

They're Watching You at Work

WHAT HAPPENS WHEN BIG DATA MEETS HUMAN RESOURCES? THE EMERGING PRACTICE OF "PEOPLE ANALYTICS" IS ALREADY TRANSFORMING HOW EMPLOYERS HIRE, FIRE, AND PROMOTE.

By Don Peck

In 2003, thanks to Michael Lewis and his best seller *Moneyball*, the general manager of the Oakland A's, Billy Beane, became a star. The previous year, Beane had turned his back on his scouts and had instead entrusted player-acquisition decisions to mathematical models developed by a young, Harvard-trained statistical wizard on his staff. What happened next has become baseball lore. The A's, a small-market team with a paltry budget, ripped off the longest winning streak in American League history and rolled up 103 wins for the season. Only the mighty Yankees, who had spent three times as much on player salaries, won as many games. The team's success, in turn, launched a revolution. In the years that followed, team after team began to use detailed predictive models to assess players' potential and monetary value, and the early adopters, by and large, gained a measurable competitive edge over their more hidebound peers.

That's the story as most of us know it. But it is incomplete. What would seem at first glance to be nothing but a memorable tale about baseball may turn out to be the opening chapter of a much larger story about jobs. Predictive statistical analysis, harnessed to big data, appears poised to alter the way millions of people are hired and assessed.

Yes, unavoidably, *big data*. As a piece of business jargon, and even more so as an invocation of coming disruption, the term has quickly grown tiresome. But there is no denying the vast increase in the range and depth of information that's routinely captured about how we behave, and the new kinds of analysis that this enables. By one estimate, more than 98 percent of the world's information is now stored digitally, and the volume of that data has quadrupled since 2007. Ordinary people at work and at home generate much of this data, by sending e-mails, browsing the Internet, using social media, working on crowd-sourced projects, and more—and in doing so they have unwittingly helped launch a grand new societal project. "We are in the midst of a great infrastructure project that in some ways rivals those of the past, from Roman aqueducts to the Enlightenment's Encyclopédie," write Viktor Mayer-Schönberger and Kenneth Cukier in their recent book, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. "The project is datafication. Like those other infrastructural advances, it will bring about fundamental changes to society."

Some of the changes are well known, and already upon us. Algorithms that predict stock-price

movements have transformed Wall Street. Algorithms that chomp through our Web histories have transformed marketing. Until quite recently, however, few people seemed to believe this data-driven approach might apply broadly to the labor market.

But it now does. According to John Hausknecht, a professor at Cornell's school of industrial and labor relations, in recent years the economy has witnessed a "huge surge in demand for workforce-analytics roles." Hausknecht's own program is rapidly revising its curriculum to keep pace. You can now find dedicated analytics teams in the human-resources departments of not only huge corporations such as Google, HP, Intel, General Motors, and Procter & Gamble, to name just a few, but also companies like McKee Foods, the Tennessee-based maker of Little Debbie snack cakes. Even Billy Beane is getting into the game. Last year he appeared at a large conference for corporate HR executives in Austin, Texas, where he reportedly stole the show with a talk titled "The Moneyball Approach to Talent Management." Ever since, that headline, with minor modifications, has been plastered all over the HR trade press.

The application of predictive analytics to people's careers—an emerging field sometimes called "people analytics"—is enormously challenging, not to mention ethically fraught. And it can't help but feel a little creepy. It requires the creation of a vastly larger box score of human performance than one would ever encounter in the sports pages, or that has ever been dreamed up before. To some degree, the endeavor touches on the deepest of human mysteries: how we grow, whether we flourish, what we become. Most companies are just beginning to explore the possibilities. But make no mistake: during the next five to 10 years, new models will be created, and new experiments run, on a very large scale. Will this be a good development or a bad one—for the economy, for the shapes of our careers, for our spirit and self-worth? Earlier this year, I decided to find out.

Ever since we've had companies, we've had managers trying to figure out which people are best suited to working for them. The techniques have varied considerably. Near the turn of the 20th century, one manufacturer in Philadelphia made hiring decisions by having its foremen stand in front of the factory and toss apples into the surrounding scrum of job-seekers. Those quick enough to catch the apples and strong enough to keep them were put to work.

In those same times, a different (and less bloody) Darwinian process governed the selection of executives. Whole industries were being consolidated by rising giants like U.S. Steel, DuPont, and GM. Weak competitors were simply steamrolled, but the stronger ones were bought up, and their founders typically were offered high-level jobs within the behemoth. The approach worked pretty well. As Peter Cappelli, a professor at the Wharton School, has written, "Nothing in the science of prediction and selection beats observing actual performance in an equivalent role."

By the end of World War II, however, American corporations were facing severe talent shortages. Their senior executives were growing old, and a dearth of hiring from the Depression through the war had resulted in a shortfall of able, well-trained managers. Finding people who had the potential to rise quickly through the ranks became an overriding preoccupation of American businesses. They began to devise a formal hiring-and-management system based in part on new studies of human behavior, and in part on military techniques developed during both world wars, when huge mobilization efforts and mass casualties created the need to get the right people into the right roles as efficiently as possible. By

the 1950s, it was not unusual for companies to spend days with young applicants for professional jobs, conducting a battery of tests, all with an eye toward corner-office potential. “P&G picks its executive crop right out of college,” *BusinessWeek* noted in 1950, in the unmistakable patter of an age besotted with technocratic possibility. IQ tests, math tests, vocabulary tests, professional-aptitude tests, vocational-interest questionnaires, Rorschach tests, a host of other personality assessments, and even medical exams (who, after all, would want to hire a man who might die before the company’s investment in him was fully realized?)—all were used regularly by large companies in their quest to make the right hire.

The process didn’t end when somebody started work, either. In his classic 1956 cultural critique, *The Organization Man*, the business journalist William Whyte reported that about a quarter of the country’s corporations were using similar tests to evaluate managers and junior executives, usually to assess whether they were ready for bigger roles. “Should Jones be promoted or put on the shelf?” he wrote. “Once, the man’s superiors would have had to thresh this out among themselves; now they can check with psychologists to see what the tests say.”

Remarkably, this regime, so widespread in corporate America at mid-century, had almost disappeared by 1990. “I think an HR person from the late 1970s would be stunned to see how casually companies hire now,” Peter Cappelli told me—the days of testing replaced by a handful of ad hoc interviews, with the questions dreamed up on the fly. Many factors explain the change, he said, and then he ticked off a number of them: Increased job-switching has made it less important and less economical for companies to test so thoroughly. A heightened focus on short-term financial results has led to deep cuts in corporate functions that bear fruit only in the long term. The Civil Rights Act of 1964, which exposed companies to legal liability for discriminatory hiring practices, has made HR departments wary of any broadly applied and clearly scored test that might later be shown to be systematically biased. Instead, companies came to favor the more informal qualitative hiring practices that are still largely in place today.

But companies abandoned their hard-edged practices for another important reason: many of their methods of evaluation turned out not to be very scientific. Some were based on untested psychological theories. Others were originally designed to assess mental illness, and revealed nothing more than where subjects fell on a “normal” distribution of responses—which in some cases had been determined by testing a relatively small, unrepresentative group of people, such as college freshmen. When William Whyte administered a battery of tests to a group of corporate presidents, he found that not one of them scored in the “acceptable” range for hiring. Such assessments, he concluded, measured not potential but simply conformity. Some of them were highly intrusive, too, asking questions about personal habits, for instance, or parental affection. Unsurprisingly, subjects didn’t like being so impersonally poked and prodded (sometimes literally).

For all these reasons and more, the idea that hiring was a science fell out of favor. But now it’s coming back, thanks to new technologies and methods of analysis that are cheaper, faster, and much-wider-ranging than what we had before. For better or worse, a new era of technocratic possibility has begun.

Consider Knack, a tiny start-up based in Silicon Valley. Knack makes app-based video games, among

them Dungeon Scrawl, a quest game requiring the player to navigate a maze and solve puzzles, and Wasabi Waiter, which involves delivering the right sushi to the right customer at an increasingly crowded happy hour. These games aren't just for play: they've been designed by a team of neuroscientists, psychologists, and data scientists to suss out human potential. Play one of them for just 20 minutes, says Guy Halfteck, Knack's founder, and you'll generate several megabytes of data, exponentially more than what's collected by the SAT or a personality test. How long you hesitate before taking every action, the sequence of actions you take, how you solve problems—all of these factors and many more are logged as you play, and then are used to analyze your creativity, your persistence, your capacity to learn quickly from mistakes, your ability to prioritize, and even your social intelligence and personality. The end result, Halfteck says, is a high-resolution portrait of your psyche and intellect, and an assessment of your potential as a leader or an innovator.

When Hans Haringa heard about Knack, he was skeptical but intrigued. Haringa works for the petroleum giant Royal Dutch Shell—by revenue, the world's largest company last year. For seven years he's served as an executive in the company's GameChanger unit: a 12-person team that for nearly two decades has had an outsize impact on the company's direction and performance. The unit's job is to identify potentially disruptive business ideas. Haringa and his team solicit ideas promiscuously from inside and outside the company, and then play the role of venture capitalists, vetting each idea, meeting with its proponents, dispensing modest seed funding to a few promising candidates, and monitoring their progress. They have a good record of picking winners, Haringa told me, but identifying ideas with promise has proved to be extremely difficult and time-consuming. The process typically takes more than two years, and less than 10 percent of the ideas proposed to the unit actually make it into general research and development.

When he heard about Knack, Haringa thought he might have found a shortcut. What if Knack could help him assess the people proposing all these ideas, so that he and his team could focus only on those whose ideas genuinely deserved close attention? Haringa reached out, and eventually ran an experiment with the company's help.

Over the years, the GameChanger team had kept a database of all the ideas it had received, recording how far each had advanced. Haringa asked all the idea contributors he could track down (about 1,400 in total) to play Dungeon Scrawl and Wasabi Waiter, and told Knack how well three-quarters of those people had done as idea generators. (Did they get initial funding? A second round? Did their ideas make it all the way?) He did this so that Knack's staff could develop game-play profiles of the strong innovators relative to the weak ones. Finally, he had Knack analyze the game-play of the remaining quarter of the idea generators, and asked the company to guess whose ideas had turned out to be best.

When the results came back, Haringa recalled, his heart began to beat a little faster. Without ever seeing the ideas, without meeting or interviewing the people who'd proposed them, without knowing their title or background or academic pedigree, Knack's algorithm had identified the people whose ideas had panned out. The top 10 percent of the idea generators as predicted by Knack were in fact those who'd gone furthest in the process. Knack identified six broad factors as especially characteristic of those whose ideas would succeed at Shell: "mind wandering" (or the tendency to follow interesting,

unexpected offshoots of the main task at hand, to see where they lead), social intelligence, “goal-orientation fluency,” implicit learning, task-switching ability, and conscientiousness. Haringa told me that this profile dovetails with his impression of a successful innovator. “You need to be disciplined,” he said, but “at all times you must have your mind open to see the other possibilities and opportunities.”

What Knack is doing, Haringa told me, “is almost like a paradigm shift.” It offers a way for his GameChanger unit to avoid wasting time on the 80 people out of 100—nearly all of whom look smart, well-trained, and plausible on paper—whose ideas just aren’t likely to work out. If he and his colleagues were no longer mired in evaluating “the hopeless folks,” as he put it to me, they could solicit ideas even more widely than they do today and devote much more careful attention to the 20 people out of 100 whose ideas have the most merit.

Haringa is now trying to persuade his colleagues in the GameChanger unit to use Knack’s games as an assessment tool. But he’s also thinking well beyond just his own little part of Shell. He has encouraged the company’s HR executives to think about applying the games to the recruitment and evaluation of all professional workers. Shell goes to extremes to try to make itself the world’s most innovative energy company, he told me, so shouldn’t it apply that spirit to developing its own “human dimension”?

“It is the whole man The Organization wants,” William Whyte wrote back in 1956, when describing the ambit of the employee evaluations then in fashion. Aptitude, skills, personal history, psychological stability, discretion, loyalty—companies at the time felt they had a need (and the right) to look into them all. That ambit is expanding once again, and this is undeniably unsettling. Should the ideas of scientists be dismissed because of the way they play a game? Should job candidates be ranked by what their Web habits say about them? Should the “data signature” of natural leaders play a role in promotion? These are all live questions today, and they prompt heavy concerns: that we will cede one of the most subtle and human of skills, the evaluation of the gifts and promise of other people, to machines; that the models will get it wrong; that some people will never get a shot in the new workforce.

It’s natural to worry about such things. But consider the alternative. A mountain of scholarly literature has shown that the intuitive way we now judge professional potential is rife with snap judgments and hidden biases, rooted in our upbringing or in deep neurological connections that doubtless served us well on the savanna but would seem to have less bearing on the world of work.

What really distinguishes CEOs from the rest of us, for instance? In 2010, three professors at Duke’s Fuqua School of Business asked roughly 2,000 people to look at a long series of photos. Some showed CEOs and some showed nonexecutives, and the participants didn’t know who was who. The participants were asked to rate the subjects according to how “competent” they looked. Among the study’s findings: CEOs look significantly more competent than non-CEOs; CEOs of large companies look significantly more competent than CEOs of small companies; and, all else being equal, the more competent a CEO looked, the fatter the paycheck he or she received in real life. And yet the authors found no relationship whatsoever between how competent a CEO looked and the financial performance of his or her company.

Examples of bias abound. Tall men get hired and promoted more frequently than short men, and make more money. Beautiful women get preferential treatment, too—unless their breasts are too large. According to a national survey by the Employment Law Alliance a few years ago, most American workers don't believe attractive people in their firms are hired or promoted more frequently than unattractive people, but the evidence shows that they are, overwhelmingly so. Older workers, for their part, are thought to be more resistant to change and generally less competent than younger workers, even though plenty of research indicates that's just not so. Workers who are too young or, more specifically, are part of the Millennial generation are tarred as entitled and unable to think outside the box.

Malcolm Gladwell recounts a classic example in *Blink*. Back in the 1970s and '80s, most professional orchestras transitioned one by one to "blind" auditions, in which each musician seeking a job performed from behind a screen. The move was made in part to stop conductors from favoring former students, which it did. But it also produced another result: the proportion of women winning spots in the most-prestigious orchestras shot up fivefold, notably when they played instruments typically identified closely with men. Gladwell tells the memorable story of Julie Landsman, who, at the time of his book's publication, in 2005, was playing principal French horn for the Metropolitan Opera, in New York. When she'd finished her blind audition for that role, years earlier, she knew immediately that she'd won. Her last note was so true, and she held it so long, that she heard delighted peals of laughter break out among the evaluators on the other side of the screen. But when she came out to greet them, she heard a gasp. Landsman had played with the Met before, but only as a substitute. The evaluators knew her, yet only when they weren't aware of her gender—only, that is, when they were forced to make not a personal evaluation but an impersonal one—could they hear how brilliantly she played.

We may like to think that society has become more enlightened since those days, and in many ways it has, but our biases are mostly unconscious, and they can run surprisingly deep. Consider race. For a 2004 study called "Are Emily and Greg More Employable Than Lakisha and Jamal?," the economists Sendhil Mullainathan and Marianne Bertrand put white-sounding names (Emily Walsh, Greg Baker) or black-sounding names (Lakisha Washington, Jamal Jones) on similar fictitious résumés, which they then sent out to a variety of companies in Boston and Chicago. To get the same number of callbacks, they learned, they needed to either send out half again as many résumés with black names as those with white names, or add eight extra years of relevant work experience to the résumés with black names.

I talked with Mullainathan about the study. All of the hiring managers he and Bertrand had consulted while designing it, he said, told him confidently that Lakisha and Jamal would get called back more than Emily and Greg. Affirmative action guaranteed it, they said: recruiters were bending over backwards in their search for good black candidates. Despite making conscious efforts to find such candidates, however, these recruiters turned out to be excluding them unconsciously at every turn. After the study came out, a man named Jamal sent a thank-you note to Mullainathan, saying that he'd started using only his first initial on his résumé and was getting more interviews.

Perhaps the most widespread bias in hiring today cannot even be detected with the eye. In a recent

survey of some 500 hiring managers, undertaken by the Corporate Executive Board, a research firm, 74 percent reported that their most recent hire had a personality “similar to mine.” Lauren Rivera, a sociologist at Northwestern, spent parts of the three years from 2006 to 2008 interviewing professionals from elite investment banks, consultancies, and law firms about how they recruited, interviewed, and evaluated candidates, and concluded that among the most important factors driving their hiring recommendations were—wait for it—shared leisure interests. “The best way I could describe it,” one attorney told her, “is like if you were going on a date. You kind of *know* when there’s a match.” Asked to choose the most-promising candidates from a sheaf of fake résumés Rivera had prepared, a manager at one particularly buttoned-down investment bank told her, “I’d have to pick Blake and Sarah. With his lacrosse and her squash, they’d really get along [with the people] on the trading floor.” Lacking “reliable predictors of future performance,” Rivera writes, “assessors purposefully used their own experiences as models of merit.” Former college athletes “typically prized participation in varsity sports above all other types of involvement.” People who’d majored in engineering gave engineers a leg up, believing they were better prepared.

Given this sort of clubby, insular thinking, it should come as no surprise that the prevailing system of hiring and management in this country involves a level of dysfunction that should be inconceivable in an economy as sophisticated as ours. Recent survey data collected by the Corporate Executive Board, for example, indicate that nearly a quarter of all new hires leave their company within a year of their start date, and that hiring managers wish they’d never extended an offer to one out of every five members on their team. A survey by Gallup this past June, meanwhile, found that only 30 percent of American workers felt a strong connection to their company and worked for it with passion. Fifty-two percent emerged as “not engaged” with their work, and another 18 percent as “actively disengaged,” meaning they were apt to undermine their company and co-workers, and shirk their duties whenever possible. These headline numbers are skewed a little by the attitudes of hourly workers, which tend to be worse, on average, than those of professional workers. But really, what further evidence do we need of the abysmal status quo?

Because the algorithmic assessment of workers’ potential is so new, not much hard data yet exist demonstrating its effectiveness. The arena in which it has been best proved, and where it is most widespread, is hourly work. Jobs at big-box retail stores and call centers, for example, warm the hearts of would-be corporate Billy Beanes: they’re pretty well standardized, they exist in huge numbers, they turn over quickly (it’s not unusual for call centers, for instance, to experience 50 percent turnover in a single year), and success can be clearly measured (through a combination of variables like sales, call productivity, customer-complaint resolution, and length of tenure). Big employers of hourly workers are also not shy about using psychological tests, partly in an effort to limit theft and absenteeism. In the late 1990s, as these assessments shifted from paper to digital formats and proliferated, data scientists started doing massive tests of what makes for a successful customer-support technician or salesperson. This has unquestionably improved the quality of the workers at many firms.

Teri Morse, the vice president for recruiting at Xerox Services, oversees hiring for the company’s 150 U.S. call and customer-care centers, which employ about 45,000 workers. When I spoke with her in July, she told me that as recently as 2010, Xerox had filled these positions through interviews and a

few basic assessments conducted in the office—a typing test, for instance. Hiring managers would typically look for work experience in a similar role, but otherwise would just use their best judgment in evaluating candidates. In 2010, however, Xerox switched to an online evaluation that incorporates personality testing, cognitive-skill assessment, and multiple-choice questions about how the applicant would handle specific scenarios that he or she might encounter on the job. An algorithm behind the evaluation analyzes the responses, along with factual information gleaned from the candidate's application, and spits out a color-coded rating: red (poor candidate), yellow (middling), or green (hire away). Those candidates who score best, I learned, tend to exhibit a creative but not overly inquisitive personality, and participate in at least one but not more than four social networks, among many other factors. (Previous experience, one of the few criteria that Xerox had explicitly screened for in the past, turns out to have no bearing on either productivity or retention. Distance between home and work, on the other hand, is strongly associated with employee engagement and retention.)

When Xerox started using the score in its hiring decisions, the quality of its hires immediately improved. The rate of attrition fell by 20 percent in the initial pilot period, and over time, the number of promotions rose. Xerox still interviews all candidates in person before deciding to hire them, Morse told me, but, she added, “We’re getting to the point where some of our hiring managers don’t even want to interview anymore”—they just want to hire the people with the highest scores.

The online test that Xerox uses was developed by a small but rapidly growing company based in San Francisco called Evolv. I spoke with Jim Meyerle, one of the company's co-founders, and David Ostberg, its vice president of workforce science, who described how modern techniques of gathering and analyzing data offer companies a sharp edge over basic human intuition when it comes to hiring. Gone are the days, Ostberg told me, when, say, a small survey of college students would be used to predict the statistical validity of an evaluation tool. “We’ve got a data set of 347,000 actual employees who have gone through these different types of assessments or tools,” he told me, “and now we have performance-outcome data, and we can split those and slice and dice by industry and location.”

Evolv's tests allow companies to capture data about everybody who applies for work, and everybody who gets hired—a complete data set from which sample bias, long a major vexation for industrial-organization psychologists, simply disappears. The sheer number of observations that this approach makes possible allows Evolv to say with precision which attributes matter more to the success of retail-sales workers (decisiveness, spatial orientation, persuasiveness) or customer-service personnel at call centers (rapport-building). And the company can continually tweak its questions, or add new variables to its model, to seek out ever stronger correlates of success in any given job. For instance, the browser that applicants use to take the online test turns out to matter, especially for technical roles: some browsers are more functional than others, but it takes a measure of savvy and initiative to download them.

There are some data that Evolv simply won't use, out of a concern that the information might lead to systematic bias against whole classes of people. The distance an employee lives from work, for instance, is never factored into the score given each applicant, although it is reported to some clients. That's because different neighborhoods and towns can have different racial profiles, which means that

scoring distance from work could violate equal-employment-opportunity standards. Marital status? Motherhood? Church membership? “Stuff like that,” Meyerle said, “we just don’t touch”—at least not in the U.S., where the legal environment is strict. Meyerle told me that Evolv has looked into these sorts of factors in its work for clients abroad, and that some of them produce “startling results.” Citing client confidentiality, he wouldn’t say more.

Meyerle told me that what most excites him are the possibilities that arise from monitoring the entire life cycle of a worker at any given company. This is a task that Evolv now performs for Transcom, a company that provides outsourced customer-support, sales, and debt-collection services, and that employs some 29,000 workers globally. About two years ago, Transcom began working with Evolv to improve the quality and retention of its English-speaking workforce, and three-month attrition quickly fell by about 30 percent. Now the two companies are working together to marry pre-hire assessments to an increasing array of post-hire data: about not only performance and duration of service but also who trained the employees; who has managed them; whether they were promoted to a supervisory role, and how quickly; how they performed in that role; and why they eventually left.

The potential power of this data-rich approach is obvious. What begins with an online screening test for entry-level workers ends with the transformation of nearly every aspect of hiring, performance assessment, and management. In theory, this approach enables companies to fast-track workers for promotion based on their statistical profiles; to assess managers more scientifically; even to match workers and supervisors who are likely to perform well together, based on the mix of their competencies and personalities. Transcom plans to do all these things, as its data set grows ever richer. This is the real promise—or perhaps the hubris—of the new people analytics. Making better hires turns out to be not an end but just a beginning. Once all the data are in place, new vistas open up.

For a sense of what the future of people analytics may bring, I turned to Sandy Pentland, the director of the Human Dynamics Laboratory at MIT. In recent years, Pentland has pioneered the use of specialized electronic “badges” that transmit data about employees’ interactions as they go about their days. The badges capture all sorts of information about formal and informal conversations: their length; the tone of voice and gestures of the people involved; how much those people talk, listen, and interrupt; the degree to which they demonstrate empathy and extroversion; and more. Each badge generates about 100 data points a minute.

Pentland’s initial goal was to shed light on what differentiated successful teams from unsuccessful ones. As he described last year in the *Harvard Business Review*, he tried the badges out on about 2,500 people, in 21 different organizations, and learned a number of interesting lessons. About a third of team performance, he discovered, can usually be predicted merely by the number of face-to-face exchanges among team members. (Too many is as much of a problem as too few.) Using data gathered by the badges, he was able to predict which teams would win a business-plan contest, and which workers would (rightly) say they’d had a “productive” or “creative” day. Not only that, but he claimed that his researchers had discovered the “data signature” of natural leaders, whom he called “charismatic connectors” and all of whom, he reported, circulate actively, give their time democratically to others, engage in brief but energetic conversations, and listen at least as much as

they talk. In a development that will surprise few readers, Pentland and his fellow researchers created a company, Sociometric Solutions, in 2010, to commercialize his badge technology.

Pentland told me that no business he knew of was yet using this sort of technology on a permanent basis. His own clients were using the badges as part of consulting projects designed to last only a few weeks. But he doesn't see why longer-term use couldn't be in the cards for the future, particularly as the technology gets cheaper. His group is developing apps to allow team members to view their own metrics more or less in real time, so that they can see, relative to the benchmarks of highly successful employees, whether they're getting out of their offices enough, or listening enough, or spending enough time with people outside their own team.

Whether or not we all come to wear wireless lapel badges, *Star Trek*-style, plenty of other sources could easily serve as the basis of similar analysis. Torrents of data are routinely collected by American companies and now sit on corporate servers, or in the cloud, awaiting analysis. Bloomberg reportedly logs every keystroke of every employee, along with their comings and goings in the office. The Las Vegas casino Harrah's tracks the smiles of the card dealers and waitstaff on the floor (its analytics team has quantified the impact of smiling on customer satisfaction). E-mail, of course, presents an especially rich vein to be mined for insights about our productivity, our treatment of co-workers, our willingness to collaborate or lend a hand, our patterns of written language, and what those patterns reveal about our intelligence, social skills, and behavior. As technologies that analyze language become better and cheaper, companies will be able to run programs that automatically trawl through the e-mail traffic of their workforce, looking for phrases or communication patterns that can be statistically associated with various measures of success or failure in particular roles.

When I brought this subject up with Erik Brynjolfsson, a professor at MIT's Sloan School of Management, he told me that he believes people analytics will ultimately have a vastly larger impact on the economy than the algorithms that now trade on Wall Street or figure out which ads to show us. He reminded me that we've witnessed this kind of transformation before in the history of management science. Near the turn of the 20th century, both Frederick Taylor and Henry Ford famously paced the factory floor with stopwatches, to improve worker efficiency. And at mid-century, there was that remarkable spread of data-driven assessment. But there's an obvious and important difference between then and now, Brynjolfsson said. "The quantities of data that those earlier generations were working with," he said, "were infinitesimal compared to what's available now. There's been a real sea change in the past five years, where the quantities have just grown so large—petabytes, exabytes, zetta—that you start to be able to do things you never could before."

It's in the inner workings of organizations, says Sendhil Mullainathan, the economist, where the most-dramatic benefits of people analytics are likely to show up. When we talked, Mullainathan expressed amazement at how little most creative and professional workers (himself included) know about what makes them effective or ineffective in the office. Most of us can't even say with any certainty how long we've spent gathering information for a given project, or our pattern of information-gathering, never mind know which parts of the pattern should be reinforced, and which jettisoned. As Mullainathan put it, we don't know our own "production function."

The prospect of tracking that function through people analytics excites Mullainathan. He sees it not only as a boon to a business's productivity and overall health but also as an important new tool that individual employees can use for self-improvement: a sort of radically expanded *The 7 Habits of Highly Effective People*, custom-written for each of us, or at least each type of job, in the workforce.

Perhaps the most exotic development in people analytics today is the creation of algorithms to assess the potential of all workers, across all companies, all the time.

This past summer, I sat in on a sales presentation by Gild, a company that uses people analytics to help other companies find software engineers. I didn't have to travel far: Atlantic Media, the parent company of *The Atlantic*, was considering using Gild to find coders. (No sale was made, and there is no commercial relationship between the two firms.)

In a small conference room, we were shown a digital map of Northwest Washington, D.C., home to *The Atlantic*. Little red pins identified all the coders in the area who were proficient in the skills that an Atlantic Media job announcement listed as essential. Next to each pin was a number that ranked the quality of each coder on a scale of one to 100, based on the mix of skills Atlantic Media was looking for. (No one with a score above 75, we were told, had ever failed a coding test by a Gild client.) If we'd wished, we could have zoomed in to see how *The Atlantic's* own coders scored.

The way Gild arrives at these scores is not simple. The company's algorithms begin by scouring the Web for any and all open-source code, and for the coders who wrote it. They evaluate the code for its simplicity, elegance, documentation, and several other factors, including the frequency with which it's been adopted by other programmers. For code that was written for paid projects, they look at completion times and other measures of productivity. Then they look at questions and answers on social forums such as Stack Overflow, a popular destination for programmers seeking advice on challenging projects. They consider how popular a given coder's advice is, and how widely that advice ranges.

The algorithms go further still. They assess the way coders use language on social networks from LinkedIn to Twitter; the company has determined that certain phrases and words used in association with one another can distinguish expert programmers from less skilled ones. Gild knows these phrases and words are associated with good coding because it can correlate them with its evaluation of open-source code, and with the language and online behavior of programmers in good positions at prestigious companies.

Here's the part that's most interesting: having made those correlations, Gild can then score programmers who *haven't* written open-source code at all, by analyzing the host of clues embedded in their online histories. They're not all obvious, or easy to explain. Vivienne Ming, Gild's chief scientist, told me that one solid predictor of strong coding is an affinity for a particular Japanese manga site.

Why would good coders (but not bad ones) be drawn to a particular manga site? By some mysterious alchemy, does reading a certain comic-book series improve one's programming skills? "Obviously, it's not a causal relationship," Ming told me. But Gild does have 6 million programmers in its database, she said, and the correlation, even if inexplicable, is quite clear.

Gild treats this sort of information gingerly, Ming said. An affection for a Web site will be just one of dozens of variables in the company's constantly evolving model, and a minor one at that; it merely "nudges" an applicant's score upward, and only as long as the correlation persists. Some factors are transient, and the company's computers are forever crunching the numbers, so the variables are always changing. The idea is to create a sort of pointillist portrait: even if a few variables turn out to be bogus, the overall picture, Ming believes, will be clearer and truer than what we could see on our own.

Gild's CEO, Sheeroy Desai, told me he believes his company's approach can be applied to any occupation characterized by large, active online communities, where people post and cite individual work, ask and answer professional questions, and get feedback on projects. Graphic design is one field that the company is now looking at, and many scientific, technical, and engineering roles might also fit the bill. Regardless of their occupation, most people leave "data exhaust" in their wake, a kind of digital aura that can reveal a lot about a potential hire. Donald Kluemper, a professor of management at the University of Illinois at Chicago, has found that professionally relevant personality traits can be judged effectively merely by scanning Facebook feeds and photos. LinkedIn, of course, captures an enormous amount of professional data and network information, across just about every profession. A controversial start-up called Klout has made its mission the measurement and public scoring of people's online social influence.

These aspects of people analytics provoke anxiety, of course. We would be wise to take legal measures to ensure, at a minimum, that companies can't snoop where we have a reasonable expectation of privacy—and that any evaluations they might make of our professional potential aren't based on factors that discriminate against classes of people.

But there is another side to this. People analytics will unquestionably provide many workers with more options and more power. Gild, for example, helps companies find undervalued software programmers, working indirectly to raise those people's pay. Other companies are doing similar work. One called Entelo, for instance, specializes in using algorithms to identify potentially unhappy programmers who might be receptive to a phone call (because they've been unusually active on their professional-networking sites, or because there's been an exodus from their corner of their company, or because their company's stock is tanking). As with Gild, the service benefits the worker as much as the would-be employer.

Big tech companies are responding to these incursions, and to increasing free agency more generally, by deploying algorithms aimed at keeping their workers happy. Dawn Klinghoffer, the senior director of HR business insights at Microsoft, told me that a couple of years ago, with attrition rising industry-wide, her team started developing statistical profiles of likely leavers (hires straight from college in certain technical roles, for instance, who had been with the company for three years and had been promoted once, but not more than that). The company began various interventions based on these profiles: the assignment of mentors, changes in stock vesting, income hikes. Microsoft focused on two business units with particularly high attrition rates—and in each case reduced those rates by more than half.

Over time, better job-matching technologies are likely to begin serving people directly, helping them

see more clearly which jobs might suit them and which companies could use their skills. In the future, Gild plans to let programmers see their own profiles and take skills challenges to try to improve their scores. It intends to show them its estimates of their market value, too, and to recommend coursework that might allow them to raise their scores even more. Not least, it plans to make accessible the scores of typical hires at specific companies, so that software engineers can better see the profile they'd need to land a particular job. Knack, for its part, is making some of its video games available to anyone with a smartphone, so people can get a better sense of their strengths, and of the fields in which their strengths would be most valued. (Palo Alto High School recently adopted the games to help students assess careers.) Ultimately, the company hopes to act as matchmaker between a large network of people who play its games (or have ever played its games) and a widening roster of corporate clients, each with its own specific profile for any given type of job.

Knack and Gild are very young companies; either or both could fail. But even now they are hardly the only companies doing this sort of work. The digital trail from assessment to hire to work performance and work engagement will quickly discredit models that do not work—but will also allow the models and companies that survive to grow better and smarter over time. It is conceivable that we will look back on these endeavors in a decade or two as nothing but a fad. But early evidence, and the relentlessly empirical nature of the project as a whole, suggests otherwise.

When I began my reporting for this story, I was worried that people analytics, if it worked at all, would only widen the divergent arcs of our professional lives, further gilding the path of the meritocratic elite from cradle to grave, and shutting out some workers more definitively. But I now believe the opposite is likely to happen, and that we're headed toward a labor market that's fairer to people at every stage of their careers. For decades, as we've assessed people's potential in the professional workforce, the most important piece of data—the one that launches careers or keeps them grounded—has been educational background: typically, whether and where people went to college, and how they did there. Over the past couple of generations, colleges and universities have become the gatekeepers to a prosperous life. A degree has become a signal of intelligence and conscientiousness, one that grows stronger the more selective the school and the higher a student's GPA, that is easily understood by employers, and that, until the advent of people analytics, was probably unrivaled in its predictive powers. And yet the limitations of that signal—the way it degrades with age, its overall imprecision, its many inherent biases, its extraordinary cost—are obvious. "Academic environments are artificial environments," Laszlo Bock, Google's senior vice president of people operations, told *The New York Times* in June. "People who succeed there are sort of finely trained, they're conditioned to succeed in that environment," which is often quite different from the workplace.

One of the tragedies of the modern economy is that because one's college history is such a crucial signal in our labor market, perfectly able people who simply couldn't sit still in a classroom at the age of 16, or who didn't have their act together at 18, or who chose not to go to graduate school at 22, routinely get left behind for good. That such early factors so profoundly affect career arcs and hiring decisions made two or three decades later is, on its face, absurd.

But this relationship is likely to loosen in the coming years. I spoke with managers at a lot of

companies who are using advanced analytics to reevaluate and reshape their hiring, and nearly all of them told me that their research is leading them toward pools of candidates who didn't attend college—for tech jobs, for high-end sales positions, for some managerial roles. In some limited cases, this is because their analytics revealed no benefit whatsoever to hiring people with college degrees; in other cases, and more often, it's because they revealed signals that function far better than college history, and that allow companies to confidently hire workers with pedigrees not typically considered impressive or even desirable. Neil Rae, an executive at Transcom, told me that in looking to fill technical-support positions, his company is shifting its focus from college graduates to “kids living in their parents' basement”—by which he meant smart young people who, for whatever reason, didn't finish college but nevertheless taught themselves a lot about information technology. Laszlo Bock told me that Google, too, is hiring a growing number of nongraduates. Many of the people I talked with reported that when it comes to high-paying and fast-track jobs, they're reducing their preference for Ivy Leaguers and graduates of other highly selective schools.

This process is just beginning. Online courses are proliferating, and so are online markets that involve crowd-sourcing. Both arenas offer new opportunities for workers to build skills and showcase competence. Neither produces the kind of instantly recognizable signals of potential that a degree from a selective college, or a first job at a prestigious firm, might. That's a problem for traditional hiring managers, because sifting through lots of small signals is so difficult and time-consuming. (Is it meaningful that a candidate finished in the top 10 percent of students in a particular online course, or that her work gets high ratings on a particular crowd-sourcing site?) But it's completely irrelevant in the field of people analytics, where sophisticated screening algorithms can easily make just these sorts of judgments. That's not only good news for people who struggled in school; it's good news for people who've fallen off the career ladder through no fault of their own (older workers laid off in a recession, for instance) and who've acquired a sort of professional stink that is likely undeserved.

Ultimately, all of these new developments raise philosophical questions. As professional performance becomes easier to measure and see, will we become slaves to our own status and potential, ever-focused on the metrics that tell us how and whether we are measuring up? Will too much knowledge about our limitations hinder achievement and stifle our dreams? All I can offer in response to these questions, ironically, is my own gut sense, which leads me to feel cautiously optimistic. But most of the people I interviewed for this story—who, I should note, tended to be psychologists and economists rather than philosophers—share that feeling.

Scholarly research strongly suggests that happiness at work depends greatly on feeling a sense of agency. If the tools now being developed and deployed really can get more people into better-fitting jobs, then those people's sense of personal effectiveness will increase. And if those tools can provide workers, once hired, with better guidance on how to do their jobs well, and how to collaborate with their fellow workers, then those people will experience a heightened sense of mastery. It is possible that some people who now skate from job to job will find it harder to work at all, as professional evaluations become more refined. But on balance, these strike me as developments that are likely to make people happier.

Nobody imagines that people analytics will obviate the need for old-fashioned human judgment in the workplace. Google's understanding of the promise of analytics is probably better than anybody else's, and the company has been changing its hiring and management practices as a result of its ongoing analyses. (Brainteasers are no longer used in interviews, because they do not correlate with job success; GPA is not considered for anyone more than two years out of school, for the same reason—the list goes on.) But for all of Google's technological enthusiasm, these same practices are still deeply human. A real, live person looks at every résumé the company receives. Hiring decisions are made by committee and are based in no small part on opinions formed during structured interviews.

One only has to look to baseball, in fact, to see where this all may be headed. In their forthcoming book, *The Sabermetric Revolution*, the sports economist Andrew Zimbalist and the mathematician Benjamin Baumer write that the analytical approach to player acquisition employed by Billy Beane and the Oakland A's has continued to spread through Major League Baseball. Twenty-six of the league's 30 teams now devote significant resources to people analytics. The search for ever more precise data—about the spin rate of pitches, about the muzzle velocity of baseballs as they come off the bat—has intensified, as has the quest to turn those data into valuable nuggets of insight about player performance and potential. Analytics has taken off in other pro sports leagues as well. But here's what's most interesting. The big blind spots initially identified by analytics in the search for great players are now gone—which means that what's likely to make the difference again is the human dimension of the search.

The A's made the playoffs again this year, despite a small payroll. Over the past few years, the team has expanded its scouting budget. "What defines a good scout?" Billy Beane asked recently. "Finding out information other people can't. Getting to know the kid. Getting to know the family. There's just some things you need to find out in person."

This article available online at:

<http://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/>

Copyright © 2014 by The Atlantic Monthly Group. All Rights Reserved.