacy invasions and participation in hidden labor practices. Each of these enhances the business model of Google’s parent company, Alphabet, and reinforces its market dominance across a host of vertical and horizontal markets.²² In 2011, the Federal Trade Commission started looking into Google’s near-monopoly status and market dominance and the harm this could cause consumers. By March 16, 2012, Google was trading on NASDAQ at \$625.04 a share, with a market capitalization of

just over \$203 billion. At the time of the hearings, Google’s latest income statement, for December 2011, showed gross profit at \$24.7 billion. It had \$43.3 billion cash on hand and just \$6.21 billion in debt. Google held 66.2% of the search engine market industry in 2012. Google Search’s profits have only continued to grow, and its holdings have become so significant that the larger company has renamed itself Alphabet, with Google Search as but one of many holdings. By the final writing of this book in August 2017, Alphabet was trading at \$936.38 on NASDAQ, with a market capitalization of \$649.49 billion.

The public is aware of the role of search in everyday life, and people’s opinions on search are alarming. Recent data from tracking surveys and consumer-behavior trends by the comScore Media Metrix consumer panel conducted by the Pew Internet and American Life Project show that search engines are as important to Internet users as email is. Over sixty million Americans engage in search, and for the most part, people report that they are satisfied with the results they find in search engines. The 2005 and 2012 Pew reports on “search engine use” reveal that 73% of all Americans have used a search engine, and 59% report using a search engine every day.²³ In 2012, 83% of search engine users used Google. But Google Search prioritizes its own interests, and this is something far less visible to the public. Most people surveyed could not tell the difference between paid advertising and “genuine” results.

If search is so trusted, then why is a study such as this one needed? The exploration beyond that first simple search is the substance of this book. Throughout the discussion of these and other results, I want to emphasize the main point: there is a missing social context in commercial digital media platforms, and it matters, particularly for marginalized groups that are problematically represented in stereotypical or pornographic ways, for those who are bullied, and for those who are consistently targeted. I use only a handful of illustrative searches to underscore the point and to raise awareness—and hopefully intervention—of how important what we find on the web through commercial search engines is to society.

Search Results as Power

Search results reflect the values and norms of the search company’s commercial partners and advertisers and often reflect our lowest and

most demeaning beliefs, because these ideas circulate so freely and so often that they are normalized and extremely profitable. Search results are more than simply what is popular. The dominant notion of search results as being both “objective” and “popular” makes it seem as if misogynist or racist search results are a simple mirror of the collective. Not only do problematic search results seem “normal,” but they seem completely unavoidable as well, even though these ideas have been thoroughly debunked by scholars. Unfortunately, users of Google give consent to the algorithms’ results through their continued use of the product, which is largely unavoidable as schools, universities, and libraries integrate Google products into our educational experiences.²⁴

Google’s monopoly status,²⁵ coupled with its algorithmic practices of biasing information toward the interests of the neoliberal capital and social elites in the United States, has resulted in a provision of information that purports to be credible but is actually a reflection of advertising interests. Stated another way, it can be argued that Google functions in the interests of its most influential paid advertisers or through an intersection of popular and commercial interests. Yet Google’s users think of it as a public resource, generally free from commercial interest. Further complicating the ability to contextualize Google’s results is the power of its social hegemony.²⁶ Google benefits directly and materially from what can be called the “laborainment”²⁷ of users, when users consent to freely give away their labor and personal data for the use of Google and its products, resulting in incredible profit for the company.

There are many cases that could be made to show how overreliance on commercial search by the public, including librarians, information professionals, and knowledge managers—all of whom are susceptible to overuse of or even replacement by search engines—is something that we must pay closer attention to right now. Under the current algorithmic constraints or limitations, commercial search does not provide appropriate social, historical, and contextual meaning to already overracialized and hypersexualized people who materially suffer along multiple axes. In the research presented in this study, the reader will find a more meaningful understanding of the kind of harm that such limitations can cause for users reliant on the web as an artifact of both formal and informal culture.²⁸ In sum, search results play a powerful role in providing fact and authority to those who see them, and as such, they must

be examined carefully. Google has become a central object of study for digital media scholars,²⁹ due to recognition on these scholars’ parts of the power and impact wielded by the necessity to begin most engagements with social media via a search process and the near universality with which Google has been adopted and embedded into all aspects of the digital media landscape to respond to that need. This work is addressing a gap in scholarship on how search works and what it biases, public trust in search, the relationship of search to information studies, and the ways in which African Americans, among others, are mediated and commodified in Google.

To start revealing some of the processes involved, it is important to think about how results appear. Although one might believe that a query to a search engine will produce the most relevant and therefore useful information, it is actually predicated on a matrix of ways in which pages are hyperlinked and indexed on the web.³⁰ Rendering web content (pages) findable via search engines is an expressly social, economic, and human project, which several scholars have detailed. These renderings are delivered to users through a set of steps (algorithms) implemented by programming code and then naturalized as “objective.” One of the reasons this is seen as a neutral process is because algorithmic, scientific, and mathematical solutions are evaluated through procedural and mechanistic practices, which in this case includes tracing hyperlinks among pages. This process is defined by Google’s founders, Sergey Brin and Larry Page, as “voting,” which is the term they use to describe how search results move up or down in a ranked list of websites. For the most part, many of these processes have been automated, or they happen through graphical user interfaces (GUIs) that allow people who are not programmers (i.e., not working at the level of code) to engage in sharing links to and from websites.³¹

Research shows that users typically use very few search terms when seeking information in a search engine and rarely use advanced search queries, as most queries are different from traditional offline information-seeking behavior.³² This front-end behavior of users appears to be simplistic; however, the information retrieval systems are complex, and the formulation of users’ queries involves cognitive and emotional processes that are not necessarily reflected in the system design.³³ In essence, while users use the simplest queries they can in a

search box because of the way interfaces are designed, this does not always reflect how search terms are mapped against more complex thought patterns and concepts that users have about a topic. This disjunction between, on the one hand, users' queries and their real questions and, on the other, information retrieval systems makes understanding the complex linkages between the content of the results that appear in a search and their import as expressions of power and social relations of critical importance.

The public generally trusts information found in search engines. Yet much of the content surfaced in a web search in a commercial search engine is linked to paid advertising, which in part helps drive it to the top of the page rank, and searchers are not typically clear about the distinctions between "real" information and advertising. Given that advertising is a fundamental part of commercial search, using content analysis to make sense of what *actually* is served up in search is appropriate and consistent with the articulation of feminist critiques of the images of women in print advertising.³⁴ These scholars have shown the problematic ways that women have been represented—as sex objects, incompetent, dependent on men, or underrepresented in the workforce³⁵—and the content and representation of women and girls in search engines is consistent with the kinds of problematic and biased ideas that live in other advertising channels. Of course, this makes sense, because Google Search is in fact an advertising platform, not intended to solely serve as a public information resource in the way that, say, a library might. Google creates advertising algorithms, not information algorithms.

To understand search in the context of this book, it is important to look at the description of the development of Google outlined by the former Stanford computer science graduate students and cofounders of the company, Sergey Brin and Larry Page, in "The Anatomy of a Large-Scale Hypertextual Web Search Engine." Their paper, written in graduate school, serves as the architectural framework for Google's PageRank. In addition, it is crucial to also look at the way that citation analysis, the foundational notion behind Brin and Page's idea, works as a bibliometric project that has been extensively developed by library and information science scholars. Both of these dynamics are often misunderstood because they do not account for the complexities of human intervention involved in vetting of information, nor do they pay attention

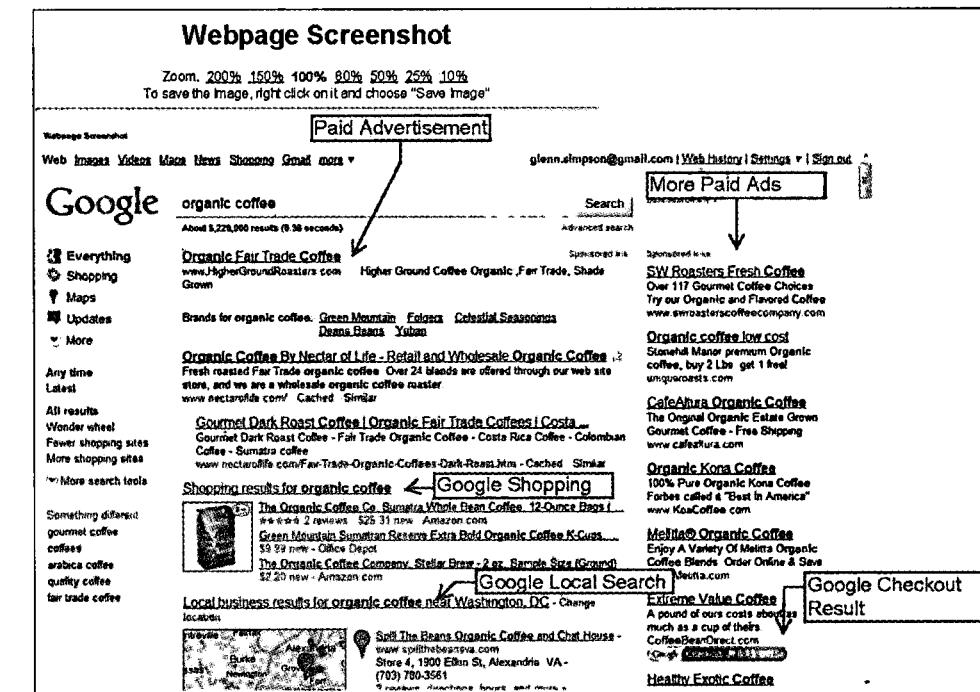


Figure 1.11. Example of Google's prioritization of its own properties in web search. Source: Inside Google (2010).

to the relative weight or importance of certain types of information.³⁶ For example, in the process of citing work in a publication, all citations are given equal weight in the bibliography, although their relative importance to the development of thought may not be equal at all. Additionally, no relative weight is given to whether a reference is validated, rejected, employed, or engaged—complicating the ability to know what a citation actually *means* in a document. Authors who have become so mainstream as not to be cited, such as not attributing modern discussions of class or power dynamics to Karl Marx or the notion of "the individual" to the scholar of the Italian Renaissance Jacob Burckhardt, mean that these intellectual contributions may undergird the framework of an argument but move through works without being cited any longer. Concepts that may be widely understood and accepted ways of knowing are rarely cited in mainstream scholarship, an important dynamic that

Linda Smith, former president of the Association for Information Science and Technology (ASIS&T) and associate dean of the Information School at the University of Illinois at Urbana-Champaign, argues is part of the flawed system of citation analysis that deserves greater attention if bibliometrics are to serve as a legitimating force for valuing knowledge production.

Brin and Page saw the value in using works that others cite as a model for thinking about determining what is legitimate on the web, or at least to indicate what is popular based on many people acknowledging particular types of content. In terms of outright co-optation of the citation, vis-à-vis the hyperlink, Brin and Page were aware of some of the challenges I have described. They were clearly aware from the beginning of the potential for “gaming” the system by advertising companies or commercial interests, a legitimated process now known as “search engine optimization,” to drive ads or sites to the top of a results list for a query, since clicks on web links can be profitable, as are purchases gained by being vetted as “the best” by virtue of placement on the first page of PageRank. This is a process used for web results, not paid advertising, which is often highlighted in yellow (see figure 1.6). Results that appear not to be advertising are in fact influenced by the advertising algorithm. In contrast to scientific or scholarly citations, which once in print are persistent and static, hyperlinking is a dynamic process that can change from moment to moment.³⁷ As a result, the stability of results in Google ranking shifts and is prone to being affected by a number of processes that I will cover, primarily search engine optimization and advertising. This means that results shift over time. The results of what is most hyperlinked using Google’s algorithm today will be different at a later date or from the time that Google’s web-indexing crawlers move through the web until the next cycle.³⁸

Citation importance is a foundational concept for determining scholarly relevance in certain disciplines, and citation analysis has largely been considered a mechanism for determining whether a given article or scholarly work is important to the scholarly community. I want to revisit this concept because it also has implications for thinking about the legitimization of information, not just citability or popularity. It is also a function of human beings who are engaged in a curation practice, not entirely left to automation. Simply put, if scholars choose to

cite a study or document, they have signaled its relevance; thus, human beings (scholars) are involved in making decisions about a document’s relevance, although all citations in a bibliography do not share the same level of meaningfulness. Building on this concept of credibility through citation, PageRank is what Brin and Page call the greater likelihood that a document is relevant “if there are many pages that point to it” versus “the probability that the random surfer visits a page.”³⁹ In their research, which led to the development of Google Search, Brin and Page discuss the possibility of monopolizing and manipulating keywords through commercialization of the web search process. Their information-retrieval goal was to deliver the most relevant or very best ten or so documents out of the possible number of documents that could be returned from the web. The resulting development of their search architecture is PageRank—a system that is based on “the objective measure of its citation importance that corresponds well with people’s subjective idea of importance.”⁴⁰

One of the most profound parts of Brin and Page’s work is in appendix A, in which they acknowledge the ways that commercial interests can compromise the quality of search result retrieval. They state, citing Ben Bagdikian, “It is clear that a search engine which was taking money for showing cellular phone ads would have difficulty justifying the page that our system returned to its paying advertisers. For this type of reason and historical experience with other media, we expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers.”⁴¹ Brin and Page outline a clear roadmap for how bias would work in advertising-oriented search and the effects this would have, and they directly suggest that it is in the consumer’s interest not to have search compromised by advertising and commercialism. To some degree, PageRank was intended to be a measure of relevance based on popularity—including what both web surfers and web designers link to from their sites. As with academic citations, Brin and Page decided that citation analysis could be used as a model for determining whether web links could be ranked according to their importance by measuring how much they were back-linked or hyperlinked to or from. Thus, the model for web indexing pages was born. However, in the case of citation analysis, a scholarly author goes through several stages of vetting and credibility testing, such as the peer-review process,

before work can be published and cited. In the case of the web, such credibility checking is not a factor in determining what will be hyper-linked. This was made explicitly clear in the many news reports covering the 2016 U.S. presidential election, where clickbait and manufactured “news” from all over the world clouded accurate reporting of facts on the presidential candidates.

Another example of the shortcomings of removing this human curation or decision making from the first page of results at the top of PageRank, in addition to the results that I found for “black girls,” can be found in the more public dispute over the results that were returned on searches for the word “Jew,” which included a significant number of anti-Semitic pages. As can be seen by Google’s response to the results of a keyword search for “Jew,” Google takes little responsibility toward the ways that it provides information on racial and gendered identities, which are curated in more meaningful ways in scholarly databases. Siva Vaidhyanathan’s 2011 book *The Googlization of Everything (And Why We Should Worry)* chronicles recent attempts by the Jewish community and Anti-Defamation League to challenge Google’s priority ranking to the first page of anti-Semitic, Holocaust-denial websites. So troublesome were these search results that in 2011, Google issued a statement about its search process, encouraging people to use “Jews” and “Jewish people” in their searches, rather than the seemingly pejorative term “Jew”—claiming that the company can do nothing about the word’s co-optation by White supremacist groups (see figure 1.12).

Google, according to its own disclaimer, will only remove pages that are considered unlawful, as is the case in France and Germany, where selling or distributing neo-Nazi materials is prohibited. Without such limits on derogatory, racist, sexist, or homophobic materials, Google allows its algorithm—which is, as we can see, laden with what Diaz calls “sociopolitics”—to stand without debate while protesting its inability to remove pages. As recently as June 27, 2012, Google settled a claim by the French antiracism organization the International League Against Racism over Google’s use of ethnic identity—“Jew”—in association with popular searches.⁴² Under French law, racial identity markers cannot be stored in databases, and the auto-complete techniques used in the Google search box link names of people to the word “Jew” on the basis of past user searches. What this recent case points to is another effort to

Google

An explanation of our search results

If you recently used Google to search for the word “Jew,” you may have seen results that were very disturbing. We assure you that the views expressed by the sites in your results are not in any way endorsed by Google. We’d like to explain why you’re seeing these results when you conduct this search.

A site’s ranking in Google’s search results relies heavily on computer algorithms using thousands of factors to calculate a page’s relevance to a given query. Sometimes subtleties of language cause anomalies to appear that cannot be predicted. A search for “Jew” brings up one such unexpected result.

If you use Google to search for “Judaism,” “Jewish” or “Jewish people,” the results are informative and relevant. So why is a search for “Jew” different? One reason is that the word “Jew” is often used in an anti-Semitic context. Jewish organizations are more likely to use the word “Jewish” when talking about members of their faith. The word has become somewhat charged linguistically, as noted on websites devoted to Jewish topics such as these:

- <http://www.jewishworldreview.com/columns/jonah081500.asp>

Someone searching for information on Jewish people would be more likely to enter terms like “Judaism,” “Jewish people,” or “Jews” than the single word “Jew.” In fact, prior to this incident, the word “Jew” only appeared about once in every 10 million search queries. Now it’s likely that the great majority of searches on Google for “Jew” are by people who have heard about this issue and want to see the results for themselves.

The beliefs and preferences of those who work at Google, as well as the opinions of the general public, do not determine or impact our search results. Individual citizens and public interest groups do periodically urge us to remove particular links or otherwise adjust search results. Although Google reserves the right to address such requests individually, Google views the comprehensiveness of our search results as an extremely important priority. Accordingly, we do not remove a page from our search results simply because its content is unpopular or because we receive complaints concerning it. We will, however, remove pages from our results if we believe the page (or its site) violates our Webmaster Guidelines, if we believe we are required to do so by law, or at the request of the webmaster who is responsible for the page.

We apologize for the upsetting nature of the experience you had using Google and appreciate your taking the time to inform us about it.

Sincerely,
The Google Team

P.S. You may be interested in some additional information the Anti-Defamation League has posted about this issue at http://www.adl.org/rumors/google_search_rumors.asp. In addition, we call your attention to Google’s search results on this topic.

©2011 Google

Figure 1.12. Explanation of results by Google. Source: www.google.com/explanation.html (originally available in 2005).

redefine distorted images of people in new media. These cases of distortion, however, continue to accumulate.

The public's as well as the Jewish community's interest in accurate information about Jewish culture and the Holocaust should be enough motivation to provoke a national discussion about consumer harm, to which my research shows we can add other cultural and gender-based identities that are misrepresented in search engines. However, Google's assertion that its search results, though problematic, were computer generated (and thus not the company's fault) was apparently a good-enough answer for the Anti-Defamation League (ADL), which declared, "We are extremely pleased that Google has heard our concerns and those of its users about the offensive nature of some search results and the unusually high ranking of peddlers of bigotry and anti-Semitism."⁴³ The ADL does acknowledge on its website its gratitude to Sergey Brin, cofounder of Google and son of Russian Jewish immigrants, for his personal letter to the organization and his mea culpa for the "Jew" search-term debacle. The ADL generously stated in its press release about the incident that Google, as a resource to the public, should be forgiven because "until the technical modifications are implemented, Google has placed text on its site that gives users a clear explanation of how search results are obtained. Google searches are automatically determined using computer algorithms that take into account thousands of factors to calculate a page's relevance."⁴⁴

If there is a technical fix, then what are the constraints that Google is facing such that eight years later, the issue has yet to be resolved? A search for the word "Jew" in 2012 produces a beige box at the bottom of the results page from Google linking to its lengthy disclaimer about the results—which remain a mix of both anti-Semitic and informative sites (see figure 1.13). That Google places the responsibility for bad results back on the shoulders of information searchers is a problem, since most of the results that the public gets on broad or open-ended racial and gendered searches are out of their control and entirely within the control of Google Search.

It is important to note that Google has conceded the fact that anti-Semitism as the primary information result about Jewish people is a problem, despite its disclaimer that tries to put the onus for bad results on the searcher. In Germany and France, for example, it is illegal to sell

Ad - Why this ad?

Offensive Search Results
www.google.com/explanation
 We're disturbed about these results as well. Please read our note here.

Searches related to Jew

jew jokes	jew watch
jew definition	jew urban dictionary
Jewish jokes	jew pictures
famous Jews	jew beard

Goooooooooooooogle >

1 2 3 4 5 6 7 8 9 10 [Next](#)

[Advanced search](#) [Search Help](#) [Give us feedback](#)

[Google Home](#) [Advertising Programs](#) [Business Solutions](#) [Privacy & Terms](#)
[About Google](#)

Figure 1.13. Google's bottom-of-the-page beige box regarding offensive results, which previously took users to "An Explanation of Our Search Results." Source: www.google.com/explanation (no longer available).

Nazi memorabilia, and Google has had to put in place filters that ensure online retailers of such are not visible in search results. In 2002, Benjamin Edelman and Jonathan Zittrain at Harvard University's Berkman Center for Internet and Society concluded that Google was filtering its search results in accordance with local law and precluding neo-Nazi organizations and content from being displayed.⁴⁵ While this indicates that Google can in fact remove objectionable hits, it is equally troubling, because the company provided search results without informing searchers that information was being deleted. That is to say that the results were presented as factual and complete without mention of omission. Yahoo!, another leading U.S. search engine, was forced into a protracted legal battle in France for allowing pro-Nazi memorabilia to be sold through its search engine, in violation of French law. What these cases point to is that search results are deeply contextual and easily manipulated, rather than objective, consistent, and transparent, and that they can be legitimated only in social, political, and historical context.

The issue of unlawfulness over the harm caused by derogatory results is a question of considerable debate. For example, in the United States, where free speech protections are afforded to all kinds of speech, including hate speech and racist or sexist depictions of people and communities, there is a higher standard of proof required to show harm toward disenfranchised or oppressed people. We need legal protections now more than ever, as automated decision-making systems wield greater power in society.

Gaming the System: Optimizing and Co-opting Results in Search Engines

Google's advertising tool or optimization product is AdWords. AdWords allows anyone to advertise on Google's search pages and is highly customizable. With this tool, an advertiser can set a maximum amount of money that it wants to spend on a daily basis for advertising. The model for AdWords is that Google will display ads on search pages that it believes are relevant to the kind of search query that is taking place by a user. If a user clicks on an ad, then the advertiser pays. And Google incentivizes advertisers by suggesting that their ads will show up in searches and display, but the advertiser (or Google customer) pays for the ad only when a user (Google consumer) clicks on the advertisement, which is the cost per click (CPC). The advertiser selects a series of "keywords" that it believes closely align with its product or service that it is advertising, and a customer can use a Keyword Estimator tool in order to see how much the keywords they choose to associate with their site might cost. This advertising mechanism is an essential part of how PageRank prioritizes ads on a page, and the association of certain keywords with particular industries, products, and services derives from this process, which works in tandem with PageRank.

In order to make sense of the specific results in keyword searches, it is important to know how Google's PageRank works, what commercial processes are involved in PageRank, how search engine optimization (SEO) companies have been developed to influence the process of moving up results,⁴⁶ and how Google bombing⁴⁷ occurs on occasion. Google bombing is the practice of excessively hyperlinking to a website (repeatedly coding HTML to link a page to a term or phrase) to cause it to

rise to the top of PageRank, but it is also seen as a type of "hit and run" activity that can deliberately co-opt terms and identities on the web for political, ideological, and satirical purposes. Judit Bar-Ilan, a professor of information science at Bar-Ilan University, has studied this practice to see if the effect of forcing results to the top of PageRank has a lasting effect on the result's persistence, which can happen in well-orchestrated campaigns. In essence, Google bombing is the process of co-opting content or a term and redirecting it to unrelated content. Internet lore attributes the creation of the term "Google bombing" to Adam Mathes, who associated the term "talentless hack" with a friend's website in 2001. Practices such as Google bombing (also known as Google washing) are impacting both SEO companies and Google alike. While Google is invested in maintaining the quality of search results in PageRank and policing companies that attempt to "game the system," as Brin and Page foreshadowed, SEO companies do not want to lose ground in pushing their clients or their brands up in PageRank.⁴⁸ SEO is the process of "using a range of techniques, including augmenting HTML code, web page copy editing, site navigation, linking campaigns and more, in order to improve how well a site or page gets listed in search engines for particular search topics,"⁴⁹ in contrast to "paid search," in which the company pays Google for its ads to be displayed when specific terms are searched. A media spectacle of this nature is the case of Senator Rick Santorum, Republican of Pennsylvania, whose website and name were associated with insults in order to drive objectionable content to the top of PageRank.⁵⁰ Others who have experienced this kind of co-optation of identity or less-than-desirable association of their name with an insult include former president George W. Bush and the pop singer Justin Bieber.

All of these practices of search engine optimization and Google bombing can take place independently of and in concert with the process of crawling and indexing the web. In fact, being found gives meaning to a website and creates the conditions in which a ranking can happen. Search engine optimization is a major factor in findability on the web. What is important to note is that search engine optimization is a multibillion-dollar industry that impacts the value of specific keywords; that is, marketers are invested in using particular keywords, and keyword combinations, to optimize their rankings.

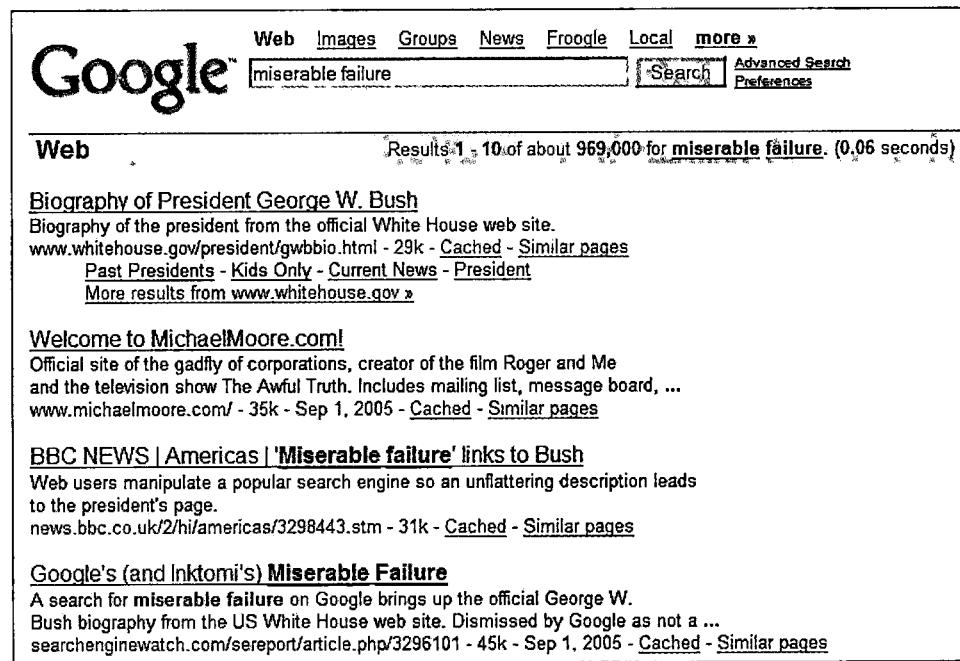


Figure 1.14. Example of a Google bomb on George W. Bush and the search terms “miserable failure,” 2005.

Despite the widespread beliefs in the Internet as a democratic space where people have the power to dynamically participate as equals, the Internet is in fact organized to the benefit of powerful elites,⁵¹ including corporations that can afford to purchase and redirect searches to their own sites. What is most popular on the Internet is not wholly a matter of what users click on and how websites are hyperlinked—there are a variety of processes at play. Max Holloway of *Search Engine Watch* notes, “Similarly, with Google, when you click on a result—or, for that matter, don’t click on a result—that behavior impacts future results. One consequence of this complexity is difficulty in explaining system behavior. We primarily rely on performance metrics to quantify the success or failure of retrieval results, or to tell us which variations of a system work better than others. Such metrics allow the system to be continuously improved upon.”⁵² The goal of combining search terms, then, in the context of the landscape of the search engine optimization logic, is only the beginning.

Much research has now been done to dispel the notion that users of the Internet have the ability to “vote” with their clicks and express interest in individual content and information, resulting in democratic practices online.⁵³ Research shows the ways that political news and information in the blogosphere are mediated and directed such that major news outlets surface to the top of the information pile over less well-known websites and alternative news sites in the blogosphere, to the benefit of elites.⁵⁴ In the case of political information seeking, research has shown how Google directs web traffic to mainstream corporate news conglomerates, which increases their ability to shape the political discourse. Google too is a mediating platform that, at least at one moment in time, in September 2011, allowed the porn industry to take precedence in the representations of Black women and girls over other possibilities among at least eleven and a half billion documents that could have been indexed.⁵⁵ That moment in 2011 is, however, emblematic of Google’s ongoing dynamic. It has since produced many more problematic results.

As the Federal Communications Commission declares broadband “the new common medium,”⁵⁶ the role of search engines is taking on even greater importance to “the widest possible dissemination of information from diverse and antagonistic sources . . . essential to the welfare of the public.”⁵⁷ This political economy of search engines and traditional advertisers includes search engine optimization companies that operate in a secondary or gray market (often in opposition to Google). Ultimately, the results we get are about the financial interest that Google or SEOs have in helping their own clients optimize their rankings. In fact, Google is in the business of selling optimization. Extensive critiques of Google have been written on the political economy of search⁵⁸ and the way that consolidations in the search engine industry market contribute to the erosion of public resources, in much the way that the media scholars Robert McChesney, former host of nationally syndicated radio show *Media Matters*, and John Nichols, a writer for the *Nation*, critique the consolidation of the mass-media news markets. Others have spoken to the inherent democratizing effect of search engines, such that search is adding to the diversity of political organization and discourse because the public is able to access more information in the marketplace of ideas.⁵⁹ Mounting evidence shows that automated decision-making systems are disproportionately harmful to the most vulnerable and the

least powerful, who have little ability to intervene in them—from misrepresentation to prison sentencing to accessing credit and other life-impacting formulas.

This landscape of search engines is important to consider in understanding the meaning of search for the public, and it serves as a basis for examining why information quality online is significant. We must trouble the notion of Google as a public resource, particularly as institutions become more reliant on Google when looking for high-quality, contextualized, and credible information. This shift from public institutions such as libraries and schools as brokers of information to the private sector, in projects such as Google Books, for example, is placing previously public assets in the hands of a multinational corporation for private exploitation. Information is a new commodity, and search engines can function as private information enclosures.⁶⁰ We need to make more visible the commercial interests that overdetermine what we can find online.

The Enclosure of the Public Domain through Search Engines

At the same time that search engines have become the dominant portal for information seeking by U.S. Internet users, the rise of commercial mediation of information in those same search engines is further enclosing the public domain. Decreases in funding for public information institutions such as libraries and educational institutions and shifts of responsibility to individuals and the private sector have reframed the ways that the public conceives of what can and should be in the public domain. Yet Google Search is conceived of as a public resource, even though it is a multinational advertising company. These shifts of resources that were once considered public have been impacted by increased intellectual property rights, licensing, and publishing agreements for companies and private individuals in the domain of copyrights, patents, and other legal protections. The move of community-based assets and culture to private hands is arguably a crisis that has rolled back the common good, but there are still possible strategies that can be explored for maintaining what can remain in the public domain. Commercial control over the Internet, often considered a “commons,” has moved it further away from the public through a series of national and international regulations and intellectual and commercial borders that exist in the management of the

network.⁶¹ Beyond the Internet and the control of the network, public information—whether delivered over the web or not—continues to be outsourced to the private sphere, eroding the public information commons that has been a basic tenet of U.S. democracy.

The critical media scholar Herbert Schiller, whose work foreshadowed many of the current challenges in the information and communications landscape, provides a detailed examination of the impact of outsourcing and deregulation in the spheres of communication and public information. His words are still timely: “The practice of selling government (or any) information serves the corporate user well. Ordinarily individual users go to the end of the dissemination queue. Profoundly antidemocratic in its effect, privatizing and/or selling information, which at one time was considered public property, has become a standard practice in recent years.”⁶² What this critique shows is that the privatization and commercial nature of information has become so normalized that it not only becomes obscured from view but, as a result, is increasingly difficult to critique within the public domain. The Pew Internet and American Life Project corroborates that the public trusts multinational corporations that provide information over the Internet and that there is a low degree of distrust of the privatization of information.⁶³ Part of this process of acquiescence to the increased corporatization of public life can be explained by the economic landscape, which is shaped by military-industrial projects such as the Internet that have emerged in the United States,⁶⁴ increasing the challenge of scholars who are researching the impact of such shifts in resources and accountability. Molly Niesen at the University of Illinois has written extensively on the loss of public accountability by federal agencies such as the Federal Trade Commission (FTC), which is a major contribution to our understanding of where the public can focus attention on policy interventions.⁶⁵ We should leverage her research to think about the FTC as the key agency to manage and intervene in how corporations control the information landscape.

The Cultural Power of Algorithms

The public is minimally aware of these shifts in the cultural power and import of algorithms. In a 2015 study by the Pew Research Center,



Figure 1.15. *Forbes*'s online reporting (and critique) of the Epstein and Robertson study.

"American's Privacy Strategies Post-Snowden," only 34% of respondents who were aware of the surveillance that happens automatically online through media platforms, such as search behavior, email use, and social media, reported that they were shifting their online behavior because of concerns of government surveillance and the potential implications or harm that could come to them.⁶⁶ Little of the American public knows that online behavior has more importance than ever. Indeed, Internet-based activities are dramatically affecting our notions of how democracy and freedom work, particularly in the realm of the free flow of information and communication. Our ability to engage with the information landscape subtly and pervasively impacts our understanding of the world and each other.

An example of how information flow and bias in the realm of politics have recently come to the fore can be found in an important new study about how information bias can radically alter election outcomes. The former editor of *Psychology Today* and professor Robert Epstein and Ronald Robertson, the associate director of the American Institute for Behavioral Research and Technology, found in their 2013 study that democracy was at risk because manipulating search rankings could shift voters' preferences, substantially and without their awareness. In their study, they note that the tenor of stories about a candidate in search engine results, whether favorable or unfavorable, dramatically af-

fected the way that people voted. Seventy-five percent of participants were not aware that the search results had been manipulated. The researchers concluded, "The outcomes of real elections—especially tight races—can conceivably be determined by the strategic manipulation of search engine rankings and . . . that the manipulation can be accomplished without people being aware of it. We speculate that unregulated search engines could pose a serious threat to the democratic system of government."⁶⁷

In March 2012, the Pew Internet and American Life Project issued an update to its 2005 "Search Engine Users" study. The 2005 and 2012 surveys tracking consumer-behavior trends from the comScore Media Metrix consumer panel show that search engines are as important to Internet users as email is. In fact, the *Search Engine Use 2012* report suggests that the public is "more satisfied than ever with the quality of search results."⁶⁸ Further findings include the following:

- 73% of all Americans have used a search engine, and 59% report using a search engine every day.
- 83% of search engine users use Google.

Especially alarming is the way that search engines are increasingly positioned as a trusted public resource returning reliable and credible information. According to Pew, users report generally good outcomes and relatively high confidence in the capabilities of search engines:

- 73% of search engine users say that most or all the information they find as they use search engines is accurate and trustworthy.

Yet, at the same time that search engine users report high degrees of confidence in their skills and trust in the information they retrieve from engines, they have also reported that they are naïve about how search engines work:

- 62% of search engine users are not aware of the difference between paid and unpaid results; that is, only 38% are aware, and only 8% of search engine users say that they can always tell which results are paid or sponsored and which are not.

- In 2005, 70% of search engine users were fine with the concept of paid or sponsored results, but in 2012, users reported that they are not okay with targeted advertising because they do not like having their online behavior tracked and analyzed.
- In 2005, 45% of search engine users said they would stop using search engines if they thought the engines were not being clear about offering some results for pay.
- In 2005, 64% of those who used engines at least daily said search engines are a fair and unbiased source of information; the percentage increased to 66% in 2012.

Users in the 2012 Pew study also expressed concern about personalization:

- 73% reported that they would *not be okay* with a search engine keeping track of searches and using that information to personalize future search results. Participants reported that they feel this to be an invasion of privacy.

In the context of these concerns, a 2011 study by the researchers Martin Feuz and Matthew Fuller from the Centre for Cultural Studies at the University of London and Felix Stalder from the Zurich University of the Arts found that personalization is not simply a service to users but rather a mechanism for better matching consumers with advertisers and that Google's personalization or aggregation is about actively matching people to groups, that is, categorizing individuals.⁶⁹ In many cases, different users are seeing similar content to each other, but users have little ability to see how the platform is attempting to use prior search history and demographic information to shape their results. Personalization is, to some degree, giving people the results they want on the basis of what Google knows about its users, but it is also generating results for viewers to see what Google Search thinks might be good for advertisers by means of compromises to the basic algorithm. This new wave of interactivity, without a doubt, is on the minds of both users and search engine optimizing companies and agencies. Google applications such as Gmail or Google Docs and social media sites such as Facebook track identity and previous searches in order to surface targeted ads for users by analyzing users' web traces. So not only do search engines increasingly remember the digital traces of where we have been

and what links we have clicked in order to provide more custom content (a practice that has begun to gather more public attention after Google announced it would use past search practices and link them to users in its privacy policy change in 2012),⁷⁰ but search results will also vary depending on whether filters to screen out porn are enabled on computers.⁷¹

It is certain that information that surfaces to the top of the search pile is not exactly the same for every user in every location, and a variety of commercial advertising, political, social, and economic decisions are linked to the way search results are coded and displayed. At the same time, results are generally quite similar, and complete search personalization—customized to very specific identities, wants, and desires—has yet to be developed. For now, this level of personal-identity personalization has less impact on the variation in results than is generally believed by the public.

Losing Control of Our Images and Ourselves in Search

It is well known that traditional media have been rife with negative or stereotypical images of African American / Black people,⁷² and the web as the locus of new media is a place where traditional media interests are replicated. Those who have been inappropriately and unfairly represented in racist and sexist ways in old media have been able to cogently critique those representations and demand expanded representations, protest stereotypes, and call for greater participation in the production of alternative, nonstereotypical or oppressive representations. This is part of the social charge of civil rights organizations such as the Urban League⁷³ and the National Association for the Advancement of Colored People, which monitor and report on minority misrepresentations, as well as celebrate positive portrayals of African Americans in the media.⁷⁴ At a policy level, some civil rights organizations and researchers such as Darnell Hunt, dean of the division of social science and department chair of sociology at UCLA,⁷⁵ have been concerned with media representations of African Americans, and mainstream organizations such as Free Press have been active in providing resources about the impact of the lack of diversity, stereotyping, and hate speech in the media. Indeed, some of these resources have been directed toward net-neutrality issues

and closing the digital divide.⁷⁶ Media advocacy groups that focus on the pornification of women or the stereotyping of people of color might turn their attention toward the Internet as another consolidated media resource, particularly given the evidence showing Google's information and advertising monopoly status on the web.

Bias in Search

"Traffic Report: How Google Is Squeezing Out Competitors and Muscling Into New Markets," by ConsumerWatchdog.org's Inside Google (June 2010), details how Google effectively blocks sites that it competes with and prioritizes its own properties to the top of the search pile (YouTube over other video sites, Google Maps over MapQuest, and Google Images over Photobucket and Flickr). The report highlights the process by which Universal Search is not a neutral and therefore universal process but rather a commercial one that moves sites that buy paid advertising to the top of the pile. Amid these practices, the media, buttressed by an FTC investigation,⁷⁷ have suggested that algorithms are not at all unethical or harmful because they are free services and Google has the right to run its business in any way it sees fit. Arguably, this is true, so true that the public should be thoroughly informed about the ways that Google biases information—toward largely stereotypic and decontextualized results, at least when it comes to certain groups of people. Commercial platforms such as Facebook and YouTube go to great lengths to monitor uploaded user content by hiring web content screeners, who at their own peril screen illicit content that can potentially harm the public.⁷⁸ The expectation of such filtering suggests that such sites vet content on the Internet on the basis of some objective criteria that indicate that some content is in fact quite harmful to the public. New research conducted by Sarah T. Roberts in the Department of Information Studies at UCLA shows the ways that, in fact, commercial content moderation (CCM, a term she coined) is a very active part of determining what is allowed to surface on Google, Yahoo!, and other commercial text, video, image, and audio engines.⁷⁹ Her work on video content moderation elucidates the ways that commercial digital media platforms currently outsource or in-source image and video content filtering to comply with their terms of use

agreements. What is alarming about Roberts's work is that it reveals the processes by which content is already being screened and assessed according to a continuum of values that largely reflect U.S.-based social norms, and these norms reflect a number of racist and stereotypical ideas that make screening racism and sexism and the abuse of humans in racialized ways "in" and perfectly acceptable, while other ideas such as the abuse of animals (which is also unacceptable) are "out" and screened or blocked from view. She details an interview with one of the commercial content moderators (CCMs) this way:

We have very, very specific itemized internal policies . . . the internal policies are not made public because then it becomes very easy to skirt them to essentially the point of breaking them. So yeah, we had very specific internal policies that we were constantly, we would meet once a week with SecPol to discuss, there was one, blackface is not technically considered hate speech by default. Which always rubbed me the wrong way, so I had probably ten meltdowns about that. When we were having these meetings discussing policy and to be fair to them, they always listened to me, they never shut me up. They didn't agree, and they never changed the policy but they always let me have my say, which was surprising. (Max Breen, MegaTech CCM Worker).

The MegaTech example is an illustration of the fact that social media companies and platforms make active decisions about what kinds of racist, sexist, and hateful imagery and content they will host and to what extent they will host it. These decisions may revolve around issues of "free speech" and "free expression" for the user base, but on commercial social media sites and platforms, these principles are always counterbalanced by a profit motive; if a platform were to become notorious for being too restrictive in the eyes of the majority of its users, it would run the risk of losing participants to offer to its advertisers. So MegaTech erred on the side of allowing more, rather than less, racist content, in spite of the fact that one of its own CCM team members argued vociferously against it and, by his own description, experienced emotional distress ("meltdowns") around it.⁸⁰

This research by Roberts, particularly in the wake of leaked reports from Facebook workers who perform content moderation, suggests that people and policies are put in place to navigate and moderate content on the web. Egregious and racist content, content that is highly profitable, proliferates because many tech platforms are interested in attracting the interests and attention of the majority in the United States, not of racialized minorities.

Challenging Race- and Gender-Neutral Narratives

These explorations of web results on the first page of a Google search also reveal the default identities that are protected on the Internet or are less susceptible to marginalization, pornification, and commodification. The research of Don Heider, the dean of Loyola University Chicago's School of Communication, and Dustin Harp, an assistant professor in the Department of Communication at the University of Texas, Arlington, shows that even though women constitute just slightly over half of Internet users, women's voices and perspectives are not as loud and do not have as much impact online as those of men. Their work demonstrates how some users of the Internet have more agency and can dominate the web, despite the utopian and optimistic view of the web as a socially equalizing and democratic force.⁸¹ Recent research on the male gaze and pornography on the web argue that the Internet is a communications environment that privileges the male, pornographic gaze and marginalizes women as objects.⁸² As with other forms of pornographic representations, pornography both structures and reinforces the domination of women, and the images of women in advertising and art are often "constructed for viewing by a male subject,"⁸³ reminiscent of the journalist and producer John Berger's canonical work *Ways of Seeing*, which describes this objectification in this way: "Women are depicted in a quite different way from men—not because the feminine is different from the masculine—but because the 'ideal' spectator is always assumed to be male and the image of the woman is designed to flatter him."⁸⁴

The previous articulations of the male gaze continue to apply to other forms of advertising and media—particularly on the Internet—and the pornification of women on the web is an expression of racist and sexist hierarchies. When these images are present, White women are the

norm, and Black women are overrepresented, while Latinas are underrepresented.⁸⁵ Tracey A. Gardner characterizes the problematic characterizations of African American women in pornographic media by suggesting that "pornography capitalizes on the underlying historical myths surrounding and oppressing people of color in this country which makes it racist."⁸⁶ These characterizations translate from old media representations to new media forms. Structural inequalities of society are being reproduced on the Internet, and the quest for a race-, gender-, and class-less cyberspace could only "perpetuate and reinforce current systems of domination."⁸⁷

More than fifteen years later, the present research corroborates these concerns. Women, particularly of color, are represented in search queries against the backdrop of a White male gaze that functions as the dominant paradigm on the Internet in the United States. The Black studies and critical Whiteness scholar George Lipsitz, of the University of California, Santa Barbara, highlights the "possessive investment in Whiteness" and the ways that the American construction of Whiteness is more "nonracial" or null. Whiteness is more than a legal abstraction formulated to conceptualize and codify notions of the "Negro," "Black Codes," or the racialization of diverse groups of African peoples under the brutality of slavery—it is an imagined and constructed community uniting ethnically diverse European Americans. Through cultural agreements about who subtly and explicitly constitutes "the other" in traditional media and entertainment such as minstrel shows, racist films and television shows produced in Hollywood, and Wild West narratives, Whiteness consolidated itself "through inscribed appeals to the solidarity of White supremacy."⁸⁸ The cultural practices of our society—which I argue include representations on the Internet—are part of the ways in which race-neutral narratives have increased investments in Whiteness. Lipsitz argues it this way:

As long as we define social life as the sum total of conscious and deliberate individual activities, then only *individual* manifestations of personal prejudice and hostility will be seen as racist. Systemic, collective, and co-ordinated behavior disappears from sight. Collective exercises of group power relentlessly channeling rewards, resources, and opportunities from one group to another will not appear to be "racist" from this perspective

because they rarely announce their intention to discriminate against individuals. But they work to construct racial identities by giving people of different races vastly different life chances.⁸⁹

Consistent with trying to make sense of the ways that racial order is built, maintained, and made difficult to parse, Charles Mills, in his canonical work, *The Racial Contract*, put it this way:

One could say then, as a general rule, that *white misunderstanding, misrepresentation, evasion, and self-deception on matters related to race* are among the most pervasive mental phenomena of the past few hundred years, a cognitive and moral economy psychically required for conquest, colonization and enslavement. And these phenomena are in no way *accidental*, but *prescribed* by the Racial Contract, which requires a certain schedule of structured blindness and opacities in order to establish and maintain the white polity.⁹⁰

This, then, is a challenge, because in the face of rampant denial in Silicon Valley about the impact of its technologies on racialized people, it becomes difficult to foster an understanding and appropriate intervention into its practices. Group identity as invoked by keyword searches reveals this profound power differential that is reflected in contemporary U.S. social, political, and economic life. It underscores how much engineers have control over the mechanics of sense making on the web about complex phenomena. It begs the question that if the Internet is a tool for progress and advancement, as has been argued by many media scholars, then cui bono—to whose benefit is it, and who holds the power to shape it? Tracing these historical constructions of race and gender offline provides more information about the context in which technological objects such as commercial search engines function as an expression of a series of social, political, and economic relations—relations often obscured and normalized in technological practices, which most of Silicon Valley's leadership is unwilling to engage with or take up.⁹¹

Studying Google keyword searches on identity, and their results, helps further thinking about what this means in relationship to marginalized groups in the United States. I take up the communications

scholar Norman Fairclough's rationale for doing this kind of critique of the discourses that contribute to the meaning-making process as a form of "critical social science."⁹² To contextualize my method and its appropriateness to my theoretical approach, I note here that scholars who work in critical race theory and Black feminism often use a qualitative method such as close reading, which provides more than numbers to explain results and which focuses instead on the material conditions on which these results are predicated.

Challenging Cybertopias

All of this leads to more discussion about ideologies that serve to stabilize and normalize the notion of commercial search, including the still-popular and ever-persistent dominant narratives about the neutrality and objectivity of the Internet itself—beyond Google and beyond utopian visions of computer software and hardware. The early cybertarian John Perry Barlow's infamous "A Declaration of the Independence of Cyberspace" argued in part, "We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth. We are creating a world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity."⁹³ Yet the web is not only an intangible space; it is also a physical space made of brick, mortar, metal trailers, electronics containing magnetic and optical media, and fiber infrastructure. It is wholly material in all of its qualities, and our experiences with it are as real as any other aspect of life. Access to it is predicated on telecommunications companies, broadband providers, and Internet service providers (ISPs). Its users live on Earth in myriad human conditions that make them anything but immune from privilege and prejudice, and human participation in the web is mediated by a host of social, political, and economic access points—both locally in the United States and globally.⁹⁴

Since Barlow's declaration, many scholars have challenged the utopian ideals associated with the rise of the Internet and its ability to free us, such as those espoused by Barlow, linking them to neoliberal notions of individualism, personal freedom, and individual control. These linkages are important markers of the shift from public- or state-sponsored

institutions, including information institutions, as the arbiters of social freedoms to the idea that free markets, corporations, and individualized pursuits should serve as the locus of social organization. These ideas are historically rooted in notions of the universal human being, unmarked by difference, that serve as the framework for a specific tradition of thinking about *individual* pursuits of equality. Nancy Leys Stepan of Cornell University aptly describes an enduring feature of the past 270 years of liberal individualism, reinvoked by Enlightenment thinkers during the rising period of modern capitalism:

Starting in the seventeenth century, and culminating in the writings of the new social contract philosophers of the eighteenth century, a new concept of the political individual was formulated—an abstract and innovative concept, an apparent oxymoron—the imagined *universal individual* who was the bearer of equal political rights. The genius of this concept, which opened the door to the modern polis, was that it defined at least theoretically, an individual being who could be imagined so stripped of individual substantiation and specification (his unique self), that he could stand for every man. Unmarked by the myriad specificities (e.g., of wealth, rank, education, age, sex) that make each person unique, one could imagine an abstract, non-specific individual who expressed a common psyche and political humanity.⁹⁵

Of course, these notions have been consistently challenged, yet they still serve as the basis for beliefs in an ideal of an unmarked humanity—nonracialized, nongendered, and without class distinction—as the final goal of human transcendence. This teleology of the abstracted individual is challenged by the inevitability of such markers and the ways that the individual particularities they signal afford differential realities and struggles, as well as privileges and possibilities. Those who become “marked” by race, gender, or sexuality as other are deviations from the universal human—they are often lauded for “transcending” their markers—while others attempt to “not see color” in a failing quest for colorblindness. The pretext of universal humanity is never challenged, and the default and idealized human condition is unencumbered by racial and gender distinction. This subtext is an important part of the narrative that somehow personal liberties can be realized through

technology because of its ability to supposedly strip us of our specifics and make us equal. We know, of course, that nothing could be further from the truth. Just ask the women of #Gamergate⁹⁶ and observe the ways that racist, sexist, and homophobic comments and trolling occur every minute of every hour of every day on the web.

As I have suggested, there are many myths about the Internet, including the notion that what rises to the top of the information pile is *strictly* what is most popular as indicated by hyperlinking. Were that even true, what is most popular is not necessarily what is *most true*. It is on this basis that I contend there is work to be done to contextualize and reveal the many ways that Black women are embedded within the most popular commercial search engine—Google Search—and that this embeddedness warrants an exploration into the complexities of whether the content surfaced is a result of popularity, credibility, commerciality, or even a combination thereof. Using the flawed logic of democracy in web rankings, the outcome of the searches I conducted would suggest that both sexism and pornography are the most “popular” values on the Internet when it comes to women, especially women and girls of color. In reality, there is more to result ranking than just how we “vote” with our clicks, and various expressions of sexism and racism are related.



Student protests on college campuses have led to calls for increased support of students of color, but one particular request became a matter of national policy that led to a threat to the Library of Congress's budget in the summer of 2016. In February 2014, a coalition of students at Dartmouth College put forward "The Plan for Dartmouth's Freedom Budget: Items for Transformative Justice at Dartmouth" (the "Freedom Plan"),¹ which included a line item to "ban the use of 'illegal aliens,' 'illegal immigrants,' 'wetback,' and any racially charged term on Dartmouth-sanctioned programming materials and locations." The plan also demanded that "the library search catalog system shall use undocumented instead of 'illegal' in reference to immigrants." Lisa Peet, reporting for *Library Journal*, noted,

The replacement of the subject heading was the culmination of a two-year grassroots process that began when Melissa Padilla, class of 2016, first noticed what she felt were inappropriate search terms while researching a paper on undocumented students at Dartmouth's Baker-Berry Library in 2013. While working with research and instruction services librarian Jill Baron, Padilla told LJ [*Library Journal*], she realized that nearly every article or book she looked at was categorized with the subject heading "Illegal aliens."²

The Dartmouth College librarians became deeply engaged in petitioning the Library of Congress. According to Peet, "Baron, DeSantis, and research and instruction services librarian Amy Witzel proposed that the students gather documentation to prove that 'Illegal aliens' is not a preferred term, and to find evidence that better terms—such as 'Undocumented immigrant,' which was their initial suggestion for a replacement—were in common use. At that point news organizations such as the Associated Press, *USA Today*, ABC, the *Chicago Tribune*,

and the *LA Times* had already committed not to use the term 'Illegal' to describe an individual."³ Though unsuccessful in 2015, the librarians' case to the Library of Congress had gained traction, and the librarian and professor Tina Gross at St. Cloud State University began organizing caucuses and committees in the American Libraries Association, including the subject analysis committee, social responsibilities round table, and REFORMA, which advocates for library services for Latinos and those who speak Spanish. Social media campaigns ensued, organized under the Twitter hashtags #DropTheWord and #NoHumanBeingIsIllegal.⁴ By March 29, 2016, Dartmouth College's student-led organization the Coalition for Immigration Reform, Equality (CoFired) and DREAMers announced in a press release that after a two-year battle, in partnership with campus librarians and the American Libraries Association, "the Library of Congress will replace the term 'illegal aliens' with 'noncitizens' and 'unauthorized immigrants' in its subject headings."⁵

"Illegal Alien" Revisited

The struggle over reclassifying undocumented immigrants was part of a long history of naming members of society as problem people. In many ways, this effort to eliminate "illegal alien" was similar to the ways that Jewish people were once classified by the Library of Congress as the "Jewish question," later to be reclassified in 1984 as "Jews," and Asian Americans were once classified as the "Yellow Peril."⁶ Control over identity is political and often a matter of public policy. Almost as soon as the successful change was approved, House Republicans introduced HR 4926 on April 13, 2016, also known as the "Stopping Partisan Policy at the Library of Congress Act," sponsored by Rep. Diane Black (R-TN). In essence, the bill threatened the Library's budget, and Black suggested that the effort to change the Library of Congress Subject Headings (LCSH) was a matter of "caving to the whims of left-wing special interests and attempting to mask the grave threat that illegal immigration poses to our economy, our national security, and our sovereignty."⁷

The battle over how people are conceptualized and represented is ongoing and extends beyond the boundaries of institutions such as the Library of Congress or corporations such as Alphabet, which owns and manages Google Search. Jonathan Furner, a professor of information

studies at UCLA, suggests that information institutions and systems, which I argue extend from state-supported organizations such as the Library of Congress to the Internet, are participating in “legitimizing the ideology of dominant groups” to the detriment of people of color.⁸ His case study of the Dewey Decimal Classification (DDC) system, for example, underscores the problematic conceptualizations of race and culture and efforts to “deracialize” the library and classification schemes.⁹ Furner offers several strategies for thinking about how to address these issues, using critical race theory as the guiding theoretical and methodological model. I believe these strategies are of great value to thinking about the information studies issues at hand in this research:

- admission on the part of designers that bias in classification schemes exists, and indeed is an inevitable result of the ways in which they are currently structured;
- recognition that adherence to a policy of neutrality will contribute little to eradication of that bias and indeed can only extend its life; [and]
- construction, collection and analysis of narrative expressions of the feelings, thoughts, and beliefs of classification-scheme users who identify with particular racially-defined populations.¹⁰

While the web-indexing process is not the same as classification systems such as DDC, the application of the theoretical model is still valid for thinking about conceptualizing algorithms and indexing models that could actively intervene in the default normativity of racism and sexism in information resources.

Problems in Classifying People

The idea of classification as a social construct is not new. A. C. Foskett suggests that classificationists are the products of their times.¹¹ In the work of Nicholas Hudson of the University of British Columbia on the origins of racial classification in the eighteenth century, he suggests that during the Enlightenment, Europeans began to construct “imagined communities,” citing Benedict Anderson’s term.¹² He says, “This mental image of a community of like-minded individuals, sharing a ‘general

will’ or a common national ‘soul,’ was made possible by the expansion of print-culture, which stabilized national languages and gave wide access to a common literary tradition.”¹³ Classification systems, then, are part of the scientific approach to understanding people and societies, and they hold the power biases of those who are able to propagate such systems. The invention of print culture accelerated the need for information classification schemes, which were often developing in tandem with the expansion of popular, scholarly, and scientific works.¹⁴ Traces of previous works defining the scientific classification of native peoples as “savage” and claims about Europeans as the “superior race,” based on prior notions of peoples and nations, began to emerge and be codified in the eighteenth century. Extensive histories have been written of how racial classification emerged in the eighteenth and nineteenth centuries in North America as a paradigm of differentiation that would support the exclusion of native and African people from social and political life.

By the nineteenth century, the processes involved in the development of racial classification marked biological rather than cultural difference and were codified to legally deny rights to property ownership and citizenship. These historical practices undergird the formation of racial classification, which is both assumed and legitimated in classification systems. Without an examination of the historical forces at play in the development of such systems, the replication and codification of people of African descent into the margins goes uncritically examined. This process can be seen in knowledge organization that both privileges and subordinates through information hierarchies such as catalogs and classification systems. The field of library science has been implicated in the organization of people and critiqued for practices that perpetuate power by privileging some sectors of society at the expense of others.

Traditional library and information science (LIS) organization systems such as subject cataloging and classification are an important part of understanding the landscape of how information science has inherited and continues biased practices in current system designs, especially on the web.

Opportunities abound for the interdisciplinarity of LIS to extend more deeply into cultural and feminist studies, because these social science fields provide powerful and important social context for information about people that can help frame how that information is organized

and made available. To date, much of the attention to information organization, storage, and retrieval processes has been influenced and, more importantly, funded by scientific research needs stemming from World War II and the Cold War.¹⁵ The adoption of critical race theory as a stance in the field would mean examining the beliefs about the neutrality and objectivity of the entire field of LIS and moving toward undoing racist classification and knowledge-management practices. Such a stance would be a major contribution that could have impact on the development of new approaches to organizing and accessing knowledge about marginalized groups.

If the information-retrieval priority of making access to recorded information efficient and expedient is the guiding process in the development of technical systems, from databases to web search engines, then what are the distinguishing data markers that define information about racialized people and women in the United States? What have primarily been missing from the field of information science, and to a lesser degree library science, are the issues of representation that are most often researched in the fields of African American studies, gender studies, communications, and increasingly digital media studies. Information organization is a matter of sociopolitical and historical processes that serve particular interests.

A Short History of Misrepresentation in Classifying People

In order to understand how racial and gender representations in Google Search express the same traditional bias that exists in other organizational systems, an overview of how women and non-Whites have been historically represented in information categorization environments is in order. The issue of misrepresentations of women and people of color in classification systems has been significantly critiqued.¹⁶ Hope A. Olson, an associate dean and professor at the School of Information Studies at the University of Wisconsin, Milwaukee, has contributed among the most important theories on the social construction of classification that many of us in the field assign to our students as a way of fostering greater awareness about the power that library, museum, and information professionals hold. Those who have the power to design systems—classification or technical—hold the ability to prioritize

hierarchical schemes that privilege certain types of information over others. An example of these biases include the cataloging of people as subjects in the Library of Congress Subject Headings (LCSH), which serve as a foundational and authoritative framework for categorizing information in libraries in the United States. The LCSH have been noted to be fraught with bias, and the radical librarian Sanford Berman details the ways that this bias has reflected Western perspectives:

Since the first edition of *Library of Congress Subject Headings* appeared 60 years ago, American and other libraries have increasingly relied on this list as the chief authority—if not the sole basis—for subject cataloging. There can be no quarrel about the practical necessity for such labor-saving, worry-reducing work, nor—abstractly—about its value as a global standardizing agent, as a means for achieving some uniformity in an area that would otherwise be chaotic. . . . But in the realm of headings that deal with people and cultures—in short, with humanity—the LC list can only “satisfy” parochial, jingoistic Europeans and North Americans, white-hued, at least nominally Christian (and preferably Protestant) in faith, comfortably situated in the middle- and higher-income brackets, largely domiciled in suburbia, fundamentally loyal to the Established Order, and heavily imbued with the transcendent, incomparable glory of Western Civilization.¹⁷

Eventually the LCSH abolished labels such as “Yellow Peril” and “Jewish Question” or made substitutions in the catalog, changing “Race Question” or “Negroes” to “Race Relations” and “Afro-Americans,”¹⁸ but the establishment of such headings and the subsequent decade-long struggles to undo them underscored Berman’s point about Western racial bias. (In fact, it was Berman who led the field in calling for anti-racist interventions into library catalogs in the 1970s.) Patriarchy, like racism, has been the fundamental organizing point of view in the LCSH. The ways in which women were often categorized was not much better, with headings such as “Women as Accountants” in lieu of the now-preferred “Women Accountants”; women were consistently an aberration to the assumed maleness of a subject area.¹⁹

Furthermore, efforts at self-identity from the perspective of marginalized and oppressed groups such as the Roma or Romanies cannot es-

cape the stigmatizing categorization of their culture as “Gypsies,” even though their “see also” designation to “rogues and vagabonds” was finally dropped from the LCSH.²⁰ A host of other problematic naming conventions including “Oriental” instead of “Asian” and the location of Christianity at the top of the religious hierarchy, with all deviations moving toward the classification of “Primitive,” suggests that there is still work to be done in properly addressing and classifying groups of people around identity.²¹ Olson says, “the problem of bias in classification can be linked to the nature of classification as a social construct. It reflects the same biases as the culture that creates it.”²² These types of biases are often seen in offline information practices where conquest is a means of erasing the history of one dynasty or culture by the subsequent regime.²³ Olson’s research has already shown that classifications reflect the philosophical and ideological presumptions of dominant cultures over subordinate cultures or groups. For example, in traditional Dewey Decimal Classification (DDC), over 80% of its religion section is devoted exclusively to Christianity, even though there are greater numbers of other religious texts and literature.²⁴ Olson points to the Library of Congress Classification (LCC) and its biases toward North American and European countries in volumes on the law, with far fewer allocations of space for Asia, Eurasia, Africa, the Pacific area, and Antarctica, reflecting the discourse of the powerful and the presumption of marginality for all others.²⁵

In this respect, Olson reminds us that the ordering of information provided in classification schemes “tends to reflect the most mainstream version of these relationships” because “classificatory structures are developed by the most powerful discourses in a society. The result is the marginalization of concepts outside the mainstream.”²⁶ In other words, the most mainstream (e.g., White, heterosexual, Christian, middle-class) controlling regimes in society will privilege themselves and diminish or subdue all others in the organization of what constitutes legitimate knowledge. When we inherit privilege, it is based on a massive knowledge regime that foregrounds the structural inequalities of the past, buttressed by vast stores of texts, images, and sounds saved in archives, museums, and libraries. Certainly, classification systems have some boundaries and limits, as they are often defined in whole by what is included and what is excluded.²⁷ In the case of most library databases in

the United States, Eurocentrism will dominate the canons of knowledge. Knowledge management reflects the same social biases that exist in society, because human beings are at the epicenter of information curation. These practices of the past are part of the present, and only committed and protracted investments in repairing knowledge stores to reflect and recenter all communities can cause a shift toward equality and inclusion in the future. This includes reconciling our brutal past rather than obscuring or minimizing it. In this way, we have yet to fully confront our histories and reconstitute libraries and museums toward reconciliation and reparation.

Search engines, like other databases of information, are equally bounded, limited to providing only information based on what is indexed within the network. Who has access to provide information in the network certainly impacts whether information can be found and surfaced to anyone looking for it. Olson’s research points to the ways that some discourses are represented with more power, even if their social classifications are relatively small:

In North American society, taking away women, African Americans, Hispanic Americans, French Canadians, Native peoples, Asian Americans, lesbians and gay men, people with disabilities, anyone who is not Christian, working class and poor people, and so forth, one is left with a very small “core.” An image that shows the complexity of these overlapping categories is that of a huge Venn diagram with many sets limited by Boolean ANDs. The white AND male AND straight AND European AND Christian AND middle-class AND able-bodied AND Anglo mainstream becomes a very small minority . . . , and each set implies what it is not. The implication of this image is that not every person, not every discourse, not every concept, has equal weight. Some discourses simply wield more power than others.²⁸

Arguably, if education is based in evidence-based research, and knowledge is a means of liberation in society, then the types of knowledge that widely circulate provide a crucial site of investigation. How oppressed people are represented, or misrepresented, is an important element of engaging in efforts to bring about social, political, and economic justice.

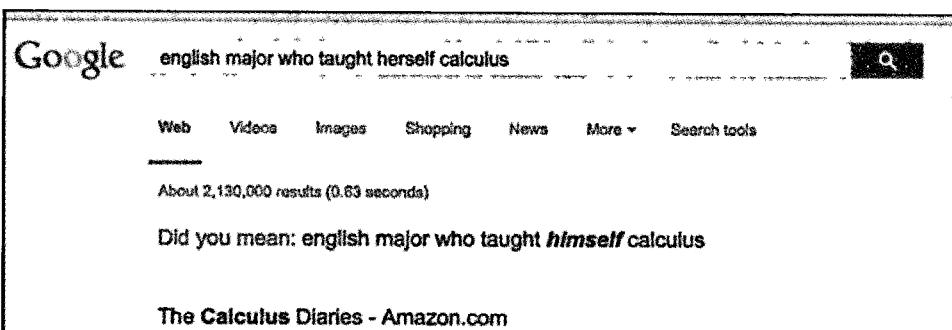


Figure 5.1. Google autocorrects to “himself” rather than “herself.” Search sent to me by a colleague, June 16, 2016.

We have to ask ourselves what it means in practical terms to search for concepts about gender, race, and ethnicity only to find information lacking or misrepresentative, whether in the library database or on the open web. Olson’s notion that cultural metaphor is the basis of the construction of classification systems means these cultural metaphors are profoundly represented in the notions of the “Jewish Question” or the “Race Question.” These subject headings suggest both an answer and a point of view from which the problems of Jews and race are presupposed. Simply put, to phrase “Jewish” or “race” as a question or problem to be answered suggests a point of view on the part of the cataloger that is quite different from how a Jewish person or a racialized person might frame themselves. It is here that the context and point of view of library and information science professionals who are responsible for framing people and communities as “problems” and “questions” is important. By examining the ways that Black people specifically have been constructed in the knowledge schemes, the African American studies professor and philosopher Cornel West aptly describes the positionality of how this community is depicted in the West:

Black people as a problem-people rather than people with problems; black people as abstractions and objects rather than individuals and persons; black and white worlds divided by a thick wall (or a “Veil”) . . . ; black rage, anger, and fury concealed in order to assuage white fear and anxiety; and black people rootless and homeless on a perennial journey

to discover who they are in a society content to see blacks remain the permanent underdog.²⁹

The library scholar Joan K. Marshall points to the way this idea was expressed in the Library of Congress when “N*ggers” was a legitimate subject category, reflecting the “social backgrounds and intellectual levels” of users, concretizing oppressive race relations.³⁰ Difference, in the case of the Library of Congress, is in direct relation to Whiteness as the norm. No one has made this more clear than Berman, whose groundbreaking work on Library of Congress Subject Headings has forever changed the field. He notes that in the case of both Jews and the representations of race, these depictions are not without social context:

For the image of the Jew to arouse any feelings, pro or con, he [sic] had to be generalized, abstracted, depersonalized. It is always possible for the personal, individual case to contradict a general assertion by providing living, concrete proof to the contrary. For the Jews to become foils of a mass movement, they had to be converted into objectified symbols so as to become other than human beings.³¹

In the case of Google, because it is a commercial enterprise, the discussions about its similar information practices are situated under the auspices of free speech and protected corporate speech, rather than being posited as an information resource that is working in the public domain, much like a library. An alternative possibility could be that corporate free speech in the interests of advertisers could be reprioritized against the harm that sexist and racist speech on the Internet could have on those who are harmed by it. This is the value of using critical race theory—considering that free speech may in fact not be a neutral notion but, rather, a conception that when implemented in particular ways silences many people in the interests of a few.

The disclaimer by Google for the problem of searching for the word “Jew” leading to White supremacist, Holocaust-denial web results is surprisingly similar to the construction of Jewish identity in the LCSH. Both systems reflect the nature of the relationship between Jewish and non-Jewish Europeans and North Americans. This is no surprise, given that hyperlinking and indexing are directly derived from library science

citation-analysis practices. This linkage between the indexing practices of the World Wide Web and the traditional classification systems of knowledge structures such as the Library of Congress is important. Both systems rely on human decisions, whether given over en masse to artificial intelligence and algorithms or left to human beings to catalog. The representation of people and cultures in information systems clearly reflects the social context within which the subjects exist. In the case of search engines, not unlike cataloging systems, the social context and histories of exploitation or objectification are not taken into explicit consideration—rather, they are disavowed. What can be retrieved by information seekers is mediated by the technological system—be it a catalog or an index of web pages—by the system design that otherizes. In the case of the web, old cataloging and bibliometric practices are brought into the modern systems design.

Library science scholars know that bibliographic and naming controls are central to making knowledge discoverable.³² Part of the issue is trying to understand who the audience is for knowledge and naming and organizing information in ways that can be discovered by the public. Berman cites Joan Marshall's critiques of the underlying philosophy of the Library of Congress's subject-cataloging practices and the ways that they constitute an audience through organizational bias, wherein a "majority reader" is established as a norm and, in the case of the Library of Congress, is often "white, Christian (usually Protestant) and male."³³ Indeed, these scholars are taking note of the influence that categorization systems have on knowledge organization and access. What is particularly important in the interrogation of these marginalizing information-management systems is Berman's reference to the Algerian psychologist Franz Fanon's articulation of the mechanics of cultural "brain washing" that occurs through racist cataloging practices.³⁴ Berman underscores that the problems of racial representation and racism are deeply connected to words and images and that a racist worldview is embedded in cataloging practices that serve to bolster the image and domination of Western values and people (i.e., White, European, and North Americans over people of African descent). The library practitioner Matthew Reidsma gave a recent gift to the profession when he blogged about library discovery systems, or search interfaces, that are just as troubled as commercial interfaces. In his blog post, he details the limitations of

The screenshot shows a blog post from matthew.reidsma.com/articles/173. The post is titled "ALGORITHMIC BIAS IN LIBRARY DISCOVERY SYSTEMS". It is dated March 11, 2016, and includes a "sPrev" link. The content discusses the move to centralized discovery layers in academic libraries, noting the ascendence of Google and libraries' focus on user experience. Unlike vendor-specific tools, discovery presents a simplified picture of the research process. The post highlights the familiar single search box and the aggregated results from various providers, comparing it to the Google model. A small profile picture of the author, Matthew Reidsma, is visible at the top left.

Figure 5.2. A call to the profession to address algorithmic bias in library discovery systems by Matthew Reidsma attempts to influence the field of information studies. Source: Reidsma, 2016.

databases, the kinds of gender biases that are present in discovery tools, and how little innovation has been brought to bear in resolving some of the contradictions we know about.³⁵

I sought to test the call that Reidsma made to the profession to interrogate library information management tools by conducting searches in a key library database. I looked in the largest library image database available to academic libraries, ArtStor, and found troublesome practices of metadata management there too. Undoubtedly, these kinds of cataloging stances can be evaluated in the context of the field of library science, which is largely averse to teaching and talking about race and the White racial gaze on information. I have published several articles with colleagues about the challenges of teaching about race in the library and information studies classroom and the importance of integrating theory and training of information workers around issues of social justice in the profession. I interpret these kinds of cataloging mishaps as a result of the investment of the profession in colorblind ideology. Unable and unequipped to think through the complexities of systems of racialization, the profession writ large struggles to find frameworks to



Figure 5.3. Search in ArtStor for “black history” features the work of a series of European and White American artists, March 2, 2016. The first result is work by Thomas Waterman Wood.

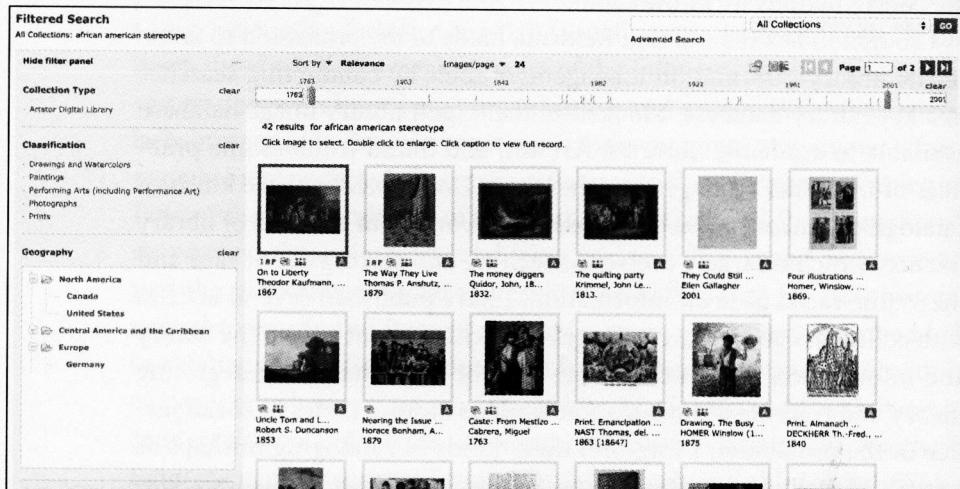


Figure 5.4. *On to Liberty*, an oil on canvas by Theodore Kauffman, a German painter, is the first item under “African American stereotype.”

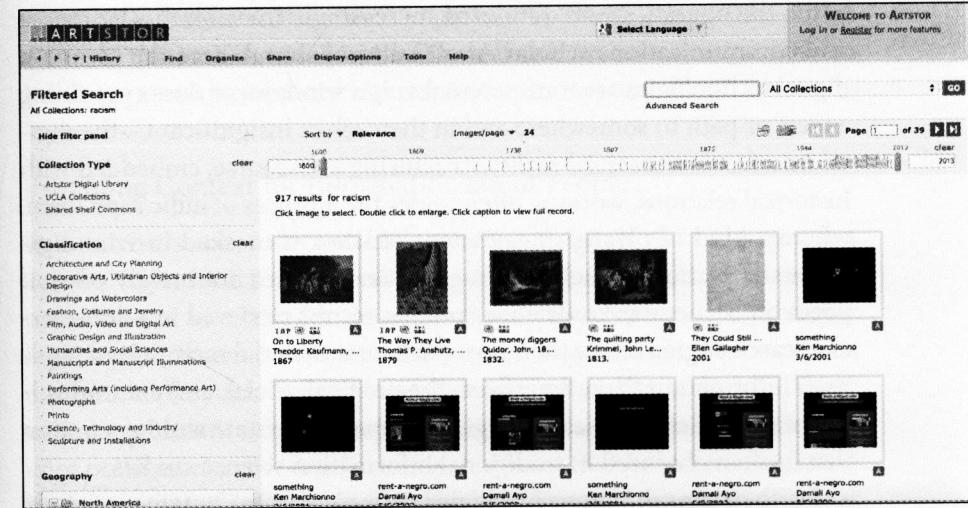


Figure 5.5. A satirical piece by the artist Damali Ayo and her online piece *Rent-A-Negro*, which is a critique of liberal racial ideologies that tokenize African Americans. The work is cataloged as “racism.”

think critically about the long-term consequences of misidentification of people and, in this case, concepts about works of art.

Search as a Source of Reality

Indeed, problematic results in ArtStor are just a small window into a long and troubled history of misrepresentation in library subject cataloging and classification systems, which are faithful reflections of the problematic representations in mainstream U.S. culture. Our ability to recognize these challenges can be enhanced by asking questions about how technological practices are embedded with values, which often obscure the social realities within which representations are formed. The interface of the search engine as a mechanism for accessing the Internet is not immune, nor impartial, to the concerns of embedded value systems. Search is also more than the specific mathematical algorithms and deep-machine learning developed by computer scientists and software engineers to index upward of a trillion pages of information and move some from the universal data pile to the first page of results on a computer screen. The interface on the screen presents an information reality,

while the operations are rendered increasingly invisible.³⁶ The media and communications scholar Alex Galloway destabilizes the idea that digital technologies are transparent, benign windows or doors providing a view or path to somewhere and in themselves insignificant—the digital interface is a material reality structuring a discourse, embedded with historical relations, working often under the auspices of ludic capitalism, where a kind of playful engagement of labor is masked in vital digital media platforms such as Google.³⁷ Search does not merely present pages but structures knowledge, and the results retrieved in a commercial search engine create their own particular material reality. Ranking is itself information that also reflects the political, social, and cultural values of the society that search engine companies operate within, a notion that is often obscured in traditional information science studies.

Further, new digital technologies may constitute containers for old media discourses, and the web interface (such as a plain Google search box) is a transitional format from previous media forms.³⁸ Certainly in the case of digital technology such as commercial search engines, the interface converges with the media itself. Commercial search, in the case of Google, is not simply a harmless portal or gateway; it is in fact a creation or expression of commercial processes that are deeply rooted in social and historical production and organization processes. John Battelle, who has carefully traced the history of Google, describes search as the product of our needs and desires, aggregated by companies:

Link by link, click by click, search is building possibly the most lasting, ponderous, and significant cultural artifact in the history of humankind: the Database of Intentions. The Database of Intentions is simply this: the aggregate results of every search ever entered, every result list ever tendered, and every path taken as a result. . . . This information represents the real-time history of post-Web culture—a massive clickstream database of desires, needs, wants, and preferences that can be discovered, subpoenaed, archived, tracked and exploited for all sorts of ends.³⁹

Undoubtedly, search is also pivotal in the development of artificial intelligence. In many ways, Google Search is an attempt to use computer science as a basis for sorting and making decisions about the relevance and quality of information rather than human sorting and web-indexing

practices—practices that search engine companies such as Yahoo! and those of the past invested in heavily and that were both expensive to implement and limited and less responsive in real time.⁴⁰

Providing Context for Information about People

In a narrow sense, information is a series of signals and messages that can be expressed through mathematics, algorithms, and statistical probabilities. In a broader sense, however, Tefko Saracevic, a professor emeritus of information science at Rutgers, suggests that information is constituted through “cognitive processing and understanding.”⁴¹ There is a pivotal relationship between information and users that is dependent on human understanding. It is this point that I want to emphasize in the context of information retrieval: information provided to a user is deeply contextualized and stands within a frame of reference. For this reason, it is important to study the social context of those who are organizing information and the potential impacts of the judgments inherent in informational organization processes. Information must be treated in a context; “it involves motivation or intentionality, and therefore it is connected to the expansive social context or horizon, such as culture, work, or problem-at-hand,” and this is fundamental to the origins of information science and to information retrieval.⁴² Information retrieval as a practice has become a highly commercialized industry, predicated on federally funded experiments and research initiatives, leading to the formation of profitable ventures such as Yahoo! and Google, and a focus on information relevance continues to be of importance to the field. Information science is essentially deeply entwined with the history of library science and has primarily been concerned with the collection, storage and retrieval, and access to and use of information. Saracevic notes that “the domain of information science is the transmission of the universe of human knowledge in recorded form, centering on manipulation (representation, organization, and retrieval) of information, rather than knowing information.”⁴³ This foregrounds the ways that representations in search engines are decontextualized in one specific type of information-retrieval process, particularly for groups whose images, identities, and social histories are framed through forms of systemic domination. Although there is a long, broad, and historical context for

addressing categorizations, the impact of learning from these traditions has not yet been fully realized.⁴⁴

Attention to “the universe of human knowledge” is suggestive for contextualizing information-retrieval practices this way, leading to inquiries into the ways current information-retrieval practices on the web, via commercial search engines, make some types of information available and suppress others. The present focus on the types of information presented in identity-based searches shows that they are removed from the social context of the historical representations and struggles over disempowering forms of representation. These critiques have been levied toward other media practices such as television and print culture. Whether human beings believe that the information delivered in search is relevant has consistently been the basis of judgment about information quality,⁴⁵ but what is underdiscussed is that retrieval of information in commercial platforms such as web-based search engines is not unique to the individual searcher. A web-based commercial search engine does not entirely “know” who a user is, and it is not customizing everything to our personal and political tastes, although it is aggregating us to people it thinks are similar to us on the basis of what is known through our digital traces.

Finding Culturally Situated Information on the Web

The field of LIS is significantly engaged in information classification and organization work, which can inform the framework for thinking about developing ICTs that are focused on surfacing prioritized results, such as the search engine. Critical race theory in this process of developing information-organization tools is of great value, particularly when thinking about the phenomenon of excessive recall of documents on the web that are irrelevant or decontextualized. Responses to the kinds of problematic biases in large commercial search engines are part of the growing motivation behind a host of culturally situated search engines that are emerging, particularly Blackbird (www.blackbirdhome.com), a Mozilla Firefox browser designed to help surface content of greater relevance to African Americans. Blackbird has been met with mixed reviews, from support and interest to wholesale rejection.⁴⁶ In any case, organizations and individuals are responding to the limits of

traditional commercial search engines through the development of such search engines. Identity-focused websites, a combination of web-based browsers and web directories, are emerging to prioritize the interests of specific communities on the basis of the human-curated practices of the past and can be seen in search engines such as BlackWebPortal (www.blackwebportal.com); GatewayBlackPortal (www.gatewayblack.com), which is based on international models such as JGrab, a Jewish search engine; BlackFind.com (www.blackfind.com); and Blackbird. Sites such as Jewogle (www.jewogle.com), which serves as an online encyclopedia of the accomplishments of Jewish people; Jewish.net (<http://jewish.net/>), which is used to “search the Jewish Web”; JewGotIt (www.jewgotit.com); and Maven Search (www.maven.co.il), which catalogs over fifteen thousand Jewish websites, have emerged in the hundreds, some tongue in cheek, across religion, culture, and national origin. Much of this is a response of communities that are seeking control over relevant content and representation, as well as access to quality information within racial or group identity.

One of the fundamental challenges for these culturally situated search engines is the way in which they make visible the contradictions and biases in search engines, which André Brock discusses in relationship to Blackbird. He notes that “Blackbird’s efforts to foreground African American content were seen as an imposition on the universal appeal of the internet, highlighting the perception of the browser as a social structure limited by Black representation.”⁴⁷ Brock’s work indicates that though there is a demand for culturally relevant Internet browsing that will help surface content of interest to Black people, its value works against norms on the web, making it less desirable.

Reproducing Social Relations through Information Technologies

Online racial disparities cannot be ignored because they are part of the context within which ICTs proliferate, and the Internet is both reproducing social relations and creating new forms of relations based on our engagement with it. Technologies and their design do not dictate racial ideologies; rather, they reflect the current climate. As users engage with technologies such as search engines, they dynamically co-construct content and the technology itself.⁴⁸ Online information and content

available in search is also structured systemically by the infusion of advertising revenue and the surveillance of user searches, which the subjects of such practices have very little ability to reshape or reformulate. Lack of attention to the current exploitative nature of online keyword searches only further entrenches the problematic identities in the media for women of color, identities that have been contested since the advent of commercial media such as broadcast, print, and radio. Noticeably absent in the discussions of search is the broader social and technical interplay that exists dynamically in the way technology is increasingly mediating public access to information, from libraries to the search engine.

Now, more than ever, a new conception of information access and quality rooted in historical, economic, and social relations could have a transformational effect on the role and consequences of search engines. It is my goal through this research to ensure that traditionally under-represented ideas and perspectives are included in the shaping of the field—to surface counternarratives that would allow for a questioning of the normalization of such practices. Rather than prioritize the dominant narratives, Internet search platforms and technology companies could allow for greater expression and serve as a democratizing tool for the public. This is rendered impossible with the current commercial practices.

What we need are public search engine alternatives, united with public-interest journalism and librarianship, to ensure that the public has access to the highest quality information available.

6

The Future of Information Culture



In March 2010, the U.S. Federal Communications Commission (FCC) put forward its ten-year broadband plan, wherein it called for high-speed Internet to become the common medium of communications in the United States.¹ This plan still governs the information landscape. The FCC envisioned that the Internet, as the common medium, would potentially displace current telecommunications and broadcast television systems as the primary communications vehicle in the public sphere. According to the report, “almost two-thirds of the time users spend online is focused on communication, information searching, entertainment or social networking.”² The plan called for increased support for broadband connectivity to facilitate Americans’ ability to access vital information, and it is focused on infrastructure, devices, accessibility, and connectivity. However, the plan made no mention of the role of search engines in the distribution of information to the public, with the exception of noting that the plan itself will be archived and made available in perpetuity in the archives of the Internet. Primary portals to the Internet, whether connecting through the dwindling publicly funded access points in libraries and schools or at home, serve as a gateway to information and cannot be ignored. Access to high-quality information, from journalism to research, is essential to a healthy and viable democracy. As information moves from the public sphere to private control by corporations, a critical juncture in the quality of information available and the public’s ability to sift and use it is at stake, as the noted political economist Herbert Schiller forewarned:

The American economy is now hostage to a relatively small number of giant private companies, with inter-locking connections, that set the national agenda. This power is particularly characteristic of the communication and information sector where the national cultural-media agenda is provided by a very small (and declining) number of integrated private

combines. The development has deeply eroded free individual expression, a vital element of a democratic society.³

An increasingly de- and unregulated commercially driven Internet raises significant issues about how information is accessed and made available. This is exacerbated by the gamification of news and headlines, as Nicole Cohen of the Institute of Communication, Culture, Information, and Technology at the University of Toronto, Mississauga, has documented in her ethnography of journalists who write for online news outlets. In her book *Writers' Rights: Freelance Journalism in a Digital Age*, she documents the increasing tensions between journalists and commercial news media organizations.⁴ In some cases, journalists are facing screens that deliver real-time analytics about the virality of their stories. Under these circumstances, journalists are encouraged to modify headlines and keywords within a news story to promote greater traction and sharing among readers. The practices Cohen details are precisely the kind of algorithmically driven analytics that place pressure on journalists to modify their content for the express purposes of increasing advertising traffic. Undoubtedly, in this case, big-data analytics have potential to significantly compromise the quality of reporting to the public.

As quality information typically provided by the public sector moves into more corporate and commercial spaces, the ability of the public to ensure protections that are necessary in a democracy is eroded, due to the cost of access. Organizations such as FreePress.org are showing how the rise of advertising and commercial interests have bankrupted the quality and content of journalism, heretofore considered a fundamental and necessary component of a democratic society. The media scholars Robert McChesney and John Nichols have noted in great historical detail and with abundant concrete evidence the importance of information in a democratic society—free from commercial interests.⁵ These rapid shifts over the past decade from the public-interest journalism environment prior to the 1990s, along with the corporate takeover of U.S. news media, have eroded the quality of information available to the public. Similarly, the move of the Internet from a publicly funded, military-academic project to a full-blown commercial endeavor has also impacted how information is made available on the web.

Media stereotypes, which include search engine results, not only mask the unequal access to social, political, and economic life in the United States as broken down by race, gender, and sexuality; they maintain it.⁶ This suggests that commercial search engines, in order to opt out of such traditional racist representations, might want, at minimum, to do something like a “disclaimer” and, at maximum, to produce a permanent “technical fix” to the proliferation of racist or sexist content. Veronica Arreola wondered as much on the *Ms.* blog in 2010 when Google Instant, a new search-enhancement tool, initially did not include the words “Latinas,” “lesbian,” and “bisexual” because of their X-rated front-page results: “You’re Google. . . . I think you could figure out how to put porn and violence-related results, say, on the second page?”⁷

It is these kinds of practices that mark the consequences of the rapid shift over the past decade from public-interest information to the corporate takeover of U.S. news media, which has made locating any kind of alternative information increasingly difficult and pushed the public toward the web. Equally, media consolidations have contributed to the erosion of professional standards such as fact checking, not misrepresenting people or situations, avoiding imposing cultural values on a group, and distinguishing between commercial and advertising interests versus editorial decisions—all of which can be applied to information provision on the web.⁸ As the search arena is consolidated under the control of a handful of corporations, it is even more crucial to pay close attention to the types of processes that are shaping the information prioritized in search engines. In practice, the higher a web page is ranked, the more it is trusted. Unlike the vetting of journalists and librarians, who are entrusted to fact check and curate information for the public according to professional codes of ethics, the legitimacy of websites’ ranking and credibility is simply taken for granted. The take-home message is that, when it comes to online commercial search engines, it is no longer enough to simply share news and education on the web; we must ask ourselves how the things we want to share are found and how the things we find have appeared.

A Monopoly on Information

Not enough attention has been paid to Google's monopoly on information in the most recent debates about network control. The focus on net neutrality in the U.S. is largely invested in concerns over the movement of packets of data across commercial networks owned by the telecommunications and cable giants, which include AT&T, Verizon, DirecTV, and Comcast. Much of the debate has focused on maintaining an open Internet, free from traffic-routing discrimination. In this context, discrimination refers to the movement of data and the rights of content providers not to have their traffic delayed or managed across the network regardless of size or content. Focus on content prioritization processes should enter the debates over net neutrality and the openness of the web when mediated by search engines, especially Google. Over the past few years, consumer watchdog organizations have been enhancing their efforts to provide data about Google's commercial practices to the public, and the Federal Trade Commission is investigating everything from Wi-Fi data harvesting of consumer data to Google's horizontal ownership and dominance of web-based services such as YouTube, AdSense, Google Maps, Blogger, Picasa, Android, Feedburner, and so on. Internet service providers have been set back by the recent U.S. court of appeals decision to protect the rights of consumers via maintaining the FCC stance on protecting net neutrality. The decision prevents Comcast from prioritizing or discriminating in traffic management over its networks. Organizations such as the Open Internet Coalition have been at the fore in lobbying Congress for protections from the prioritization of certain types of lawful Internet traffic that multinational telecommunications companies are able to promote, while simultaneously blocking access to their networks by competitors. Quietly, companies such as Google, Facebook, and Twitter that have high volumes of traffic have backed the Open Internet Coalition in an effort to ensure that they will have the necessary bandwidth to support their web-based assets that draw millions of users a day to their sites with tremendous traffic.

Outside the United States, Google has faced a host of complaints about representations of material culture and identity. In the realm of public information, the former Harvard University librarian Robert

Darnton outlined the problematic issues that arose from the Google book-digitization project. In this project, Google digitized millions of books, over ten million as of the close of 2009, opening up considerable speculation about the terms on which readers will be able to access these texts. The legal issues at play at the height of the legal battle included potential violations of antitrust law and whether public interests would prevail against monopolistic tendencies inherent in one company's control and ownership of such a large volume of digital content.⁹ Proponents of Google's project suggested that the world's largest library will make previously out-of-print and unavailable texts accessible to a new generation of readers/consumers. Opponents were fearful that Google would control the terms of access, unlike public libraries, on the basis of shareholder interests. Further challenges to this project were leveled by France and Germany, which rejected the ownership of their material culture by a U.S.-based company, claiming it is impinging on their national and cultural works.¹⁰ They suggested that the digitization of works by their national citizens of the past is an infringement on the public good, which is threatened by Google's monopoly on information. In 2013, U.S. Circuit Court Judge Denny Chin ruled that the Google book project was "fair use," serving a blow to critics, and in 2015, a hearing of the case was rejected by the U.S. Supreme Court.¹¹ An appeal to the Second Circuit, New York, affirmed Google's right to claim fair use. Despite Darnton's critique, underscored by media scholars such as Siva Vaidhyanathan, a professor of media studies and law at the University of Virginia, who has written substantially on the threats of the decision to the legal concept of "fair use," the verdict underscores the power of Google's capital and its influence, to the detriment of nations that cannot withstand its move to create the largest digital repository in the world. This includes the ability to own, categorize, and determine the conditions or terms of access to such content. In support of the position against the project before the European Commission, concerns were presented by France that "a large portion of the world's heritage books in digital format will be under the control of a single corporate entity."¹²

Closer to home, with the exception of the Anti-Defamation League's previously mentioned letter, many protests of Google's information and website representation have not been based on the way cultural identities are presented, but rather the focus has been on commercial interests

in patents, intellectual property, and even page ranking. For example, in 2003, an early lawsuit against Google focused on its prioritization of high-paying advertisers that were competing against small businesses and entities that do not index pages on the basis of the pay-per-click advertising model that has come to dominate experiences of the Internet in the United States. The lawsuit by Search King and PR Network against Google alleged that Google decreased the page rank of its clients in a direct effort to annihilate competition.¹³ Since Bob Massa, the president of Search King and PR Ad Network, issued a statement against Google's biased ranking practices, Google's business practices have been under increased scrutiny, both in the U.S. and globally.

Why Public Policy Matters

Given the controversies over commercial, cultural, and ethnic representations of information in PageRank, the question that the Federal Trade Commission might ask today, however, is whether search engines such as Google should be regulated over the values they assign to racial, gendered, and sexual identities, as evidenced by the types of results that are retrieved. At one time, the FCC enforced decency standards for media content, particularly in television, radio, and print. Many political interventions over indecency and pornography on the web have occurred since the mid-1990s, with the 1996 Communications Decency Act (CDA) being the most visible and widely contested example, particularly section 230 with respect to immunity for online companies, which cannot be found liable for content posted by third parties. Section 230 is specifically designed to protect children from online pornography, while granting the greatest rights to freedom of expression, which it does by not holding harm toward Internet service providers, search engines, or any other Internet site that is trafficking content from other people, organizations, or businesses—companies such as Google, Facebook, Verizon, AT&T, Wordpress, and Wikipedia—all of which are exempt from liability under the act.¹⁴ These were the same protections afforded to Hunter Moore and his revenge-porn site discussed in chapter 4.

The attorney Gregory M. Dickinson describes the important precedents set by a court ruling against the Internet service provider Prodigy. He suggests that the court's interpretation of Prodigy's market position

was that of a “family-friendly, carefully controlled and edited Internet provider,” which engaged in processes to filter or screen offensive content in its message boards; as such, it “had taken on the role of a newspaper-like publisher rather than a mere distributor and could therefore be held liable.”¹⁵ He underscores the importance of the court ruling in *Stratton Oakmont, Inc. v. Prodigy Services Co.* (1995) that Prodigy's engagement in some level of filtering content of an objectionable nature meant that Prodigy was responsible and liable. This, he argues, was not Congress's intent—to hold harmless any platform providing content that is obscene, pornographic, or objectionable by community standards of decency.

Commercial search engines, at present, have been able to hide behind disclaimers asserting that they are not responsible for what happens in their search engine technologies. Yet Dickinson's study of the law with respect to Prodigy raises interesting legal issues that could be explored in relationship to search engines, particularly Google, now that it has admitted to engaging in filtering practices. What is most apparent since the passage of the CDA in 1996 is that decency standards on the web and in traditional media have been fodder for “the culture wars,” and by all apparent measures, indecency is sanctioned by Congress, the FCC, and media companies themselves. These protections of immunity are mostly upheld by the *Zeran v. America Online, Inc.* (1997) ruling in the U.S. Court of Appeals for the Fourth Circuit, which found that companies are not the responsible parties or authors of problematic material distributed over their hardware, software, or infrastructure, even though section 230 was intended to have these companies self-censor indecent material. Instead, the courts have ruled that they cannot hold companies liable for not self-censoring or removing content. Complicating the issues in the 1996 act is the distinction between “computer service providers” (nonmediated content) and “information providers” (mediated content).¹⁶

During the congressional hearings that led to the Federal Trade Commission investigation of Google, the reporter Matthew Ingram suggested in a September 2011 article that “it would be hard for anyone to prove that the company's free services have injured consumers.”¹⁷ But Ingram is arguably defining “injury” a little too narrowly. Searching for “Latinas” or “Asian women” brings about results that focus on porn, dating, and fetishization. What is strikingly similar in the cases of searching

for “Jew” and for “black girls” is that objectionable results materialized in Google’s page-ranking algorithm—results that might not reflect the social or historical context of the lives of each group or their desire to be represented this way. However, what is strikingly dissimilar is that Black teenagers and girls of color have far less social, political, or economic agency than the Anti-Defamation League does. Public policy must open up avenues to explore and assess the quality of group identity information that is available to the public, a project that will certainly be hotly contested but that should still ensue.

The Web as a Source of Opportunity

The web is characterized as a source of opportunity for oppressed and marginalized people, with tremendous focus put on closing the hardware, software, and access gaps on the Internet for various communities. Among the most prevalent ideas about the political aspects of technology disenfranchisement and opportunity are theories that center on the concept of the “digital divide,” a term coined in a series of speeches and surveys by the Clinton-Gore administration and the National Telecommunications Infrastructure Administration. Digital-divide narratives have focused on three key aspects of disempowerment that have led to technological deficits between Whites and Blacks: access to computers and software, development of skills and training in computer technologies, and Internet connectivity—most recently characterized by access to broadband.¹⁸

However true the disparities between Whites and non-Whites or men and women in the traditional articulations of the digital divide, often missing from this discourse is the framework of power relations that precipitate such unequal access to social, economic, and educational resources.¹⁹ Thus, the context for discussing the digital divide in the U.S. is too narrow a framework that focuses on the skills and capabilities of people of color and women, rather than questioning the historical and cultural development of science and technology and representations prioritized through digital technologies, as well as the uneven and exploitative global distribution of resources and labor in the information and communication ecosystem. Certainly, the digital divide was an important conceptual framework to deeper engagement for poor people and

people of color, but it also created new sites of profit for multinational corporations.²⁰ Closing the digital divide through ubiquitous access, training, and the provisioning of hardware and software does address the core criticisms of the digital technology have and have-not culture in the U.S.; but much like the provisioning of other technological goods such as the telephone, it has not altered the landscape of power relations by race and gender.

Search needs to be reconciled with the critical necessity of closing the digital divide, since search is such a significant part of mediating the online experience. Digital-divide scholars have argued that increased culturally relevant engagements with technology, web presence, and skill building will contribute to greater inclusion and to greater social, political, and economic agency for historically underrepresented, marginalized, and oppressed groups.²¹ This is the thrust of the neoliberal project of “uplift” and “empowerment”—by closing the skill-based gaps in computer programming, for example. These approaches do not account for the political economy and corporate mechanisms at play, and we must ask how communities can intervene to directly shape the practices of market-dominant and well-established technology platforms that are mediating most of our web interaction.²² They also often underexamine the diasporic labor conditions facing Black women who are engaged in the raw-mineral extraction process to facilitate the manufacture of computer and mobile phone hardware. I raise this issue because research on the global digital divide, and Google’s role in it,²³ must continue to expand to include a look at the ways that Black people in the U.S. and abroad are participating and, in the case of the United States, *not* participating to a significant degree in information and communication technology industries.²⁴ This makes calls for “prosumer” participation,²⁵ as a way of conceptualizing how Black people can move beyond being simple consumers of digital technologies to producers of technological output, a far more complex discussion.

George Ritzer and Nathan Jurgenson at the University of Maryland characterize this emphasis of merging the consumptive and productive aspects of digital engagement as “a trend toward unpaid rather than paid labor and toward offering products at no cost, and the system is marked by a new abundance where scarcity once predominated.”²⁶ The critical communications scholar Dallas Smythe describes this type of prosumer-

ism as “the audience as commodity,” where users are sold to advertisers as a commodity and, in return for “free” services, users are explicitly exposed to advertising.²⁷ Christian Fuchs, the director of the Communication and Media Research Institute and Westminster Institute for Advanced Studies, discusses this accumulation strategy, bolstered by Google’s users, as a process of both prosumer commodity and audience commodity by virtue of the decentralized nature of the web.²⁸ The intensive participation of people in uploading, downloading, sharing, tagging, browsing, community building, and content generation allows for mass distribution and one-to-many or many-to-many engagements in a way that traditional media could not have done due to its centralized nature.²⁹ In Fuchs’s work on the political economy of Google, he characterizes the unpaid, user-generated content provided by its users as the basis for Google’s ability to conduct keyword searching because it indexes all user-generated content and “thereby acts as a meta-exploiter of all user-generated content producers.”³⁰ Surplus labor is created for Google through users’ engagements with its products, from Gmail to Google Scholar, the reading of blogs in Blogger/Blogspot, the use of Google Maps or Google Earth, or the watching of videos on YouTube, among many of the company’s services.³¹ The vertical offerings of Google are so great,³² coupled with its prioritization of its own properties in keyword searches, that mere use of any of these “free” tools creates billion-dollar profits for Google—profits generated from both unpaid labor from users and the delivery of audiences to advertisers. Fuchs’s work explicitly details how Google’s commodities are not its services such as Gmail or YouTube; its commodities are all of the content creators on the web whom Google indexes (the prosumer commodity) and users of their services who are exposed to advertising (audience commodity).

We are the product that Google sells to advertisers.

These aspects of software and hardware development are important, and decreased engagements of women and people of color in the high-tech design sector, coupled with increased marginalized participation in the most dangerous and volatile parts of the information and communication technology labor market, have impact on the artifacts such as search results themselves. According to U.S. Department of Labor workforce data obtained by the *Mercury News* through a Freedom of Information request, of the 5,907 top managers in the Silicon Valley of-

fices of the ten largest high-tech companies in 2005, 296 were Black or Hispanic, a 20% decline from 2000.³³ Though the scope of this book does not include a formal interrogation of Black manufacturing labor migration to outsourced ICT manufacturing outside the United States, it is worth noting that this phenomenon has implications for participation in industries that shape everything from hardware to software design, of which Google is playing a primary role. As of July 1, 2016, Google’s own diversity scorecard shows that only 2% of its workforce is African American, and Latinos represent 3%. With all of the aberrations and challenges that tech companies face in charges of data discrimination, the possibility of hiring recent graduates and advanced-degree holders in Black studies, ethnic studies, American Indian studies, gender and women’s studies, and Asian American studies with deep knowledge of history and critical theory could be a massive boon to working through the kinds of complex challenges facing society, if this is indeed the goal of the technocracy. From claims of Twitter’s racist trolling that drives people from its platform³⁴ to charges that Airbnb’s owners openly discriminate against African Americans who rent their homes³⁵ to racial profiling at Apple stores in Australia³⁶ and Snapchat’s racist filters,³⁷ there is no shortage of projects to take on in sophisticated ways by people far more qualified than untrained computer engineers, whom, through no fault of their own, are underexposed to the critical thinking and learning about history and culture afforded by the social sciences and humanities in most colleges of engineering nationwide. The lack of a diverse and critically minded workforce on issues of race and gender in Silicon Valley impacts its intellectual output.

Google is a powerful and important resource for organizing information and facilitating social cooperation and contact, while it simultaneously reinforces hegemonic narratives and exploits its users. This has widely been characterized by critical media scholars as a dialectic that has less to do with Google’s technologies and services and more to do with the organization of labor and the capitalist relations of production.³⁸ The notion that Google/Alphabet has the potential to be a democratizing force is certainly laudable, but the contradictions inherent in its projects must be contextualized in the historical conditions that both create it and are created by it. Thinking about the specifics of who benefits from these practices—from hiring to search results

to technologies of surveillance—these are problems and projects that are not equally experienced. I have written, for example, with my colleague Sarah T. Roberts about the myriad problems with a project such as Google Glass and the problems of class privilege that directly map to the failure of the project and the intensifying distrust of Silicon Valley gentrifiers in tech corridors such as San Francisco and Seattle.³⁹ The lack of introspection about the public wanting to be surveilled at the level of intensity that Google Glass provided is part of the problem: centuries-old concepts of conquest and exploration of every landscape, no matter its inhabitants, are seen as emancipatory rather than colonizing and totalizing for people who fall within its gaze. People on the street may not characterize Google Glass as a neocolonial project in the way we do, but they certainly know they do not like seeing it pointed in their direction; and the visceral responses to Google Glass wearers as “Glassholes” is just one indicator of public distrust of these kinds of privacy intrusions.

The neocolonial trajectories are not just in products such as search or Google Glass but exist throughout the networked economy, where some people serve as the most exploited workers, including child and forced laborers,⁴⁰ in such places as the Democratic Republic of Congo, mining ore called columbite-tantalite (abbreviated as “coltan”) to provide raw materials for companies such as Nokia, Intel, Sony, and Ericsson (and now Google)⁴¹ that need such minerals in the production of components such as tantalum capacitors, used to make microprocessor chips for computer hardware such as phones and computers.⁴² Others in the digital-divide network serve as supply-chain producers for hardware companies such as Apple⁴³ or Dell,⁴⁴ and this outsourced labor from the U.S. goes to low bidders that provide the cheapest labor under neoliberal economic policies of globalization.

To review, in the ecosystem, Black people provide the most grueling labor for blood minerals, and they do the dangerous, toxic work of dismantling e-waste in places such as Ghana, where huge garbage piles of poisonous waste from discarded electronics from the rest of the world are shipped. In the United States, Black labor is for the most part bypassed in the manufacturing sector, a previous site of more stable unionized employment, due to electronics and IT outsourcing to Asia. African American identities are often a commodity, exploited as titillating fodder in a network that traffics in racism, sexism, and homophobia for

profit. Meanwhile, the onus for change is placed on the backs of Black people, and Black women in the United States in particular, to play a more meaningful role in the production of new images and ideas about Black people by learning to code, as if that alone could shift the tide of Silicon Valley’s vast exclusionary practices in its products and hiring.

Michele Wallace, a professor of English at the City College of New York and the Graduate Center of the City University of New York (CUNY), notes the crisis in lack of management, design, and control that Black people have over the production of commercial culture. She states that under these conditions, Black people will be “perpetual objects of contemplation, contempt, derision, appropriation, and marginalization.”⁴⁵ Janell Hobson at the University of Albany draws important attention to Wallace’s commentary on Black women as creative producers and in the context of the information age. She confirms this this confluence of media production on the web is part of the exclusionary terrain for Black women, who are underrepresented in many aspects of the information industry.⁴⁶ I would add to her argument that while it is true that the web can serve as an alternative space for conceiving of and sharing empowered conceptions of Black people, this happens in a highly commercially mediated environment. It is simply not enough to be “present” on the web; we must consider the implications of what it means to be on the web in the “long tail” or mediated out of discovery and meaningful participation, which can have a transformative impact on the enduring and brutal economic and social disenfranchisement of African Americans, especially among Black women.

Social Inequality Will Not Be Solved by an App

An app will not save us. We will not sort out social inequality lying in bed staring at smartphones. It will not stem from simply sending emails to people in power, one person at a time. New, neoliberal conceptions of individual freedoms (especially in the realm of technology use) are oversupported in direct opposition to protections realized through large-scale organizing to ensure collective rights. This is evident in the past thirty years of active antilabor policies put forward by several administrations⁴⁷ and in increasing hostility toward unions and twenty-first-century civil rights organizations such as Black Lives Matter. These

proindividual, anticomunity ideologies have been central to the anti-democratic, anti-affirmative-action, antiwelfare, antichoice, and antirace discourses that place culpability for individual failure on moral failings of the individual, not policy decisions and social systems.⁴⁸ Discussions of institutional discrimination and systemic marginalization of whole classes and sectors of society have been shunted from public discourse for remediation and have given rise to viable presidential candidates such as Donald Trump, someone with a history of misogynistic violence toward women and anti-immigrant schemes. Despite resistance to this kind of vitriol in the national electoral body politic, society is also moving toward greater acceptance of technological processes that are seemingly benign and decontextualized, as if these projects are wholly apolitical and without consequence too. Collective efforts to regulate or provide social safety nets through public or governmental intervention are rejected. In this conception of society, individuals make choices of their own accord in the free market, which is normalized as the only legitimate source of social change.⁴⁹

It is in this broader social and political environment that the Federal Communications Commission and Federal Trade Commission have been reluctant to regulate the Internet environment, with the exception of the Children's Internet Protection Act⁵⁰ and the Child Safe Viewing Act of 2007.⁵¹ Attempts to regulate decency vis-à-vis racist, sexist, and homophobic harm have largely been unaddressed by the FCC, which places the onus for proving harm on the individual. I am trying to make the case, through the mounting evidence, that unregulated digital platforms cause serious harm. Trolling is directly linked to harassment offline, to bullying and suicide, to threats and attacks. The entire experiment of the Internet is now with us, yet we do not have enough intense scrutiny at the level of public policy on its psychological and social impact on the public.

The reliability of public information online is in the context of real, lived experiences of Americans who are increasingly entrenched in the shifts that are occurring in the information age. An enduring feature of the American experience is gross systemic poverty, whereby the largest percentages of people living below the poverty line suffering from un- and underemployment are women and children of color. The economic crisis continues to disproportionately impact poor people of

color, especially Black / African American women, men, and children.⁵² Furthermore, the gap between Black and White wealth has become so acute that a recent report by Brandeis University found that this gap quadrupled between 1984 and 2007, making Whites five times richer than Blacks in the U.S.⁵³ This is not the result of moral superiority; this is directly linked to the gamification of financial markets through algorithmic decision making. It is linked to the exclusion of Blacks, Latinos, and Native Americans from the high-paying jobs in technology sectors. It is a result of digital redlining and the resegregation of the housing and educational markets, fueled by seemingly innocuous big-data applications that allow the public to set tight parameters on their searches for housing and schools. Never before has it been so easy to set a school rating in a digital real estate application such as Zillow.com to preclude the possibility of going to "low-rated" schools, using data that reflects the long history of separate but equal, underfunded schools in neighborhoods where African Americans and low-income people live. These data-intensive applications that work across vast data sets do not show the microlevel interventions that are being made to racially and economically integrate schools to foster educational equity. They simply make it easy to take for granted data about "good schools" that almost exclusively map to affluent, White neighborhoods. We need more intense attention on how these types of artificial intelligence, under the auspices of individual freedom to make choices, forestall the ability to see what kinds of choices we are making and the collective impact of these choices in reversing decades of struggle for social, political, and economic equality. Digital technologies are implicated in these struggles.

These dramatic shifts are occurring in an era of U.S. economic policy that has accelerated globalization, moved real jobs offshore, and decimated labor interests. Claims that the society is moving toward greater social equality are undermined by data that show a substantive decrease in access to home ownership, education, and jobs—especially for Black Americans.⁵⁴ In the midst of the changing social and legal environment, inventions of terms and ideologies of "colorblindness" disingenuously purport a more humane and nonracist worldview.⁵⁵ This is exacerbated by celebrations of multiculturalism and diversity that obscure structural and social oppression in fields such as education and information sciences, which are shaping technological practices.⁵⁶ Research by Sha-

ron Tettegah, a professor of education at the University of Nevada, Las Vegas, shows that people invested in colorblindness are also less empathetic toward others.⁵⁷ Making race the problem of those who are racially objectified, particularly when seeking remedy from discriminatory practices, obscures the role of government and the public in solving systemic issues.⁵⁸

Central to these “colorblind” ideologies is a focus on the inappropriateness of “seeing race.” In sociological terms, colorblindness precludes the use of racial information and does not allow any classifications or distinctions.⁵⁹ Yet, despite the claims of colorblindness, research shows that those who report higher racial colorblind attitudes are more likely to be White and more likely to condone or not be bothered by derogatory racial images viewed in online social networking sites.⁶⁰ Silicon Valley executives, as previously noted, revel in their embrace of colorblindness as if it is an asset and not a proven liability. In the midst of reenergizing the effort to connect every American and to stimulate new economic markets and innovations that the Internet and global communications infrastructures will afford, the real lives of those who are on the margin are being reengineered with new terms and ideologies that make a discussion about such conditions problematic, if not impossible, and that place the onus of discriminatory actions on the individual rather than situating problems affecting racialized groups in social structures.⁶¹

Formulations of postracialism presume that racial disparities no longer exist, a context within which the colorblind ideology finds momentum.⁶² George Lipsitz, a critical Whiteness scholar and professor at the University of California, Santa Barbara, suggests that the challenge to recognizing racial disparities and the social (and technical) structures that instantiate them is a reflection of the possessive investment in Whiteness—which is the inability to recognize how White hegemonic ideas about race and privilege mask the ability to see real social problems.⁶³ I often challenge audiences who come to my talks to consider that at the very historical moment when structural barriers to employment were being addressed legislatively in the 1960s, the rise of our reliance on modern technologies emerged, positing that computers could make better decisions than humans. I do not think it a coincidence that when women and people of color are finally given opportunity to participate in limited spheres of decision making in society, computers are

simultaneously celebrated as a more optimal choice for making social decisions. The rise of big-data optimism is here, and if ever there were a time when politicians, industry leaders, and academics were enamored with artificial intelligence as a superior approach to sense-making, it is now. This should be a wake-up call for people living in the margins, and people aligned with them, to engage in thinking through the interventions we need.

Conclusion



Algorithms of Oppression

We have more data and technology than ever in our daily lives and more social, political, and economic inequality and injustice to go with it. In this book, I have sought to critique the political-economic framework and representative discourse that surrounds racial and gendered identities on the web, but more importantly, I have shined a light on the way that algorithms are value-laden propositions worthy of our interrogation. I am particularly mindful of the push for digital technology adoption by Black / African Americans, divorced from the context of how digital technologies are implicated in global racial power relations. I have tried to show how traditional media misrepresentations have been instantiated in digital platforms such as search engines and that search itself has been interwoven into the fabric of American culture. Although rhetoric of the information age broadly seeks to disembody users, or at least to minimize the hegemonic backdrop of the technological revolution, African Americans have embraced, modified, and contextualized technology into significantly different frameworks despite the relations of power expressed in the socio-algorithms. This book can open up a dialogue about radical interventions on socio-technical systems in a more thoughtful way that does not further marginalize people who are already in the margins. Algorithms are, and will continue to be, contextually relevant and loaded with power.

Toward an Ethical Algorithmic Future

This book opens up new lines of inquiry using what I believe can be a black feminist technology studies (BFTS) approach to Internet research. BFTS could be theorized as an epistemological approach to researching gendered and racialized identities in digital and analog

media studies, and it offers a new lens for exploring power as mediated by intersectional identities. More research on the politics, culture, and values embedded in search can help frame a broader context of African American digital technology usage and early adoption, which is largely underexamined, particularly from the perspectives of women and girls. BFTS is a way to bring more learning beyond the traditional discourse about technology consumption—and lack thereof—among Black people. Future research using this framework can surface counter-narratives about Black people and technology and can include how African American popular cultural practices are influencing non-African American youth.¹ Discourses about African Americans and women as technologically illiterate are nothing new, but dispelling the myth of Blacks / African Americans as marginal to the broadest base of digital technology users can help us define new ways of thinking about motivations in the next wave of technology innovation, design, and, quite possibly, resistance.

Algorithms and Invisibility: My Interview with Kandis

Most of the attention to the protection of online information has been argued legally as a matter of “rights.” Rights are a type of property, or entitlement, that function on the web through a variety of narratives, such as “free speech” and “freedom of expression,” all of which are constitutionally protected in the United States. The framing of web content and ownership of web URLs as “property” afforded private protections is of consequence for individuals, as noted in Jessie Daniels’s aforementioned work, which documents the misrepresentation of Dr. Martin Luther King Jr. at the site martinlutherking.org, a cloaked website managed by neo-Nazis and White supremacists at Stormfront.² Private ownership of identity on the web is a matter of who can pay and who lines up quickly enough to purchase identity markers that establish a type of official record about a person or a group of people. Indeed, anyone can own anyone else’s identity in the current digital landscape. The right to control over group and personal identity and memory must become a matter of concern for archivists, librarians, and information workers, and a matter of internet regulation and public policy.

In concluding this book, I want to extend an example beyond Google to look closely at the consequences of the lack of identity control on another platform: Yelp.

Kandis has been in business for thirty years, and her primary clients are African American. This is her story, which elucidates in a very personal way how algorithmic oppression works and is affecting her very quality of life as a small business owner who runs the only local African American hair salon within a predominantly White neighborhood, located near a prestigious college town in the United States:

When I first came and opened up my shop here, there was a strong African American community. There were Black sororities and fraternities, and they had step shows, which no longer exist anymore! The Black Student Union organization was very strong; it was the ‘80s. Everyone felt like family, and everyone knew each other. Even though I only worked in this part of town, it was almost like I went to school there too. We all knew each other, and everyone celebrated each other’s success.

I often get invited to participate in the major events and celebrations—from marriages to their parents’ funerals. For instance, I have several clients from the ‘80s who I still service to this day. Here it is twenty years later, and now I’m servicing their sons’ and daughters’ hair, who may or may not attend the same university. The relationships are intact and so strong. It’s not uncommon, even if we aren’t always in touch for the past several years, for clients who may live on the other side of the country to send a recommendation my way! I have worked in this community for thirty years. I know a lot of people. They all have to get their hair done.

I asked Kandis how has that changed:

Well, prior to the changes the Internet brought about, I never had to do much advertising because my name carried weight through the campus and the graduate school. So, if you were a Black prelaw or premed student, I would know who you were, pretty much as you were coming into the school, because people would tell them about me, and then they would book an appointment with me.

But now, since there are only a small number of African Americans at the university, and those few people aren’t looking up at each other, we are los-

ing the art of conversation and how we used to verbally pass information. So my name started dying down. It kind of reminds me of how in elementary school, there are songs that kids sing, and they continue through generations. All kids know certain nursery rhymes and songs, even the ones their parents sang when they were little. It's like those song are trapped in time.

We, as African Americans, are storytellers. My name is not being talked about anymore, because young people don't talk to each other anymore. I was able to afford a modest lifestyle. Prior to these changes, from the stock market crashing and now this way of doing business through technology, life has become an uphill battle, and I felt like I was drowning. I've actually considered leaving, but where would I go? Where would I go?

Look, I'm very much used to diversity, but when the campus stopped admitting so many Blacks, I became a minority in a way that I had never felt before. You think of the university as being a part of the community, and the community would benefit from the university. But I don't think they thought about what would happen to all the businesses who supported those students. Where would the students and faculty go to get their needs met, like their hair cared for? I mean, other students can go down the street to anyone, but the Black students have to have a car to travel thirty minutes across town? Why are they required to have a car or transportation when no one else needs that to get their hair done?

Kandis was directly impacted by the shifts away from affirmative action that decimated the admission of African Americans to colleges and universities over the past twenty years:

Sometimes people are in a highly competitive arena, and they need to go to a nonjudgmental place where they can be themselves and where they don't have to apologize for the way they speak or their culture or wonder if they go to a Caucasian hair stylist, if they can handle their hair.

To be a Black woman and to need hair care can be an isolating experience.

The quality of service I provide touches more than just the external part of someone. It's not just about their hair. A lot of people are away from their families, and they need someone to trust who will support them. People do like to be recommended to someone they can trust.

I asked Kandis how the Internet changed things for her and her business:

Back then, when I had many more clients, there was no Internet. A personal computer was almost like a concept but not a reality for most people. At the beginning, you had a lot of younger people on the cutting edge of computers or who were more computer savvy and up on the current trend. My generation was a little bit slower to comprehend and participate. The Internet has also developed a new age of the infomercial and the how-to-do-it-yourself university!

The Internet is now showing everyone how to do it themselves and how to cut out the middleman. They have created the new consumers. New consumers have less value for small businesses; they think just because they watch the new infomercial, they can do it themselves, buy it themselves, make it themselves. Also, because you can purchase everything for yourself, the new consumers now feel entitled. They feel no shame with coming in and snapping photographs. They collect your hard work, all your information in two seconds, by taking pictures of all of my products so they can go and purchase them online instead of from me.

When things started changing so fast technologically, using the computer was an easier transition for me, because I had taken a few computer [Mac] classes on campus, and I was able to adapt to what was going on. Because of this, I was familiar, and I was comfortable with exploring the unknown.

I quickly realized the Internet/Yelp told people that I did not exist.

I think algorithms don't take into consideration communities of color and their culture of trusting in the web with our personal information and that we are dealing with things that won't even allow us to give someone a review. We don't like to give our personal information out like that. And just because I don't have reviews doesn't mean that I have no value. Because the computer, or I guess the Internet/Yelp, is redefining who is valuable and where the value lies, and I believe this is false. It's not right.

The algorithm shouldn't get to decide whether I exist or not. So I had to figure it out, within my financial limitations, because now it becomes another financial burden to stay relevant in eyes of what the web is telling people about who is valuable. I had to be creative and spend a lot of time

on the computer trying to figure what was the least expensive way to be visible with the most impact.

So, when I discovered Yelp and its alleged benefits—because I don't think it really benefited me—I was forced to participate in Yelp.

I asked what that participation with Yelp was like.

They tell you that everything is free, like they are doing a community service, but later on, it's basically pay to play. They call on a regular basis to get you to spend a few hundred dollars a month to advertise with them, and if you don't, they are going to push you further down their pile. I can be sitting in my chair searching for myself and not find me or find me ten pages later. I can type in every keyword, like "African American," "Black," "relaxer," "natural," as keywords, and White businesses, White hairdressers, or White salons would clearly come up before me—along with people who have not been in business as long as me. I think they need to put in how long someone has been in business in that algorithm, because I don't think it's fair that people who are brand new are popping up before those of us who may or may not be on the Internet but have more experience and are more established.

And another thing, Black people don't "check in" and let people know where they're at when they sit in my chair. They already feel like they are being hunted; they aren't going to tell The Man where they are. I have reviews from real clients that they put into a filter because it doesn't meet their requirements of how they think someone should review me.

I asked her to tell me more about that.

I think Yelp looks at people as *their* clients, not mine. If they are your clients who are loyal to your business and you, they are not interested. They want to market to my clients, and if you review me and you've never reviewed any other businesses, they are not going to take you as a serious voice on Yelp. What is that about? They are selling to the reviewers. Since I am the only Black person in this neighborhood doing hair, that should tip the scale in my favor, but it doesn't. If they were honest, it would mean I would go to the top. But they are promoting and biasing the lay of the land in this area, which is causing me more harm by making it look like I don't exist.

I have been on Facebook and pulled up some folks who work at Yelp, and from my perspective and from what I saw, there weren't that many Black people. It wasn't diverse. You can see everyone on FB, and these people are not Black people. And that's a problem, because how would they know to even consider the needs of a minority or what our language is? You are telling us we have to use certain keywords, and you don't even know our language, because you think that "Black hair" means hair *color*, not *texture*! We don't call each other African American; society calls us that. Do you know what I mean? We are Black.

You know, they locked me out of Yelp. When I sent them an email asking them why I wasn't on the first page of Yelp and why, when I'm sitting in my chair, I can't find myself and why, when I used certain keywords, I couldn't find myself. I told them that by doing that, they are suggesting I don't exist. At that time, they put most of my reviews in a filter and locked me out for about four months. Every time you make a change on Yelp, there is someone checking you. If you go into your page, there is someone looking at what you are doing. I know that because if you get people inquiring about you, Yelp will call you and try to get advertising from you. And they will say that they see people trying to connect with you through Yelp, and then they will try to sell you advertising. They will try to show their value by saying they can help you get more business.

I used to have my own page, but now you have a third of a page with people who are similar to you. And if they don't choose you, they are showing your competition. For a fee, they will remove your competition, but otherwise they are showing your competition on your own page! You don't have your own page anymore; you are on a page with advertising and your competition, who are similar to your style, and they will put you up against them while searching on your own business.

They'd rather put other salons in other parts of the city to avoid driving the clients to someone who is not paying for advertisement. So I would do something like put up a picture of myself, as this was something that they suggested, and I use keywords to be found, but that doesn't help now.

Before, Yelp would encourage the business owner to upload a photo of themselves, and that was great for me. This is when being in the minority would help me stand out. That didn't last long because they stopped showing your head shot, and now they put up a map instead of allowing you to use your own image. Your own image isn't shown. Before you get to my photos and my reviews, there are suggestions to other people who

could be as far as five miles or more away. The first one is laser hair removal and extensions. I don't do hair removal or extensions.³

Kandis pulled out her mobile phone and walked me through her Yelp page and showed me how it works.

They are already advertising against me. At the end of the page, there are more competitors. I have to pay to take that off. My reviews are in the middle. You can't control the photos anymore. They have a "people who viewed this also viewed . . ." section, and it goes to other businesses. I have to pay to get that taken off. They have a few reviews now that are blocked, that they felt that I had asked my clients to do, and because these people haven't reviewed other people, they don't show.

So if you get two reviews in one day and haven't had any in six months, they think you have done something, and they block my reviews.⁴ The basic principle of Yelp is to supposedly lead people with unbiased decisions when choosing a good business. How? You tell me? Can they honestly do this when they're in the business of selling the advertisement?

They control the algorithm, which controls who can write the reviews. All this has a major influence on where you're placed on the list. You hope and pray that your customers, who may be from a different generation or culture, will participate in their construct. It just isn't as random as one may think.

There is no algorithm that can replace human dignity. They created a system that simulates a value, based on their own algorithm, so Yelp can be the number-one beneficiary. When companies like Yelp shake the tree for low-hanging fruit, this affects mostly small businesses and the livelihoods of real people who will never work for corporate America. The key is to be independent of these companies, because they never stop. They have new goals, and they come back with new visions. And it's not like a real contract where you can argue and negotiate. The scale is unbalanced; you can't negotiate.

I verified all of the claims that Kandis was making by visiting her page and the pages of other small business to see how they placed her competitors. Indeed, several times when I thought I was clicking on reviews of her small business or getting more information, I was instead clicking on competitors and led away from the business I was investigating. I share Kandis's experience to demonstrate the way that both the interface

and the algorithmic design is taking on new dimensions of control and influence over her representation and access to information about her business. She has so little ability to impact the algorithm, and when she tries, the company subverts her ability to be racially and gender recognized—a type of recognition that is essential to her success as a business owner. The attempts at implementing a colorblind algorithm in lieu of human decision making has tremendous consequences. In the case of Kandis, what the algorithm says and refuses to say about her identity and the identity of her customers has real social and economic impact.

Imagining Alternatives: Toward Public Noncommercial Search

Neoliberal impulses in the United States to support market-driven information portals such as Google Search have consequences for finding high-quality information on the Internet about people and communities, since this is the primary pathway to navigating the web. This is one of the many contradictions of the current for-profit search and cloud-computing industry. Future research efforts might address questions that can help us understand the role of the design of platforms, interfaces, software, and experiences as practices that are culturally and gender situated and often determined by economic imperatives, power, and values. Such an agenda could forward a commitment to ensuring that pornographic or exploitive websites do not stand as the default identification for women on the web. Despite a climate wherein everything driven by market interests is considered the most expedient and innovative way of generating solutions, we see the current failings. Calling attention to these practices, however unpopular it might be, is necessary to foster a climate where information can be trusted and found to be reliable. What is needed is a decoupling of advertising and commercial interests from the ability to access high-quality information on the Internet, especially given its increasing prominence as the common medium in the United States.

When using a digital media platform, be it Google Search or Yelp or some other ranking algorithmic decision's default settings, it is possible to believe that it is normal to see a list of only a handful of possible results on the first page of a search, but this "normal" is a direct result of the way that human beings have consciously designed both software and hardware to function this way and no other.

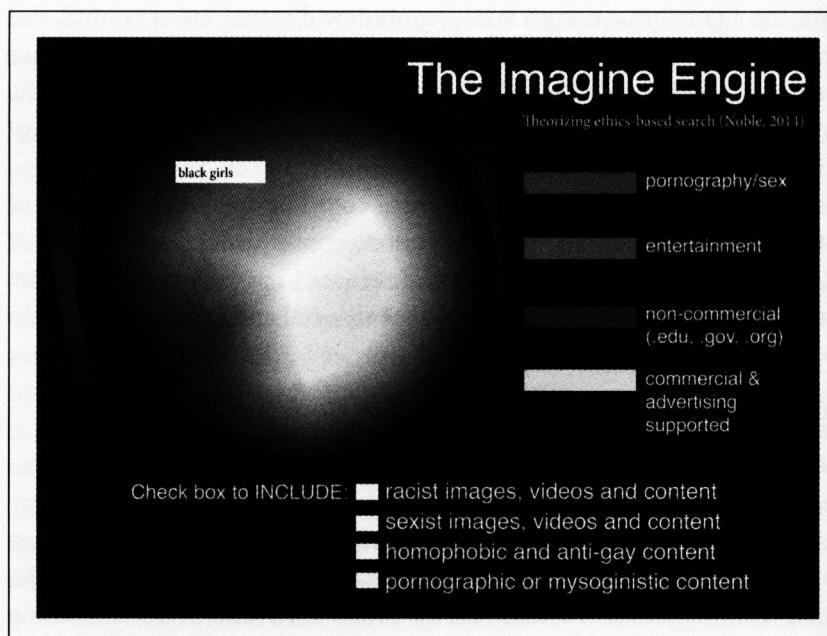


Figure C.1. S. U. Noble's interface of transparency: the imagine engine.

Imagine instead that all of our results were delivered in a visual rainbow of color that symbolized a controlled set of categories such that everything on the screen that was red was pornographic, everything that was green was business or commerce related, everything orange was entertainment, and so forth. In this kind of scenario, we could see the entire indexable web and click on the colors we are interested in and go deeply into the shades we want to see. Indeed, we can and should imagine search with a variety of other possibilities. In my own imagination and in a project I am attempting to build, access to information on the web could be designed akin to the color-picker tool or some other highly transparent interface, so that users could find nuanced shades of information and easily identify the borderlands between news and entertainment, or entertainment and pornography, or journalism and academic scholarship. In this scenario, I might also be able to quickly identify the blogosphere and personal websites.

Such imaginings are helpful in an effort to denaturalize and reconceptualize how information could be provided to the public vis-à-vis

the search engine. In essence, we need greater transparency and public pressure to slow down the automation of our worst impulses. We have automated human decision making and then disavowed our responsibility for it. Without public funding and adequate information policy that protects the rights to fair representation online, an escalation in the erosion of quality information to inform the public will continue.

Where Are Black Girls Now?

Since I began the pilot study in 2010 and collected data through 2016, some things have changed. In 2012, I wrote an article for *Bitch Magazine*, which covers popular culture from a feminist perspective, after some convincing from my students that this topic is important to all people—not just Black women and girls. I argued that we all want access to credible information that does not foster racist or sexist views of one another. I cannot say that the article had any influence on Google in any definitive way, but I have continued to search for Black girls on a regular basis, at least once a month. Within about six weeks of the article hitting newsstands, I did another search for “black girls,” and I can report that Google had changed its algorithm to some degree about five months after that article was published. After years of featuring pornography as the primary representation of Black girls, Google made modifications to its algorithm, and the results as of the conclusion of this research can be seen if figure C.2.

No doubt, as I speak around the world on this subject, audiences are often furiously doing searches from their smart phones, trying to reconcile these issues with the momentary results. Some days they are horrified, and other times, they are less concerned, because some popular and positive issue or organization has broken through the clutter and moved to a top position on the first page. Indeed, as this book was going into production, news exploded of biased information about the U.S. presidential election flourishing through Google and Facebook, which had significant consequences in the political arena.

I encourage us all to take notice and to reconsider the affordances and the consequences of our hyperreliance on these technologies as they shift and take on more import over time. What we need now, more than ever, is public policy that advocates protections from the effects of unregulated and unethical artificial intelligence.

Google black girls

All Images Videos Shopping News More Search tools Sign In

About 301,000,000 results (0.42 seconds)

Black Girls Code
www.blackgirlscodes.com/ ~ Black Girls Code ~
 Black Girls Code, BlackGirlsCode, STEM education San Francisco, Technology training for girls, diversity learning, Social Entrepreneurship in San Francisco, ...

Black Girls Rock
www.blackgirlsrockinc.com/ ~
 2016 Black Girls Rock! HONOREES: Rihanna ... TUNE IN TO CELEBRATE A NIGHT OF BLACK GIRL MAGIC WITH Black Girls Rock!

Read More - Black Girls Rock
www.blackgirlsrockinc.com/home/ ~
 Amanda Sternberg is an actress, fashionista and social activist. She has been featured in many TV shows such as Hunger Games and Sleep Hollow. She also ...

In the news

 7-Year-Old Writes Book To Show Black Girls They Are Princesses
 Huffington Post - 2 days ago
 Morgan Taylor wants to make sure every little black and brown girl feels like royalty. Which is ...

Black Girls Code relaunching Saturday with gaming workshop
 Miami Herald - 51 mins ago

[More news for black girls](#)

Black Girls Rock! Homepage - BET.com
www.bet.com/shows/black-girls-rock.html ~ BET ~
 Check out all things Black Girls Rock! here. ... Rihanna Owns Her Role as a Rock Star - Black Girls Rock!, 2016, Highlights, Hilary Clinton, Beverly Bond ...

9 Types of Black Girls - YouTube
<https://www.youtube.com/watch?v=HEQYU6SIAz>
 Aug 12, 2015 - 30 LIKES FOR 9 TYPES OF BLACK FRIENDS!!!! These are 9 Types of Black Girls you are bound to come across in your lifetime. Most likely ...

(@darkskin.blackgirls) • Instagram photos and videos
https://www.instagram.com/darkskin_blackgirls/ ~
 Dedicated to DarkSkin Women #BlackLivesMatter Business inquiries: darkskin_blackgirls@yahoo.com DM For a Feature Tag them if you know their ...

Black Girls RUN!
blackgirlrun.com/ ~
 Welcome to the MOVEMENT | Black Girls RUN!Black Girls RUN

Black Girls Are Easy |
blackgirlseasay.com/ ~
 Black Girls Are Easy has been voted the top dating and relationship advice website of the year

Why Are Black Girls Disproportionally Pushed Out of Schools? - The ...
www.theatlantic.com/education/archive/2016/03/the...-of-black-girls-in.../473718/
 The "good girl" and "bad girl" dichotomy, as chronicled by Monique W. Morris in *Pushout: The Criminalization of Black Girls in Schools*, is a ...

Study: Black Girls Are Being Pushed Out of School : Code Switch : NPR
www.npr.org/sections/codeswitch/.../study-black-girls-are-being-pushed-out-of-school
 News surrounding a confrontation in a Baltimore school is raising new questions about the role race plays in discipline for black girls. Baltimore ...

I Have a Problem With #BlackGirlMagic - Black Girls Aren't Magical ...
www.elle.com/life-love/a33180/why-i-dont-love-blackgirlmagic/

Figure C.2. My last Google search on "black girls," June 23, 2016.

Epilogue



Between the time I wrote this book and the day it went into production, the landscape of U.S. politics was radically altered with the presidential defeat on November 8, 2016, of former secretary of state Hillary Clinton by Donald Trump. Within days, media pundits and pollsters were trying to make sense of the upset, the surprise win by Trump, particularly since Clinton won the popular vote by close to three million votes.

Immediately, there were claims that "fake news" circulating online was responsible for the outcome. Indeed, as I gave talks about this book in the weeks after the election, I could only note in my many public talks that "as I've argued for years about the harm toward women and girls through commercial information bias circulating through platforms like Google, no one has seemed to care until it threw a presidential election." Notably, one remarkable story about disinformation (patently false information intended to deceive) made headlines about the election results.

This new political landscape has dramatically altered the way we might think about public institutions being a major force in leveling the playing field of information that is curated in the public interest. And it will likely be the source of a future book that recontextualizes what information means in the new policy regime that ensues under the leadership of avowed White supremacists and disinformation experts who have entered the highest levels of public governance.

Agencies that could have played a meaningful role in supporting research about the role of information and research in society, including the Institute for Museum and Library Services, the National Endowment for the Humanities, and the National Endowment for the Arts, are all under the threat of being permanently defunded and dismantled as of the moment this book goes into production. In fact, public research universities are also facing serious threats in cuts to federal funding because of their lack of compliance with the new administration's policies. This has



Figure E.1. Google search for “final election results” leads to fake news. Source: *Washington Post*, November 14, 2016.

If you head to Google to learn the final results of the presidential election, the search engine helpfully walks through the final electoral vote tallies and number of seats won by each party in the House and Senate. Under that, Google lists some related news articles. At the top this morning, with an accompanying photo: a story arguing that Donald Trump won both the popular and electoral votes.

In the news:

FINAL ELECTION 2016 NUMBERS: TRUMP WON BOTH POPULAR (62.9 M / 42.2 M) AND ELECTORAL COLLEGE

Toronto - WordPress.com - 2 days ago
Chris Kuelback 45 OF 191316: <http://edition.com/election/results/trump-won.html>

Clinton vs. Trump: Popular Vote. Are There Still Uncounted Ballots?
Haley Goss - 25 hours ago
Lawrence County 2016 Election Results Final
Rigell - 25 hours ago

That's not true.

The Daily Show's Dan Amira noticed that numbers were being spread on social media that linked back to the “70 News” site. The 70 News article cites its source as this tweet.

Figure E.2. Google results on final election results incorrectly show Trump as the winner of the popular vote. Source: *Washington Post*, November 14, 2016.

Final #Election2016 numbers

#PopularVote:

#Trump: 62,972,226

#Clinton: 62,277,750

#ElectoralCollege vote

#Trump 306 #Clinton 232

Michael
@Koxinga8

[Follow](#)

Final #Election2016 Results By Numbers:-

#PopularVote

#electoralvotes

3:24 AM - 12 Nov 2016

4 1,506 1,327

Figure E.3. Circulation of false information on Twitter shows Trump as the winner of the popular vote, November 14, 2016.

so radically altered the research landscape to the political right that scientists and researchers marched on Washington, D.C., on April 22, 2017, in response to orders that government-funded scientists and researchers stop conducting and disseminating research to the public. The potential for such a precedent may extend to public research universities, or at least many faculty members are working under the premise that this may not be out of the realm of possibility over the next four to eight years.

In this book, I have argued that the neoliberal political and economic environment has profited tremendously from misinformation and mischaracterization of communities, with a range of consequences for the most disenfranchised and marginalized among us. I have also argued for increased nonprofit and public research funding to explore alternatives to commercial information platforms, which would have included support of noncommercial search engines that could serve the public and pay closer attention to the circulation of patently false or harmful information. In the current environment, I would be remiss if I did not acknowledge, on the eve of the publication of this book, that this may not be viable at all given the current policy environment that is unfolding.

My hope is that the public will reclaim its institutions and direct our resources in service of a multiracial democracy. Now, more than ever, we need libraries, universities, schools, and information resources that will help bolster and further expand democracy for all, rather than shrink the landscape of participation along racial, religious, and gendered lines. Information circulates in cultural contexts of acceptability. It is not enough to simply want the most accurate and credible information to rise to the top of a search engine, but it is certainly an important step toward impacting the broader culture of information use that helps us make decisions about the distribution of resources among the most powerful and the most disenfranchised members of our society.

In short, we must fight to suspend the circulation of racist and sexist material that is used to erode our civil and human rights. I hope this book provides some steps toward doing so.

NOTES

INTRODUCTION

1. Matsakis, 2017.
2. See Peterson, 2014.
3. This term was coined by Eli Pariser in his book *The Filter Bubble* (2011).
4. See Dewey, 2015.
5. I use phrases such as “the N-word” or “n*gger” rather than explicitly using the spelling of a racial epithet in my scholarship. As a regular practice, I also do not cite or promote non–African American scholars or research that flagrantly uses the racial epithet in lieu of alternative phrasings.
6. See Sweeney, 2009.
7. See Boyer, 2015; Craven, 2015.
8. See Noble, 2014.
9. The term “digital footprint,” often attributed to Nicholas Negroponte, refers to the online identity traces that are used by digital media platforms to understand the profile of a user. The online interactions are often tracked across a variety of hardware (e.g., mobile phones, computers, internet services) and platforms (e.g., Google’s Gmail, Facebook, and various social media) that are on the World Wide Web. Digital traces are often used in the data-mining process to profile users. A digital footprint can often include time, geographic location, and past search results and clicks that have been tracked through websites and advertisements, including cookies that are stored on a device or other hardware.
10. “Kandis” is a pseudonym.
11. See H. Schiller, 1996.

CHAPTER 1. A SOCIETY, SEARCHING

1. See UN Women 2013.
2. See Diaz, 2008; Segev, 2010; Nissenbaum and Introna, 2004.
3. See Olson, 1998; Berman, 1971; Wilson, 1968; and Furner, 2007.
4. See Daniels, 2009, 2013; Davis and Gandy, 1999.
5. See Halavais, 2009, 1–2.
6. See Angwin et al., 2016.
7. See O’Neil, 2016, 8.
8. See Levin, 2016.
9. See Kleinman, 2015.

10. The debates over Google as a monopoly were part of a congressional Antitrust Subcommittee hearing on September 21, 2011, and the discussion centered around whether Google is causing harm to consumers through its alleged monopolistic practices. Google has responded to these assertions. See Kohl and Lee, 2011.
11. See Ascher, 2017.
12. See Leonard, 2009.
13. See Daniels, 2009, 2013; Brock, 2009.
14. See Kendall, 2002.
15. See Brock, 2009.
16. See S. Harding, 1987, 7.
17. See chapter 2 for a detailed discussion of the “Jewish” disclaimer by Google.
18. See hooks, 1992; Harris-Perry, 2011; Ladson-Billings, 2009; Miller-Young, 2007; Sharpley-Whiting, 1999; C. M. West, 1995; Harris, 1995; Collins, 1991; Hull, Bell-Scott, and Smith, 1982.
19. See Collins, 1991; hooks, 1992; Harris, 1995; Crenshaw, 1991.
20. See Brock, 2007.
21. The “digital divide” is a narrative about the lack of connectivity of underserved or marginalized groups in the United States that stems from the National Telecommunications and Information Administration report on July 8, 1999, *Falling through the Net: Defining the Digital Divide*.
22. See Inside Google 2010.
23. See Fallows, 2005; Purcell, Brenner, and Rainie, 2012.
24. A detailed discussion of this subject can be found in a Google disclaimer about the results that surface when a user searches on the word “Jew.” The URL for this disclaimer (now defunct) was www.google.com/explanation.html.
25. Senate Judiciary Committee, Subcommittee on Antitrust, Competition Policy and Consumer Rights, 2011.
26. See Elad Segev’s work on Google and global inequality (2010).
27. A good discussion of the ways that Google uses crowdsourcing as an unpaid labor pool for projects such as Google Image Labeler can be found in the blog *Labor-tainment* at <http://labortainment.blogspot.com> (last accessed June 20, 2012).
28. See the work of Cameron McCarthy, professor of education at the University of Illinois at Urbana-Champaign (1994).
29. See Nissenbaum and Introna, 2004; Vaidhyanathan, 2011; Segev, 2010; Diaz, 2008; and Noble, 2014.
30. This process has been carefully detailed by Levene, 2006.
31. Blogger, Wordpress, Drupal, and other digital media platforms make the process of building and linking to other sites as simple as the press of a button, rather than having to know code to implement.
32. See Spink et al., 2001; Jansen and Pooch, 2001; Wolfram, 2008.
33. See Markey, 2007.
34. See Ferguson, Kreshel, and Tinkham, 1990.
35. See Wasson, 1973; Courtney and Whipple, 1983.

36. See Smith, 1981.
37. See Bar-Ilan, 2007.
38. Google’s official statement on how often it crawls is as follow: “Google’s spiders regularly crawl the Web to rebuild our index. Crawls are based on many factors such as PageRank”, links to a page, and crawling constraints such as the number of parameters in a URL. Any number of factors can affect the crawl frequency of individual sites. Our crawl process is algorithmic; computer programs determine which sites to crawl, how often, and how many pages to fetch from each site. We don’t accept payment to crawl a site more frequently.” See Google, “About Google’s Regular Crawling of the Web,” accessed July 6, 2012, <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=34439>.
39. Brin and Page, 1998a: 110.
40. Ibid.
41. Brin and Page, 1998b, 18, citing Bagdikian, 1983.
42. On June 27, 2012, the online news outlets *The Local* and *The Raw Story* reported on Google’s settlement of the claim based on concerns over linking the word “Jew” with popular personalities. See AFP, 2012.
43. Anti-Defamation League, 2004.
44. Ibid.
45. See Zittrain and Edelman, 2002.
46. See SEMPO, 2010.
47. A website dedicated to the history of web memes attributes the precursor to the term “Google bombing” to Archimedes Plutonium, a Usenet celebrity, who created the term “searchenginebombing” in 1997. For more information, see “Google Bombing,” Know Your Meme, accessed June 20, 2012, <http://knowyourmeme.com/memes/google-bombing>. Others still argue that the first Google bomb was created by Black Sheep, who associated the terms “French Military Victory” to a redirect to a mock page that looked like Google and listed all of the French military defeats, with the exception of the French Revolution, in which the French were allegedly successful in killing their own French citizens. The first, most infamous instance of Google bombing was the case of *Hugedisk* magazine linking the text “dumb motherfucker” to a site that supported George W. Bush. For more information, see Calore and Gilbertson, 2001.
48. Brin and Page note that in the Google prototype, a search on “cellular phone” results in PageRank making the first result a study about the risks of talking on a cell phone while driving.
49. See SEMPO, 2004, 4.
50. In 2003, the radio host and columnist Dan Savage encouraged his listeners to go to a website he created, www.santorum.com, and post definitions of the word “santorum” after the Republican senator made a series of antigay remarks that outraged the public.
51. See Hindman, 2009; Zittrain, 2008; Vaidhyanathan, 2011.
52. Steele and Iliinsky, 2010, 143.

53. See Hindman, 2009.
54. Ibid.
55. See Gulli and Signorini, 2005.
56. Federal Communications Commission, 2010.
57. Associated Press v. United States, 326 U.S. 1, 20 (1945). Diaz (2008) carefully traces the fundamental notion of deliberative democracy and its critical role in keeping the public informed, in the tradition of John Stuart Mill's treatise "On Liberty," which contends that democracy cannot flourish without public debate and discourse from the widest range of possible points of view.
58. See Van Couvering, 2004, 2008; Diaz, 2008; Noble, 2014; and Zimmer, 2009.
59. See Lev-On, 2008.
60. See Andrejevic, 2007.
61. See Goldsmith and Wu, 2006.
62. H. Schiller, 1996, 48.
63. See Fallows, 2005; Purcell, Brenner, and Rainie, 2012.
64. President Eisenhower forewarned of these projects in his farewell speech on January 17, 1961, when he said, "In the councils of government, we must guard against the acquisition of unwarranted influence, whether sought or unsought, by the military-industrial complex. The potential for the disastrous rise of misplaced power exists and will persist." Eisenhower, 1961.
65. Niesen, 2012.
66. The full report can be accessed at www.pewinternet.org.
67. See Epstein and Robertson, 2015.
68. Purcell, Brenner, and Rainie, 2012, 2. Pew reports these findings from a survey conducted from January 20 to February 19, 2012, among 2,253 adults, age eighteen and over, including 901 cellphone interviews. Interviews were conducted in English and Spanish. The margin of error for the full sample is plus or minus two percentage points.
69. Feuz, Fuller, and Stalder, 2011.
70. Google Web History is designed to track signed-in users' searches in order to better track their interests. Considerable controversy followed Google's announcement, and many online articles were published with step-by-step instructions on how to protect privacy by ensuring that Google Web History was disabled. For more information on the controversy, see Tsukayama, 2012. Google has posted official information about its project at <http://support.google.com/accounts/bin/answer.py?hl=en&answer=54068&topic=14149&ctx=topic> (accessed June 22, 2012).
71. Leigh Estabrook and Ed Lakner (2000) have conducted a national study on Internet control mechanisms used by libraries, which primarily consist of policies and user education rather than filtering. These policies and mechanisms are meant to deter users from accessing objectionable content, including pornography, but also other material that might be considered offensive.
72. See Corea, 1993; Dates, 1990; Mastro and Tropp, 2004; Stromer, Merrit, and Matabane, 1989.
73. The Chicago Urban League has developed a Digital Media Strategy that is specifically concerned with the content and images of Black people on the Internet. See the organization's website: www.thechicagourbanleague.org.
74. The NAACP Image Awards recognize positive images of Blacks in the media. See the organization's website: www.naacp.org.
75. See Hunt, Ramón, and Tran, 2016.
76. FreePress.org has a page dedicated to the issues of civil rights and media justice. See www.freepress.net/media_issues/civil_rights (accessed April 15, 2012).
77. The Federal Trade Commission is looking into the privacy issues facing Americans over Google's targeted and behavior-based advertising programs. It has also settled out of court over the Google book-digitization project, which was reported in the media as a "monopolistic online land grab" over public domain orphan works. See Yang and Easton, 2009.
78. See Roberts, 2016; Stone, 2010.
79. For more information, see Roberts, 2012.
80. Roberts, 2016.
81. See Heider and Harp, 2002; Gunkel and Gunkel, 1997; Pavlik, 1996; Kellner, 1995; Barlow, 1996.
82. See Heider and Harp, 2002.
83. Ibid., 289.
84. Berger, 1972, 64.
85. See Mayall and Russell, 1993, 295.
86. Gardner, 1980, 105–106.
87. Gunkel and Gunkel, 1997, 131.
88. Lipsitz, 1998, 370.
89. Ibid., 381.
90. See Mills, 2014.
91. See Winner, 1986; Pacey, 1983.
92. See Chouliaraki and Fairclough, 1999.
93. See Barlow, 1996.
94. See Segev, 2010.
95. Stepan, 1998, 28.
96. #Gamergate was an incident involving a group of anonymous harassers of women in the video-gaming industry, including Zoë Quinn and Brianna Wu, as well as the writer and critic Anita Sarkeesian, who faced death threats and threats of rape, among others. In response to challenges of white male supremacy, sexism, racism, and misogyny in video-game culture, many women video-game developers, feminists, and men supporting women in gaming were attacked online as well as stalked and harassed.

CHAPTER 2. SEARCHING FOR BLACK GIRLS

1. See Guynn, 2016.
2. See Hiles, 2015.

27. A comprehensive timeline of Edward Snowden's whistleblowing on the U.S. government's comprehensive surveillance program is detailed by the *Guardian* newspaper in MacAskill and Dance, 2013.
28. Tippman, 2015.
29. Kiss, 2015.
30. Goode, 2015.
31. See Robertson, 2016.
32. Ibid.

CHAPTER 5. THE FUTURE OF KNOWLEDGE IN THE PUBLIC

1. See the plan at "The Plan for Dartmouth's Freedom Budget: Items for Transformative Justice at Dartmouth," *Dartblog*, accessed August 9, 2017, www.dartblog.com/Dartmouth_Freedom_Budget_Plan.pdf.
2. Peet, 2016.
3. Ibid.
4. Ibid.
5. Qin, 2016.
6. Sanford Berman documents the sordid history of racist classification in the Library of Congress in his canonical work *Prejudices and Antipathies* (1971). A follow-up to his findings was written thirty years later by Steven A. Knowlton in the article "Three Decades since *Prejudices and Antipathies*: A Study of Changes in the Library of Congress Subject Headings" (2005).
7. Peet, 2016.
8. Furner, 2007, 148.
9. Ibid., 147.
10. Ibid., 169.
11. See Olson, 1998.
12. See Anderson, 1991, 37–46.
13. Hudson, 1996, 256; Anderson, 1991.
14. The first documented evidence of print culture is attributed to Chinese wood-block printing. See Hyatt Mayor, 1971, 1–4.
15. See Saracevic, 2009.
16. See Berman, 1971; Olson, 1998.
17. Berman, 1971, 15.
18. Ibid., 5.
19. See *ibid.*; Palmer and Malone, 2001.
20. See Berman, 1971, 5.
21. Ibid.
22. Olson, 1998, 233.
23. Ibid., 234.
24. Ibid., 234–235.
25. Ibid., 235.
26. Ibid.

27. See Cornell, 1992; Olson, 1998.
28. See Olson, 1998, 237.
29. See C. West, 1996, 84.
30. See Berman, 1971, 18.
31. Ibid., citing Mosse, 1966.
32. See Wilson, 1968, 6.
33. Berman, 1971, 19, citing Marshall, personal communication, June 23, 1970.
34. Ibid., 20.
35. Reidsma, 2016.
36. See Galloway, 2008.
37. See Galloway, Lovink, and Thacker, 2008.
38. See Galloway, 2008.
39. Battelle, 2005, 6.
40. See Brin and Page, 1998a.
41. Saracevic, 1999, 1054.
42. Ibid.
43. Saracevic, 2009, 2570.
44. See Bowker and Star, 1999.
45. See Saracevic, 2009.
46. See Brock, 2011.
47. Ibid., 1101.
48. See Fuchs, 2008.

CHAPTER 6. THE FUTURE OF INFORMATION CULTURE

1. See Federal Communications Commission, 2010
2. Ibid.
3. H. Schiller, 1996, 44.
4. Cohen, 2016.
5. See McChesney and Nichols, 2009; H. Schiller, 1996.
6. See Harris-Perry, 2011; hooks, 1992.
7. Arreola, 2010.
8. See the website of the Society of Professional Journalists Code of Ethics, www.spj.org.
9. See Darnton, 2009; Jeanneney, 2007.
10. See Jeanneney, 2007.
11. See *Authors Guild v. Google*, Case 1:05-cv-08136-DC, Document 1088, November 14, 2013.
12. Darnton, 2009, 2.
13. See *Search King v. Google*, 2003.
14. See Dickinson, 2010.
15. Ibid., 866.
16. Ibid.
17. Ingram, 2011.

18. See Wilhelm, 2006.
19. See Sinclair, 2004.
20. See Luyt, 2004.
21. See van Dijk and Hacker, 2003; and Pinkett, 2000.
22. See Rifkin, 2000.
23. See Segev, 2010.
24. See Rifkin, 1995.
25. The term “prosumer” is a portmanteau of “producer” and “consumer” that is often used to indicate a higher degree of digital literacy, economic participation, and personal control over the means of technology production. The term is mostly attributed, in this context, to Alvin Toffler, a futurist who thought that the line between traditional economic consumer and producer would eventually blur through engagements with technology and that this participation would generally lead to greater mass customization of products and services by corporations. See Toffler, 1970, 1980; Tapscott, 1996; Ritzer and Jurgenson, 2010.
26. Ritzer and Jurgenson, 2010, 14.
27. See Smythe, 1981/2006.
28. See Fuchs, 2011.
29. Ibid.
30. Ibid., 43.
31. Ibid.
32. A list of Google’s global assets and subsidiaries can be found in its SEC filings: www.sec.gov/Archives/edgar/data/1288776/000119312507044494/dex2101.htm.
33. See recent news coverage discussing U.S. Department of Labor data and the significant decline of Blacks, Latinos, and women in the Silicon Valley technology industries: Swift, 2010.
34. See Meyer, 2016.
35. See Glusac, 2016.
36. See Eddie and Prigg, 2015.
37. See Mosher, 2016.
38. See Fuchs, 2011.
39. See Noble and Roberts, 2015.
40. See Department of Labor, Office of the Secretary, “Notice of Final Determination Revising the List of Products Requiring Federal Contractor Certification as to Forced or Indentured Child Labor Pursuant to Executive Order 13126,” which prohibits coltan that has been produced by child labor from entering the United States.
41. Kristi Esseck covered this issue in her article “Guns, Money and Cell Phones” (2011). The United Nations also issued a report, submitted by Secretary General Kofi Annan, about the status of companies involved in coltan trafficking and the impact of investigations by the UN into the conflicts arising from such practices in the Democratic Republic of the Congo. The report can be accessed at www.un.org/Docs/journal/asp/ws.asp?m=S/2003/1027 (accessed July 3, 2012).

42. Coltan mining is significantly understudied by Western scholars but has been documented in many nongovernmental organizations’ reports about the near-slavery economy in the Congo that is the result of Western dependence on “conflict minerals” such as coltan that have been the basis of ongoing wars and smuggling regimes that have extended as far as Rwanda, Uganda, and Burundi. See reviews in the *New York Times* as well as a detailed overview of the conditions in the Congo due to mining by Anup Shah, at www.globalissues.org, which asserts that an elite network of multinational companies, politicians, and military leaders have essentially kept the issues from the view of the public. See Hardenaug, 2001; Shah, 2010.
43. While less formal scholarship has been dedicated to this issue, considerable media attention in 2011 and 2012 has been focused on the labor conditions in parts of China where Apple manufactures its products. While some of the details of the journalistic reporting have been prone to factual error in location and dates, there is considerable evidence that labor conditions by Apple’s supplier Foxconn are precarious and rife with human-rights abuses. See Duhigg and Barboza, 2012.
44. See Fields, 2004.
45. Wallace, 1990, 98.
46. See Hobson, 2008.
47. See Harvey, 2005.
48. See Jensen, 2005; Brown, 2003; Burdman, 2008.
49. See Harvey, 2005.
50. The Children’s Internet Protection Act (CIPA) was adopted by the FCC in 2001 and is designed to address filtering of pornographic content from any computers in federally funded agencies such as schools and libraries. The act is designed to incentivize such organizations with Universal E-Rate discounts for using filters and providing Internet safety policies. See FCC, “Children’s Internet Protection Act,” accessed August 9, 2017, www.fcc.gov/guides/childrens-internet-protection-act.
51. The Child Safe Viewing Act of 2007 is designed to regulate objectionable adult-themed material so that children cannot see it on mobile devices. The FCC is investigating the use of blocking software or devices for use on television and mobile devices through the use of a V-Chip that can allow adults to block content. See FCC, “Protecting Children from Objectionable Content on Wireless Devices,” accessed August 9, 2017, www.fcc.gov/guides/protecting-children-objectionable-content-wireless-devices.
52. The National Urban League reported in 2010 startling statistics about the economic crisis, specific to African Americans: (1) less than half of black and Hispanic families own a home (47.4% and 49.1%, respectively), compared to three-quarters of White families; and (2) Blacks and Hispanics are more than three times as likely as whites to live below the poverty line. See National Urban League, 2010.
53. See McGreal, 2010.

- 54. See Jensen, 2005; McGreal, 2010.
- 55. See Neville et al., 2012.
- 56. See Pawley, 2006.
- 57. See Tettegah, 2016.
- 58. See Brown, 2003; Crenshaw, 1991.
- 59. See Lipsitz, 1998; Brown, 2003; Burdman, 2008.
- 60. See Tynes and Markoe, 2010.
- 61. See Brown, 2003.
- 62. Ibid.
- 63. See Lipsitz, 1998; Jensen, 2005.

CONCLUSION

- 1. See Tate, 2003.
- 2. See Daniels, 2008.
- 3. This suppression of Kandis's business, according to her, is not based on her lack of popularity but, rather, on her unwillingness to pay more to Yelp to have her competitors taken off her page.
- 4. Kandis described an experience of having two of her clients who are not regular reviewers on Yelp post positive reviews about her, only to have them sequestered from her page. She described her conversations with Yelp customer service agents, from which she deduced that these reviews were seen as "fraudulent" or inauthentic reviews that she must have solicited.

REFERENCES

- AFP. (2012, May 23). Google's "Jew" Suggestion Leads to Judge Order. *The Local France*. Retrieved from thelocal.fr.
- Anderson, B. (1991). *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (2nd ed.). London and New York: Verso.
- Andrejevic, M. (2007). Surveillance in the Digital Enclosure. *Communication Review*, 10(4), 295–317.
- Angwin, J., Larson, J., Mattu, S., and Kirchner, L. (2016). Software Used to Predict Criminality Is Biased against Black People. *TruthOut*. Retrieved from www.truthout.org.
- Anti-Defamation League. (2004). ADL Praises Google for Responding to Concerns about Rankings of Hate Sites. Retrieved from www.adl.org.
- Apuzzo, M. (2015, July 22). Dylann Roof, Charleston Shooting Suspect, Is Indicted on Federal Hate Crimes. *New York Times*. Retrieved from www.nytimes.com.
- Arreola, V. (2010, October 13). Latinas: We're So Hot We Broke Google. *Ms. Magazine Blog*. Retrieved from www.msmagazine.com.
- Ascher, D. (2017). The New Yellow Journalism. Ph.D. diss., University of California, Los Angeles.
- Associated Press. (2013, January 16). Calif. Teacher with Past in Porn Loses Appeal. *USA Today*. Retrieved from www.usatoday.com.
- Associated Press v. United States. (1945). 326 U.S. 1, US Supreme Court.
- Bagdikian, B. (1983). *The Media Monopoly*. Boston: Beacon.
- Bar-Ilan, J. (2007). Google Bombing from a Time Perspective. *Journal of Computer-Mediated Communication*, 12(3), article 8. Retrieved from <http://jcmc.indiana.edu>.
- Barlow, J. P. (1996). A Declaration of the Independence of Cyberspace. Electronic Frontier Foundation. Retrieved from <http://projects.eff.org/barlow/Declaration-Final.html>.
- Barth, F. (1966). *Models of Social Organization*. London: Royal Anthropological Institute.
- Barzilai-Nahon, K. (2006). Gatekeepers, Virtual Communities and the Gated: Multidimensional Tensions in Cyberspace. *International Journal of Communications, Law and Policy*, 11, 1–28.
- Battelle, J. (2005). *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. New York: Portfolio.
- Bell, D. (1992). *Faces at the Bottom of the Well*. New York: Basic Books.