

INTRODUCTION

IMAGINE YOU'RE IN a café, perhaps the noisy one I'm sitting in at this moment. A young woman at a table to your right is typing on her laptop. You turn your head and look at her screen. She surfs the Internet. You watch.

Hours pass. She reads an online newspaper. You notice that she reads three articles about China. She scouts movies for Friday night and watches the trailer for *Kung Fu Panda*. She clicks on an ad that promises to connect her to old high school classmates. You sit there taking notes. With each passing minute, you're learning more about her. Now imagine that you could watch 150 million people surfing at the same time. That's more or less what Dave Morgan does.

"What is it about romantic-movie lovers?" Morgan asks, as we sit in his New York office on a darkening summer afternoon. The advertising entrepreneur is flush with details about our ramblings online. He can trace the patterns of our migrations, as if we were swallows or humpback whales, while we move from site to site. Recently he's become intrigued by the people who click most often on an ad for car rentals. Among them, the largest group had paid a visit to online obituary listings. That makes sense, he says, over the patter of rain against

the windows. "Someone dies, so you fly to the funeral and rent a car." But it's the second-largest group that has Morgan scratching his head. Romantic-movie lovers. For some reason Morgan can't fathom, loads of them seem drawn to a banner ad for Alamo Rent A Car.

Morgan, a cheery 43-year-old, wears his hair pushed to the side, as if when he was young his mother dipped a comb into water, drew it across, and the hair just stayed there. He grew up in Clearfield, a small town in western Pennsylvania a short drive from Punxsutawney. Every year on the second day of February, halfway between the winter solstice and the vernal equinox, a crowd in that town gathers around a large caged rodent still groggy from hibernation. They study the animal's response to its own shadow. According to ancient Celtic lore, that single bit of data tells them whether spring will come quickly or hold off until late March. Morgan has migrated as far as can be from such folk predictions. At his New York start-up, Tacoda, he hires statisticians to track our wanderings on the Web and figure out our next moves. Morgan was a pioneer in Internet advertising during the dot-com boom, starting up an agency called 24/7 Real Media. During the bust that followed he founded another company, Tacoda, and moved seamlessly into what he saw as the next big thing: helping advertisers pinpoint the most promising Web surfers for their message.

Tacoda's entire business gorges on data. The company has struck deals with thousands of online publications, from the *New York Times* to *BusinessWeek*. Their sites allow Tacoda to drop a bit of computer code called a cookie into our computers. This lets Tacoda trace our path from one site to the next. The company focuses on our behavior and doesn't bother finding out our names or other personal details. (That might provoke a backlash concerning privacy.) But Tacoda can still

learn plenty. Let's say you visit the *Boston Globe* and read a column on the Toyota Prius. Then you look at the car section on AOL. Good chance you're in the market for wheels. So Tacoda hits you at some point in your Web wanderings with a car ad. Click on it, and Tacoda gets paid by the advertiser—and gleans one more detail about you in the process. The company harvests 20 billion of these behavioral clues every day.

Sometimes Morgan's team spots groups of Web surfers who appear to move in sync. The challenge then is to figure out what triggers their movements. Once this is clear, the advertisers can anticipate people's online journeys—and sprinkle their paths with just the right ads. This requires research. Take the curious connection between fans of romance movies and the Alamo Rent A Car ad. To come to grips with it, Morgan and his colleagues have to dig deeper into the data. Do car renters arrive in larger numbers from a certain type of romance movie, maybe ones that take place in an exotic locale? Do members of this group have other favorite sites in common? The answers lie in the strings of ones and zeros that our computers send forth. Maybe the statistics will show that the apparent link between movie fans and car renters was just a statistical quirk. Or perhaps Morgan's team will unearth a broader trend, a correlation between romance and travel, lust and wanderlust. That could lead to all kinds of advertising insights. In either case, Morgan can order up hundreds of tests. With each one he can glean a little bit more about us and target the ads with ever more precision. He's taking analysis that once ran through an advertiser's gut, and replacing it with science. We're his guinea pigs—or groundhogs—and we never stop working for him.

. . .

WHEN IT COMES to producing data, we're prolific. Those of us wielding cell phones, laptops, and credit cards fatten our digital dossiers every day, simply by living. Take me. As I write on this spring morning, Verizon, my cell phone company, can pin me down within several yards of this café in New Jersey. Visa can testify that I'm well caffeinated, probably to overcome the effects of the Portuguese wine I bought last night at 8:19. This was just in time for watching a college basketball game, which, as TiVo might know, I turned off after the first half. Security cameras capture time-stamped images of me near every bank and convenience store. And don't get me started on my Web wanderings. Those are already a matter of record for dozens of Internet publishers and advertisers around the world. Dave Morgan is just one in a large and curious crowd. Late in the past century, to come up with this level of reporting, the East German government had to enlist tens of thousands of its citizens as spies. Today we spy on ourselves and send electronic updates minute by minute.

This all started with computer chips. Until the 1980s, these bits of silicon, bristling with millions of microscopic transistors, were still a novelty. But they've grown cheaper and more powerful year by year, and now manufacturers throw them into virtually anything that can benefit from a dab of smarts. They power our cell phones, the controls in our cars, our digital cameras, and, of course, our computers. Every holiday season, the packages we open bring more chips into our lives. These chips can record every instruction they receive and every job they do. They're fastidious note takers. They record the minutiae of our lives. Taken alone, each bit of information is nearly meaningless. But put the bits together, and the patterns describe our tastes and symptoms, our routines at work, the paths we tread through the mall and the supermarket. And

these streams of data circle the globe. Send a friend a smiley face from your cell phone. That bit of your behavior, that tiny gesture, is instantly rushing, with billions of others, through fiber-optic cables. It's soaring up to a satellite and back down again and checking in at a server farm in Singapore before you've put the phone back in your pocket. With so many bits flying around, the very air we breathe is teeming with motes of information.

If someone could gather and organize these far-flung electronic gestures, our lives would pop into focus. This would create an ever-changing, up-to-the-minute mosaic of human behavior. The prospect is enough to make marketers quiver with excitement. Once they have a bead on our data, they can decode our desires, our fears, and our needs. Then they can sell us precisely what we're hankering for.

But it sounds a lot simpler than it is. Sloshing oceans of data, from e-mails and porn downloads to sales receipts, create immense chaotic waves. In a single month, Yahoo alone gathers 110 billion pieces of data about its customers, according to a 2008 study by the research firm comScore. Each person visiting sites in Yahoo's network of advertisers leaves behind, on average, a trail of 2,520 clues. Piece together these details, you might think, and our portraits as shoppers, travelers, and workers would gell in an instant. Summoning such clarity, however, is a slog. When I visit Yahoo's head of research, Prabhakar Raghavan, he tells me that most of the data trove is digital garbage. He calls it "noise" and says that it can easily overwhelm Yahoo's computers. If one of Raghavan's scientists gives an imprecise computer command while trawling through Yahoo's data, he can send the company's servers whirling madly through the noise for days on end. But a timely tweak in these instructions can speed up the hunt by a factor

of 30,000. That reduces a 24-hour process to about three seconds. His point is that people with the right smarts can summon meaning from the nearly bottomless sea of data. It's not easy, but they can find us there.

The only folks who can make sense of the data we create are crack mathematicians, computer scientists, and engineers. They know how to turn the bits of our lives into symbols. Why is this necessary? Imagine that you wanted to keep track of everything you ate for a year. If you're like I was in the fourth grade, you go to the stationery store and buy a fat stack of index cards. Then, at every meal you write the different foods on fresh cards. Meat loaf. Spinach. Tapioca pudding. Cheerios. After a few days, you have a growing pile of cards. The problem is, there's no way to count or analyze them. They're just a bunch of words. These are symbols too, of course, each one representing a thing or a concept. But they are near impossible to add or subtract, or to drop into a graph illustrating a trend. Put these words in a pile, and they add up to what the specialists call "unstructured data." That's computer talk for "a big mess." A better approach would be to label all the meats with *M*, all the green vegetables with *G*, all the candies with *C*, and so on. Once the words are reduced to symbols, you can put them on a spreadsheet and calculate, say, how many times you ate meat or candy in a given week. Then you can make a graph linking your diet to changes in your weight or the pimple count on your face.

The key to this process is to find similarities and patterns. We humans do this instinctively. It's how we figured out, long ago, which plants to eat and how to talk. But while many of us were focusing on specific challenges, others were thinking more symbolically. I picture early humans sitting around a fire. Some, naturally, are jousting for the biggest piece of

meat or busy with mating rituals. But off to the side, a select few are toying with stones, thinking, "If each of these pebbles represents one mammoth, then this rock . . ." Later, notes Tobias Dantzig in *Number: The Language of Science*, the Romans used their word *calcula*, meaning "pebble," to give a name to this thought process. But the pebble was just the start. The essence of calculation was to advance from the physical pebbles to ever-higher realms of abstract reasoning.

That science developed over the centuries, and we now have experts who are comfortable working with ridiculously large numbers, the billions and trillions that the rest of us find either unimaginable or irrelevant. They are heirs to the science that turns our everyday realities into symbols. As the data we produce continues to explode and computers grow relentlessly stronger, these maestros gain in power. Two of them made a big splash in the late 1990s by founding Google. For the age we're entering, Google is the marquee company. It's built almost entirely upon math, and its very purpose is to help us hunt down data. Google's breakthrough, which transformed a simple search engine into a media giant, was the discovery that our queries—the words we type when we hunt for Web pages—are of immense value to advertisers. The company figured out how to turn our data into money. And lots of others are looking to do the same thing. Data whizzes are pouring into biology, medicine, advertising, sports, politics. They are adding us up. We are being quantified.

When this process began, a half-century ago, the first computers were primitive boxes the size of a garbage truck. They kept their distance from us, purring away in air-conditioned rooms. At this early stage, the complexity of the human animal was too much for them. They couldn't even beat us at chess. But in certain numerical domains, they showed

promise. An early test involved consumer credit. In 1956, two Stanford graduates, a mathematician named Bill Fair and his engineer friend Earl Isaac, came up with the idea of replacing loan officers with a computer. This hulking machine knew practically nothing, not even what the applicants did for a living. It certainly hadn't learned if they'd gotten a raise or filed for divorce. Legions of human loan officers, by contrast, were swimming in data. They often knew the families of the loan applicants. They were acquainted with how much the applicant had struggled in high school and how his engagement had fallen through, probably because of a drinking problem (if he was anything like his uncle). The loan officers had enough details to write sociological monographs, if they were so inclined, about the families in their towns. But they lacked a scientific system to analyze it all. Bankers depended, for the most part, on their gut.

By contrast, the computerized approach zeroed in on only a small set of numbers, most of them concerning bank balances, debts, and payment history. Bare bones. Fair and Isaac built a company to analyze the patterns of those numbers. They developed a way to determine the odds that each customer would default on a loan. Everyone got a number. These risk scores proved to be much better predictors than the gut-trusting humans. Most borrowers with high credit scores made good on their loans. And more people qualified for them. The machine, after all, didn't discriminate on the basis of anything but numbers. It was equal-opportunity banking. Like a lot of analytical systems, it was fairer. Its narrow scope, paradoxically, returned broad-minded results. What's more, a lot of people turned out to be better bets than the loan officers suspected. The market for credit expanded.

Still, the computer knew its place. It thrived in the world

of numbers, and it stayed there. Those of us who specialized in words and music and images barely noticed it. Yet over the following decades, the computer grew in power, gobbling up ever more ones and zeros per millisecond. It got cheaper and smaller, and it linked up with others around the world. It produced jaw-dropping efficiencies. And from the viewpoint of the humanities crowd (including this history major), it swallowed entire technologies. It supplanted typewriters and moved on, like an imperial force, to rout record players and film cameras. It took over the mighty telephone. Finally, in the 1990s, even those of us who had long viewed computers as aliens from the basement world of geekdom started to make room for them in our homes and offices. We learned that we could use these machines to share our words and movies and photos with the entire world.

In fact, we had little choice. The old ways were laughably slow. But there was one condition: we had to render everything we sent, the very stuff of our lives, into ones and zeros. That's how we came to deliver our riches, the key to communications on earth, to the masters of the symbolic language. Now these mathematicians and computer scientists are in a position to rule the information of our lives. I call them the Numerati.

ON A SWELTERING summer afternoon, Dave Morgan sits in his Spartan office overlooking Seventh Avenue. He has the shade drawn to keep out the heat, and he can't figure out how to turn on the fluorescent light. Sitting in the shadows, he tells me how marketing has changed over the past generation. Traditionally, he says, marketers concentrated on big groups of us. We weren't much more diverse, from their point of view,

than the lines of General Motors cars: Cadillacs and Buicks for the rich and wannabes, Chevys for the middle class, Pontiacs for young hotshots, and pickup trucks for farmers. They didn't need to know much more than that because midcentury American factories, whether they were producing blue jeans or peanut butter, were manufacturing mass quantities. Smaller, more focused runs cost too much money. Sure, certain neighborhoods in coastal cities attracted eccentrics who drove foreign cars and walked around in lederhosen or berets. But for the most part, we ate, wore, and drove what the mass-production factories churned out, and we learned about it through the mass media. This model, created in the United States, spread in the decades following World War II across Europe and to much of Asia and Latin America. It was an efficient way to reach millions of consumers with machine-made goods.

Advertising in this industrial complex, Morgan says, was simple. You cut the population into five or six demographic groups, based on income, gender, and neighborhood, and you advertised in the magazines they read and on the TV shows they watched. In an age of virtually indistinguishable products, brands were crucial. This has all changed. "In 50 years," Morgan says, through the darkness, "we've gone from a command-and-control economy to one driven by consumers." How did this happen? For starters, computers made their way into factories. This gave manufacturers new flexibility. It became much easier to tweak cereals or sodas to create nuttier or more lemony blends. With a simple command, looms weaving a striped pattern switched to plaid. This wasn't much harder than it is for me to change the font from Times to Papyrus as I write this chapter. And it meant that industry could produce thousands of new variations. At the same time, globalization

was dumping products from all over the world practically on our doorstep. Today, choices are nearly limitless. Winning in this crowded marketplace requires far more than industrial efficiency. The trick now is to deliver to each of us the precise flavor and texture and color we want, at just the right price. Consumers run the show, Morgan says. "It's not controlled by manufacturing or distribution."

This means that marketers must scope us out as individuals. One approach would be to deploy battalions of psychology and lit majors armed with clipboards to knock on our doors. That's impractical. The sensible way to study us is to track and analyze the data we never stop spewing. And Morgan is stretching beyond that. He tells me of experiments his team is developing to monitor the spark of recognition in the brain as people look at online ads. The tests focus on a brain wave called p300. (The U.S. Navy has run similar tests to see how pilots distinguish friends from foes in the air.) If a p300 wave heats up within a fraction of a second of a subject's seeing an ad, the Tacoda team will make the case that the viewer has not only looked at the spot but has processed it mentally. The next step? Figuring out which type of people process certain types of ads. Like other Numerati in a wide range of industries, Dave Morgan is scrutinizing humans and searching for hidden correlations. What do we do, he asks, that might predict what we'll do next?

WHEN I TELL people about this book, they often say, "We're just going to be numbers!"

Yes, I say, but we've long been numbers. Think of the endless rows of workers threading together electronic cables in a Mexican assembly plant or the thousands of soldiers rushing

into machine-gun fire at Verdun—even the blissed-out crowd pushing through the turnstiles at a Grateful Dead concert. From management's point of view, all of us in these scenarios might as well be nameless and faceless. We're utterly interchangeable. Turning us into simple numbers was what happened in the industrial age. That was yesterday's story.

The Numerati have much more ambitious plans for us. Forget single digits. They want to calculate for each of us a huge and complex maze of numbers and equations. These are mathematical models. Scientists have been using them for decades to simulate everything from fleets of trucks to nuclear bombs. They build them from vast collections of data, with every piece representing a fact or a probability. Each model must reflect, in numbers, the physical truth: its size and weight, the characteristics of its metal and plastics, how it responds to changes in air pressure or heat. Complex models can have thousands, or even millions, of variables. And they must interact with one another mathematically just the way they do in the real world. Building them is painstaking work. And sometimes they flop. The dramatic market convulsions of 2008, for example, stemmed from faulty models that glossed over the complexity—and the risk—associated with real estate loans.

Despite such stumbles, today's Numerati are plowing forward, with an eye on us. They're already stitching bits of our data into predictive models, and they're just getting warmed up. In the coming decade, each of us will spawn, often unwittingly, models of ourselves in nearly every walk of life. We'll be modeled as workers, patients, soldiers, lovers, shoppers, and voters. In these early days, many of the models are still primitive, making us look like stick figures. The ultimate goal, though, is to build versions of humans that are just as complex

as we are—each one unique. Add all of these efforts together, and we're witnessing (as well as experiencing) the mathematical modeling of humanity. It promises to be one of the great undertakings of the twenty-first century. It will grow in scope to include much of the physical world as mathematicians get their hands on new flows of data, from constellations of atmospheric sensors to the feeds from millions of security cameras. It's a parallel world that's taking shape, a laboratory for innovation and discovery composed of numbers, vectors, and algorithms. And you and I are in the middle of it.

What will the Numerati learn about us as they turn us into dizzying combinations of numbers? First they need to find us. Say you're a potential SUV shopper in the northern suburbs of New York, or a churchgoing, antiabortion Democrat in Albuquerque. Maybe you're a Java programmer ready to relocate to Hyderabad, or a jazz-loving, Chianti-sipping Sagittarius looking for snuggles by the fireplace in Stockholm. Heaven help us: maybe you're eager to strap bombs to your waist and climb onto a bus. Whatever you are—and each of us is a lot of things—companies and governments want to identify and locate you. Consider this: Google grew into a multibillion-dollar sensation by helping us find the right Web page. How much more valuable will it be, in every conceivable industry, to find the right person? That information is worth fortunes, and the personal data we throw off draws countless paths straight to our door. Even if you hold back your name, it's a cinch to find you. A Carnegie Mellon University study recently showed that simply by disclosing gender, birth date, and postal zip code, 87 percent of people in the United States could be pinpointed by name.

The Numerati also want to alter our behavior. If we're shopping, they want us to buy more. At the workplace, they're

out to boost our productivity. As patients, they want us healthier and cheaper. As companies such as IBM and Amazon roll out early models of us, they can predict our behavior and experiment with us. They can simulate changes in a store or an office and see how we would likely react. And they can attempt to calculate mathematically how to boost our performance. How would shoppers like you respond to a \$100 rebate on top-of-the-line Nikon cameras? How much more productive would you be at the office if you had a \$600 course on spreadsheets? How would your colleagues cope if the company eliminated their positions or folded them into operations in Bangalore? The Numerati will be placing our models in all kinds of scenarios. They'll try out different medicines or advertisements on us. They'll see how we might respond to a new exercise regimen or a job transfer to a distant division. We don't have to participate or even know that our mathematical ghosts are laboring night and day as lab rats. We'll receive the results of these studies—the optimum course—as helpful suggestions, prescriptions, or marching orders.

The exploding world of data, as we'll see, is a giant laboratory of human behavior. It's a test bed for the social sciences, for economic behavior and psychology. Researchers at companies such as Microsoft and Yahoo are busy hiring scientists from fields as diverse as medicine and linguistics to help them grapple with the bits of our lives that are pouring in. These streams of digital data don't recognize ancient boundaries. They're defined by algorithms, not disciplines. They can easily cross-fertilize. This means that psychologists, economists, biologists, and computer scientists can collaborate as never before, all of them sifting for answers through countless details of our lives. Jack Einhorn, the chief scientist at a New York media start-up called Inform Technologies, predicts that the

great discoveries of the twenty-first century will come from finding patterns in vast archives of data. "The next Jonas Salk will be a mathematician," he says, "not a doctor."

IT'S MIDSUMMER gridlock in Manhattan. By the time I reach the French bistro in Chelsea, Dave Morgan's already sitting at a table by an open window, reading e-mails on his Treo. He seems distracted as we eat, glancing from time to time at the handset. Just as the waitress drops the dessert menus on our table, his machine beeps. Morgan looks at it, apologizes, and hurries off into the summer heat. From my seat at the window, I watch him angling across the street and trotting up the far sidewalk.

The next time I see Morgan, it's October. He's moved from Tacoda's Seventh Avenue offices and is newly installed at the headquarters of AOL, high above the skating rink at Rockefeller Center. I meet him at the door of what he calls 75 Rock, and we walk to a café. He tells me that on the day we had lunch, he and his investors agreed to sell Tacoda to AOL. (The reported price was \$275 million. The Numerati, it should be noted, tend to make a lot of money.) Morgan is working, at least for the time being, as a senior advertising exec at AOL. He certainly doesn't need the salary. But he says he's tempted to stick around. By tapping AOL's resources and its millions of users, he says, he can learn even more about Web surfers and target us with ever greater precision. It's a long process, he says. "We're just at the beginning."

I ask him about the correlation he told me about earlier, the one between romantic-movie fans and Alamo Rent A Car. It takes a moment for him to recall it. "Oh yeah. They were off the charts." Did his researchers, I ask, ever come up with

an explanation for it? He nods. "It had to do with weekends." It was Alamo ads promoting "escapes" that attracted the attention of these Web surfers, he says. The romantic-movie fans booked leisure rentals, largely for weekend getaways. Perhaps they wanted to act out the kinds of scenes that drew them to the cinema. Banners for weekday rentals apparently left them cold.

This brings Morgan to a different insight, one that involves not just who we are but how we feel. No doubt plenty of romantic-movie fans, he says, rent cars for business trips. But after reading the review of the latest candlelight-and-kisses movie, they're thinking about getaways to Napa Valley or Nantucket. Work, at least for the moment, is far away. The challenge ahead is to map not just our tastes and preferences but our shifting moods. "If you think about it," he says, "the movies and music that people click on tell us a lot about their state of mind at that moment. Are they happy? Are they reflective?" He considers the trove of mood messages that pour through our cell phones. That's a new frontier and a potential gold mine of behavioral data. He goes on about the advertising possibilities of music sites, including AOLs, where they can see us clicking on cheerful, sad, or inspirational songs.

I'm not so sure about that. If I click on a happy song, I say, maybe I'm just looking for a pick-me-up. Morgan shrugs. He won't know until he does more research. This means more of our data to collect and more numbers to run. Just thinking about it makes him smile. Outside the sky grows dark, and rain scatters the crowd at Rockefeller Center. As Dave Morgan heads back to his behavioral laboratory at 75 Rock, he covers his head with his hands and sprints.