

On the Data Blasé, Statistical Imaginaries, and the Cashier-less Bodega

MARK SHEPARD

University at Buffalo

The banality of everyday data is a hallmark of the 21st century. At a moment when computational techniques involving big data and machine learning are in the process of parsing all walks of life, we might revisit notions of the gigantic, its transformation of the quantitative into the qualitative, and the corresponding subjectivities produced through this process. The cashier-less bodega provides a context to examine the implications of this new informatics regime.

I would sum up my fear about the future in one word: boring.

—J. G. Ballard¹

DATA GIGANTICISM

Despite the incessant hype of early 21st century techno-evangelists, the near-future—that point in time just beyond the present that holds the promise of radical change brought about by disruptive technologies—appears to have made manifest Ballard's fear. Between the breathless claims of proponents of “smarter cities” and the rousing critiques of “radical devices” lies the banal cloudscape of everyday data. This cloudscape is neither the highly optimized, ever-more efficient and sustainable city we've been promised, nor the spectacularly dark and sinister surveillance state of the post-Snowden era we've been warned about. This cloudscape is as broadly pervasive as it is largely invisible, although it tends to render as beige across our collective consciousness. It stretches seamlessly across public and private domains, between home and office, and throughout both online and offline environments. It blankets urban and exurban contexts equally without prejudice, although its density is perhaps most pronounced in cities. We have learned to reckon with it as we do with other matters of similarly prosaic significance: with practiced indifference.

Currently, we currently generate 2.5 quintillion bytes of data each day, with 90 percent of the total data in the world produced in the past two years. This volume is so large that our data activity is more commonly measured by the minute. In one minute we produce 456,000 tweets, send 16 million text messages, and 156 million emails. We perform more than 4 million Google searches, post more than 500,000 comments on Facebook, and stream almost 70,000 hours of video on Netflix. We make 154,200 calls on Skype, swipe 990,000 times on Tinder, and send 527,760 snap chat images. We receive more than 100 million spam emails. Every minute.²

Beyond the data we intentionally generate is data that results from our movements and transactions with various media, information and communications systems. A byproduct of the devices or systems with which we habitually engage, this data exhaust is invisibly emitted as we go about our daily business.³ Search histories, location histories, transaction histories: we leave trails of data, often unwittingly, as we surf the web, commute to work, or purchase groceries. This data exhaust in turn begets even more data. Personal profiles are derived from our online activity. Traffic congestion data for roads and highways is extrapolated from GPS-based navigation apps such as Google Maps. Retailers attempt to predict the probability that we are pregnant based on our purchasing behavior.⁴

This staggering amount of data is stored in more than 7,500 data centers worldwide that, according to the Natural Resources Defense Council (NRDC), consume up to 3% of global electricity supply.⁵ An average-sized data center draws over 100 times the power of a corporate office tower, whereas larger data centers can consume enough electricity to power a small town.⁶ The world's largest data center, the Range International Information Hub located in Langfang, China, occupies 6.3 million square feet. That's the size of the Pentagon, or the equivalent of 110 football fields.

“No type of building embodies 21st-century culture more distinctly than the data center,” observes Kazys Varnelis. “The physical reality of the cloud, they are the substance behind the portable, networked devices that we peer into as we stumble about our daily business.”⁷ Data centers and their network infrastructure are the enablers of our data-driven culture, much as train stations and railway networks enabled the new mobilities of the 19th century. Yet unlike the monuments to communications infrastructure like John Carl Warnecke's 1974 AT&T Long Lines Building centrally located in lower Manhattan, today's data center is designed not to be seen. Typically located away from dense metropolitan centers and close to fiberoptic infrastructure and cheap energy supply, contemporary data centers are the big box companions of the black boxes they facilitate. As our appetite for data increases exponentially, so too the material footprint of the immaterial cloud through which all this data circulates.

Each age or epoch has its own concept of the gigantic. The notion of the gigantic as a characteristic condition of the twentieth century is perhaps most vividly described by



Figure 1. NSA Utah Data Center, Bluffdale, Utah. Creative Commons CC0 1.0 Universal, Parker Higgins.

Heidegger as a process of transformation from the quantitative to qualitative, whereby the gigantic itself becomes a special quality, exceeding the limits of its own calculability:

The gigantic is ... that through which the quantitative becomes a special quality and thus a remarkable kind of greatness. Each historical age is not only great in a distinctive way in contrast to others; it also has, in each instance, its own concept of greatness. But as soon as the gigantic in planning and calculating and adjusting and making secure shifts over out of the quantitative and becomes a special quality, then what is gigantic, and what can seemingly always be calculated completely, becomes, precisely through this, incalculable. This becoming incalculable remains the invisible shadow that is cast around all things everywhere when man has been transformed into subiectum and the world into picture.⁸

With the data centers of the 21st century, we see a new cultural articulation of the gigantic. At a moment when big data and machine learning are being applied to all walks of life, we might revisit historical precedents of the gigantic, the transformation of the quantitative into the qualitative, and the corresponding subjectivities produced through this

process. The challenge is to focus on the cultural processes and procedures that produce the conditions that make possible the giganticism of these big black boxes—the incalculable, in Heidegger’s view—rather than unpack their deep structures and internal relations in essentialist terms.

THE DATA BLASÉ AND THE REPUTATION ECONOMY

Writing at the dawn of the 20th century, German sociologist Georg Simmel described the emergence of the modern urban subject at the tail end of the transformation from agrarian to industrial societies in continental Europe. To Simmel, the problems of modern urban life stemmed from the struggle of individuals to preserve the autonomy and individuality of their existence in face of what he characterized as overwhelming social forces. Central to his thesis is the emergence of the “blasé attitude” of “those who live in cities.”⁹ The blasé attitude was marked by a radical indifference brought about by the intensification of nervous stimulation resulting from the swift change of external and internal stimuli produced by the modern metropolis. The “rapid telescoping of changing images, pronounced differences within what is grasped at a single glance, and the unexpectedness of violent stimuli” reinforced this psychological condition “with every crossing of the street, with the tempo and multiplicity of economic,

occupational and social life.”¹⁰ This overstimulation of the senses and hyper rationalization of consciousness, Simmel believed, stemmed primarily from the domination of the money economy across all social relations.

Today, our attention is no longer divided simply within our field of vision, but between two radically different fields of vision, one human, one non-human. Whereas Simmel describes the relative independence afforded the modern urban subject in contrast to the more constrained social circles common to rural life, the contemporary urbanite is overwhelmingly dependent on connectivity to data and network services. When we retreat to the countryside, we do so not in search of “a relatively small circle almost entirely closed against neighboring foreign or otherwise antagonistic groups,”¹¹ but rather to unplug, disconnect and free ourselves from the various networked devices to which we have become invariably tethered. In place of the overstimulation and hyper rationalization of Simmel’s modern urban subject, we might posit our constant negotiation between proximate and remote places, between presence and co-presence, as complicit in the production of what we might call the “data blasé” attitude.

BLASÉ	DATA BLASÉ
Individual	Collective, aggregate
Independence	Dependence on data and network services
Money economy	Reputation economy
Attention divided within the frame	Attention divided between frames
Indifference to value distinctions between things	Indifference to value of physical, proximate interactions with others
Specialization	Normalization

Figure 2. Comparison of the blasé and the data blasé attitudes.

If Simmel tied the blasé attitude to the emergence of the money economy, the data blasé can be seen in part as a byproduct of the reputation economy, where the data generated by our activity on social media platforms determines our social value, access to services, and employment opportunities. From Uber drivers to AirBnB hosts, reputation becomes a form of currency in an economy of star ratings, and social status is measured in quantities of likes, followers, shares and retweets. In *Nosedive*, an episode of the science fiction television series *Black Mirror*, we witness the trials and tribulations of Lacie who, attempting to climb the social ladder to

a better (and cheaper) apartment, takes great pains to ensure each of her personal or business interactions results in five-star ratings. The reputation economy, however, is not merely a figment of science fiction writers’ imaginations. China is currently developing its own “Citizen Score” platform, where one’s credit score is combined with a sentiment analysis of their social media activity and that of their friends, their police record and other personal data points to create a single score by which one becomes eligible for discounts on car and apartment rentals, for instance. The system is projected to come online in the year 2020.

If the money economy reinforced the individuation and specialization of Simmel’s modern urban subject, the reputation economy favors the collective and the aggregate, where the normalization of behavioral types and ever-finer demographic categories dominate. With the integration of artificial intelligence (AI) techniques involving data mining, deep learning and sentiment analysis into these reputation systems, social pressure intensifies to conform to identifiable patterns of behavior. We become afraid to speak out or stand out for fear it might adversely impact our ability to get a job or pay less for life’s essentials. Known as “social cooling,”¹² this culture of self-censorship leads people to modify their behavior to conform to social norms legible to the algorithm in order to maintain their digital reputations.

Whereas the blasé attitude is marked by an indifference to value distinctions between things, the data blasé can be characterized in part by an indifference to the value of physical, proximate interactions with others. According to a recent study by Common Sense Media, 61% of teenagers preferred texting, video chatting or social media over talking to their friends in person.¹³ Hanging out at home while remaining in contact with close circles of friends via text messaging and social media has, for some teens, become as satisfying as gathering together in a physical location.¹⁴ This indifference to the value of physical spaces for social encounters is by no means new. Many have narrated the collapse of public space as the geography of the public sphere.¹⁵ What is perhaps notable here is the reformatting of social interaction by the affordances and constraints of social media and other communication platforms.

Yet despite the increasing frequency of security breaches and compromised personal data held by consumer credit reporting agencies, or the weaponization of social media data for partisan political warfare, we have become indifferent to the multiple ways by which we exchange our personal data for access to online services, discount pricing, and other benefits of our data-driven world. Expectations for the privacy of the personal information we share online and off are elided by what we expect to gain in return. Even post-Snowden fears of intrusive surveillance by the NSA and other government agencies are assuaged by the belief that if we’ve done nothing



Figure 3. Amazon Go store, Seattle, Washington. Creative Commons Attribution-Share Alike 4.0, Sikander Iqbal.

wrong we need not worry too much. This indifference contributes to the perceived imperceptibility of data and its implications for everyday life. That our data-driven culture appears too big to fail is society's new inconvenient truth.

THE CONVENIENCE OF GRAB AND GO RETAIL AND ITS STATISTICAL IMAGINARIES

Convenience is the name of the awards ceremony at which capitalism admires itself.

—Near Future Laboratory¹⁶

If the data center is the architectural emblem of early 21st century culture writ large, the mini market portends to be its scalar counterpart. On Monday, January 22, 2018, Amazon opened to the public its first brick-and-mortar convenience store in downtown Seattle. Occupying a mere 1,800 square feet at the base of Amazon's corporate office tower, the store, named Amazon Go, represents the online retail giant's first foray into the grocery business after purchasing super market

chain Whole Foods for \$13.7 billion just six months prior. Drawing on its expertise in reducing friction in online retail transactions, Amazon aims with Go to deliver a frictionless shopping experience to the corner bodega.

While the store in almost all respects resembles an upscale urban mini market, a few key features are noticeably absent: there are no checkout lines or cashiers in sight. Shoppers login to Amazon's smartphone app, and upon entering the store, scan a QR code displayed on their phone on a turnstile similar to those found at TSA checkpoints in airports or entrances to subway stations. As they go about their shopping, a computer vision system incorporating hundreds of cameras embedded in the ceiling employs machine learning algorithms to track which items they take off the shelves, and updates their virtual shopping carts accordingly. Upon leaving the store, the system automatically bills them for the items they took. Amazon calls this "just walk out" technology, and it enables shoppers to do just that. No waiting in lines for baskets of items to be scanned by a cashier using a barcode reader. No swiping of credit cards or exchange of cash. Grab and go has

never been easier. Just walk out. Amazon Go presents a frictionless shopping experience, where the customer—under continuous, total surveillance—is led to feel like a shoplifter.¹⁷

Grab and go shopping has been the future of retail for some time now. Since the development of contactless payment systems and self-checkout aisles, the notion of frictionless consumption has captivated the imagination of the retail industry. Minimize the ebbs in the shopping experience to maximize the flows of capital. Dilip Kumar, vice president of technology for Amazon Go, identifies “time poverty” as the top affliction of the modern urbanite when discussing Amazon’s strategy to meet the needs of their customer.¹⁸ In a culture where time is money, offering a shopping experience that saves time for customers also saves them money.

Delivering on the grab and go promise, however, was no simple task. While Amazon has said little about how the technology works, the key challenges of such a system involve: 1) identifying items that a customer places in their shopping cart, and 2) identifying the customer that placed the items in the cart. At its core, Amazon Go is a product of the same fundamental advances in artificial intelligence, computer vision, and automated decision-making that are behind recent advances in driverless cars. Specifically, the advent of deep learning techniques, cloud computing and probabilistic reasoning have made possible what even five years ago was not.

Computer vision has evolved rapidly over the past decade. Central to this evolution is the Convolutional Neural Network (CNN). With a CNN, as with the visual cortex, neurons sensitive to particular features are distributed across the field of vision. The output of these neurons are connected together in such a way so that a single neuron sensitive to, say, a box of pasta, scans the entire visual field. The resultant output indicating where the box of pasta is in the image is then fed to the next layer. This process is repeated for the entire feature set the system has been trained to recognize, mapping the spatial locations of these features within the images. These spatial maps are then fed into higher levels of the network where they are combined to recognize higher-order patterns and objects.

Until recently, the computational power needed to process the multiple inputs and outputs of complex neural networks in relatively short periods of time was prohibitive for commercial applications such as these. The ability to employ graphics processing units (GPUs) as desktop supercomputers, however, made the parallel processing of complex neural networks possible. Computations requiring hours of processing by the most powerful CPUs took mere minutes on the average GPU. With Amazon being one of the largest providers of cloud-based GPU processing in the world (Amazon AWS), access to the required computational power would not be an obstacle for Amazon Go.

Integrating observations from a large array of cameras at different points in time into a coherent picture of the state of the customer’s virtual shopping cart required pursuing a probabilistic scheme implementing Bayesian statistical methods. Bayesian statistics assumes that the world is inherently an uncertain place, and that to solve a problem, we need to embrace this uncertainty. Generally speaking, Bayes’ theorem describes the probability of an event, based on prior knowledge of conditions that might be related to the event. It compares observations of an event and the underlying probability of that event occurring to calculate the probability that the given event has occurred. Our belief that a possible state is true is a function of how strongly we believed in it before, multiplied by how much a sensor reading supports that possible state.

There are two clear advantages Amazon has in employing this kind of probabilistic reasoning. The first is that Bayes’ theorem takes into account prior probabilities, and Amazon knows the prior purchasing histories of a vast number of customers. For instance, if an Amazon Go customer buys a salmon sandwich and a bag of chips every other day around lunch time, the system can weigh its initial values to return a higher probability for those items on those days. This is one way Amazon can leverage the vast amount of data it already has on its customers to improve the accuracy of the system. The other advantage is that by translating everything into the language of probability, the system can combine readings from multiple sensors over different time periods. If we assume these observations are independent, then we can simply multiply probabilities together, using the posterior of one observation as the prior for the next.¹⁹

In illustrating the shift from disciplinary to control societies, Gilles Deleuze defines the “dividual” as the control society descendent of the disciplinary “individual.”²⁰ Whereas an individual was identified through the uniqueness of their signature in disciplinary societies, control society dividuals are identified through combinations of usernames and passwords. Deleuze’s dividual is a market segment, a demographic tranch, a body without organs distributed across a probabilistic landscape composed of data points registered in discrete time series. We are (probably) what we eat – when, where and with whom we’ve most likely eaten it.

If the architectural emblem of the 21st century—the data center—is a warehouse processing vast amounts of data we generate every minute, cashier-less bodegas like Amazon Go become the gristmills that encode our idiosyncratic shopping behaviors and purchasing habits into legible, smooth and marketable data bodies. These statistical imaginaries form the training sets for the urban mini market coming to a corner near you. At the time of this writing, Amazon has opened three additional Amazon Go stores, most recently one in Chicago. While we trade data about our shopping behavior

and purchasing habits for a quicker checkout experience, the higher-order value proposition is telegraphed by Amazon's plan to roll out 3,000 of its grab and go bodegas by the year 2021.²¹ The colonization of the corner bodega by data-driven systems is poised to dramatically expand the reach of data harvesting performed by one of the world's largest retailers. Ultimately these high-tech mini markets aim to become an interface for extracting data from even the most prosaic of our daily activities: shopping for groceries. Indeed, the future of retail would appear rather boring.

ENDNOTES

- 1 Interview with J.G. Ballard, October 30, 1982, *Re/Search* 8/9 (1984).
- 2 Bernard Marr, "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read," *Forbes*, May 21, 2018, <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#7f08740360ba>.
- 3 See James Manyika et. al., *Big Data: The Next Frontier for Innovation, Competition, and Productivity* (Washington, D.C.: McKinsey Global Institute, 2011); and Rob Kitchin, *The Data Revolution* (Thousand Oaks, CA: Sage, 2014).
- 4 Charles Duhigg, "How Companies Know Your Secrets," *New York Times*, February 16, 2012, <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>.
- 5 "12 Mind-Blowing Data Center Facts You Need to Know," Ciena official website, August 25, 2016. <https://www.ciena.com/insights/articles/Twelve-Mind-blowing-Data-Center-Facts-You-Need-to-Know.html>.
- 6 "12 Mind-Blowing Data Center Facts You Need to Know," Ciena official website.
- 7 Kazys Varnelis, "Eyes That Do Not See: Tracking the Self in the Age of the Data Center," *Harvard Design Magazine* 38 (2014). <http://www.harvarddesignmagazine.org/issues/38/eyes-that-do-not-see-tracking-the-self-in-the-age-of-the-data-center>.
- 8 Martin Heidegger, "The Age of the World Picture," *The Question Concerning Technology and Other Essays* (New York: Harper Torchbooks, 1977), 135.
- 9 For a characterization of this type, see Berthold Brecht, *Reader for Those Who Live in Cities*.
- 10 Georg Simmel, "The Metropolis and Mental Life," *Simmel on Culture* (Thousand Oaks: Sage, 1997).
- 11 Simmel, 189.
- 12 Sophie Kleber, "As AI Meets the Reputation Economy, We're All Being Silently Judged," *Harvard Business Review*, January 29, 2018. <https://hbr.org/2018/01/as-ai-meets-the-reputation-economy-were-all-being-silently-judged>.
- 13 "Social Media, Social Life Infographic" Common Sense Media official website, September 10, 2018. <https://www.commonsensemedia.org/social-media-social-life-infographic>.
- 14 Amanda Lenhart, "How Teens Hang Out and Stay in Touch with Their Closest Friends," *Teens Technology and Friendships*, Pew Research Center, August 6, 2015. <http://www.pewinternet.org/2015/08/06/chapter-2-how-teens-hang-out-and-stay-in-touch-with-their-closest-friends/>.
- 15 For a cross-section of views on the geography of the public sphere, see Jürgen Habermas, *The Structural Transformation of the Public Sphere* (Cambridge, MA: The MIT Press, 1962); Richard Sennett, *The Fall of Public Man* (New York: W.W. Norton & Company, 1977); Low and Smith, *The Politics of Public Space* (Abingdon, UK: Routledge, 2006); and Yochai Benkler, *The Wealth of Networks* (New Haven, CT: Yale University Press, 2007).
- 16 Julian Bleecker et. al., "Convenience Newspaper," *The Near Future Laboratory*, March 1, 2012. <http://nearfuturelaboratory.com/2012/03/01/corner-convenience/>.
- 17 For descriptions by reporters of the "shoplifting effect" of the Amazon Go experience, see for example Nick Wingfield, "Inside Amazon Go, a Store of the Future," *The New York Times*, January 21, 2018. <https://www.nytimes.com/2018/01/21/technology/inside-amazon-go-a-store-of-the-future.html>; and Jake Bullinger, "Amazon's Checkout-Free Store Makes Shopping Feel Like Shoplifting," *The Atlantic*, January 24, 2018. <https://www.theatlantic.com/business/archive/2018/01/amazon-go-store-checkouts-seattle/551357/>.
- 18 Drew Harwell and Abha Bhattarai, "Inside Amazon Go: The Camera-Filled Convenience Store that Watches You Back," *The Washington Post*, January 22, 2018. https://www.washingtonpost.com/news/business/wp/2018/01/22/inside-amazon-go-the-camera-filled-convenience-store-that-watches-you-back/?utm_term=.197c00effd9d.
- 19 Haomiao Huang, "How Amazon Go (Probably) Makes 'Just Walk Out' Groceries a Reality," *Ars Technica*, April 10, 2017. <https://arstechnica.com/information-technology/2017/04/how-amazon-go-probably-makes-just-walk-out-groceries-a-reality/>.
- 20 Gilles Deleuze, "Postscript on Control Societies," in *Negotiations: 1972–1990*, trans. Martin Joughin (New York: Columbia University Press, 1995), 177.
- 21 Spencer Soper, "Amazon Will Consider Opening Up to 3,000 Cashierless Stores by 2021," *Bloomberg*, September 19, 2018. <https://www.bloomberg.com/news/articles/2018-09-19/amazon-is-said-to-plan-up-to-3-000-cashierless-stores-by-2021>.