

Preface

Are you famous?

Celebrities are our contemporary era's royalty, the icons for how life ought (or ought not) to be lived. They and their scrupulously manicured personae are stand-ins for success, beauty, desire, and opulence. Yet celebrity life can appear, at least to the plebeian outside, as a golden prison, a structurally gorgeous, superficially rewarding, but essentially intolerable clink. The idea of the "right to privacy," coined back in the late nineteenth century, actually arose to address celebrity toils within this prison.¹ Newspaper photographers were climbing onto tree branches, peering into high-society homes, and taking pictures of big-money soirees so as to give noncelebrity readers a glimpse into this beau monde. Our contemporary conception of privacy was born in response to these intrusions.

For those of us on the outside looking in, our noncelebrity identity and relation to privacy may seem much simpler. Or is it? How can one really be sure one *is*—or *is not*—a celebrity? It's easy enough to do a quick test to prove it. Go to google.co.uk (note: not google.com), and search for your name. Then, scroll to the bottom of the results page. Do you see the phrase "Some results may have been removed under data protection law in Europe" (figure P.1)? The answer to this question—whether this caveat appears or not—will tell you how Google, or more precisely Google's algorithms, have categorized and thus defined you.

If you encounter this phrase, it is an algorithmic indication that Google doesn't consider you noteworthy enough to deserve the golden prison—you are not a celebrity, in the sense that the public has no right

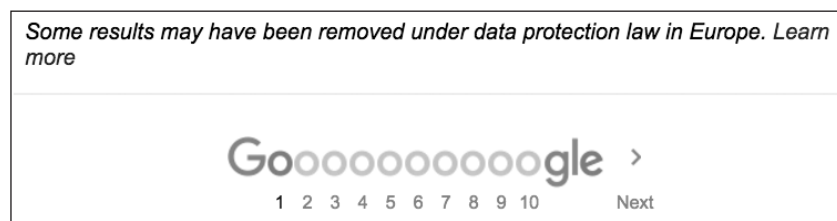


FIGURE P.1. When searching for your name on Google within the European Union, the phrase “Some results may have been removed under data protection law in Europe” appears if you are not a Google celebrity. Source: www.google.co.uk.

to know your personal details. You are just an ordinary human being. And this unadorned humanity grants you something quite valuable: you can request what European Union courts have called “the right to be forgotten.”

In 2014, Google was forced by EU courts to allow its European users the “right to be forgotten.” Since May of that year, individuals can now submit a take-down request of European search results that are “inadequate, irrelevant or no longer relevant, or excessive in relation to those purposes and in the light of the time that has elapsed.”² An avalanche of hundreds of thousands of requests followed—some from people with legitimate grievances, others from those attempting to purge references to various kinds of criminal convictions.³

In response, Google quietly introduced a new feature that automates part of this take-down process. In a letter to the European Union’s data-protection authorities, Google’s global privacy counsel, Peter Fleischer, explained that since “most name queries are for famous people and such searches are very rarely affected by a removal, due to the role played by these persons in public life, we have made a pragmatic choice not to show this notice by default for known celebrities or public figures.”⁴ In other words, whether the phrase appears or not reflects Google’s determination of whether you are a regular person who can

request the removal of certain search results or a celebrity who presumably just has to lump it when it comes to privacy.

This “pragmatic choice” is algorithmic—an automatic, data-based division between “celebrities or public figures” and the rest of us. Yet the EU legal opinion did not define who counts as a public figure. Google gets to make it up as it goes. The exact method by which it does this is unknown, hidden behind the computational curtain. *Washington Post* legal blogger Stewart Baker, who first brought this algorithmic assessment to public attention, found that Google believed pop singer Rihanna to be “Google famous” but Robyn Rihanna Fenty (the singer’s legal name) “not.”⁵ The distinction is intriguing because it shows a gap, a separation that exists primarily at the level of knowledge. In this case, Google isn’t really assessing “fame.” Google is creating its own, proprietary version of it.

Of course, fame, much like Google fame, has always been a creation. Film studies scholar Richard Dyer detailed this point many decades ago: who becomes a flesh-and-blood celebrity, or in his words, embodies a “star image,” is neither a random nor meritocratic occurrence. “A star image is made out of media texts that can be grouped together as promotion, publicity, films, and commentaries/criticism.”⁶ A star is not born. A star’s image is meticulously made and remade according to, in Dyer’s case, Hollywood.⁷

Suitably, the site of any star’s making is also a site of power. In the case of Hollywood and other mass-media industries, this power facilitates profit. But of equal importance is how these industries simultaneously produce what cultural theorist Graeme Turner calls the “raw material” of our identity.⁸ Our aspirations, desires, and expectations of life are influenced, in some way, by the star images we all encounter. To be in control of these images is to control how a social body understands itself. Or who is on the screen or onstage (and who isn’t) concretizes the epistemic possibilities of who we can imagine ourselves to be.⁹

But Google’s celebrity is functionally different. When Google determines your fame, it is not trying to package you as a “star image,” nor is

it producing “raw material” for cultural consumption. Rather, it is automatically fulfilling a legal mandate to facilitate European Union personal information take-down requests. And it is through this functional difference that Google has structurally transformed the very idea of fame itself. By developing an algorithmic metric that serves as a data-based stand-in for celebrity, Google created an entirely new index for who legally has the “right to be forgotten” (noncelebrities) and who doesn’t (celebrities).

In short, Google’s algorithmic index is an emergent way of thinking about what a celebrity is in our contemporary, data-rich communications environment. And in doing so, this emergent index fundamentally reworks our ideas around identity. In the present day of ubiquitous surveillance, who we are is not only what we think we are. Who we are is what our data is made to say about us.

Through Google’s extensive database network, celebrities are a different form of “star” and thus produce another kind of “raw material” for who we are seen to be online. And in this difference is where the line between celebrities and “the rest of us” begins to blur. A Google celebrity may not be an actual celebrity. Rather, a Google celebrity is someone whose data is algorithmically authenticated as such. And while Google users might not grace the covers of magazines, they do produce an unprecedented amount of information about themselves, their desires, and their patterns of life. This data is a different type of raw material according to a different kind of industry—what media scholar Joseph Turrow calls the “new advertising industry” and WikiLeaks described as the “global mass surveillance industry.”¹⁰

That is to say, “Google famous” may not equal famous, but “Google famous” influences which search results get censored and which lives are deemed available for public consumption. It orders the discourses of, and access to, personal privacy rights. And it inaugurates a new future for what it means to be a celebrity or public figure. But unlike the media-industry networks that painstakingly curate the raw materials of a star image, Google uses an algorithmic category of celebrity based entirely on interpretations of data.

Twentieth-century scientific positivism demands that we let data speak for itself. Following this mantra, wherever data tells us to go, we will find truth. But the data that Google uses to categorize people and assign status of identity does not speak; it is evaluated and ordered by a powerful corporation in order to avoid legal culpability. Indeed, scholars Lisa Gitelman and Virginia Jackson argue data doesn't speak but is spoken for.¹¹ Data does not naturally appear in the wild. Rather, it is collected by humans, manipulated by researchers, and ultimately massaged by theoreticians to explain a phenomenon. Who speaks for data, then, wields the extraordinary power to frame how we come to understand ourselves and our place in the world.

To participate in today's digitally networked world is to produce an impressive amount of data. From those who post personal details on Facebook to others who simply carry their cell phone with them to and from work, we leave traces of our lives in ways we never would expect. And as this data funnels into an expansive nexus of corporate and state databases, we are clearly not the ones who interpret what it means.

In the following introductory chapter, I begin to etch out what it means to be made "of data." Algorithmic interpretations about data of our web surfing, data of our faces, and even data about our friendships set new, distinct terms for identity online. And it is through these terms that our algorithmic identities are crafted—terms in which race is described by ones and zeros and emotions defined by templates of data.

We Are Data is about how algorithms assemble, and control, our datafied selves and our algorithmic futures. It's about how algorithms make our data speak as if we were a man, woman, Santa Claus, citizen, Asian, and/or wealthy. And it's also about how these algorithmically produced categories replace the politicized language of race, gender, and class with a proprietary vocabulary that speaks *for* us—to marketers, political campaigns, government dragnets, and others—whether we know about it, like it, or not. The knowledge that shapes both the world and ourselves online is increasingly being built by algorithms, data, and the logics therein.

