



JANUARY 23, 2006

COVER STORY

Math Will Rock Your World

A generation ago, quants turned finance upside down. Now they're mapping out ad campaigns and building new businesses from mountains of personal data

Neal Goldman is a math entrepreneur. He works on Wall Street, where numbers rule. But he's focusing his analytic tools on a different realm altogether: the world of words.

Goldman's startup, Inform Technologies LLC, is a robotic librarian. Every day it combs through thousands of press articles and blog posts in English. It reads them and groups them with related pieces. Inform doesn't do this work alphabetically or by keywords. It uses algorithms to analyze each article by its language and context. It then sends customized news feeds to its users, who also exist in Inform's system as -- you guessed it -- math.

How do you convert written words into math? Goldman says it takes a combination of algebra and geometry. Imagine an object floating in space that has an edge for every known scrap of information. It's called a polytope and it has near-infinite dimensions, almost impossible to conjure up in our earthbound minds. It contains every topic written about in the press. And every article that Inform processes becomes a single line within it. Each line has a series of relationships. A single article on Bordeaux wine, for example, turns up in the polytope near France, agriculture, wine, even alcoholism. In each case, Inform's algorithm calculates the relevance of one article to the next by measuring the angle between the two lines.

By the time you're reading these words, this very article will exist as a line in Goldman's polytope. And that raises a fundamental question: If long articles full of twists and turns can be reduced to a mathematical essence, what's next? Our businesses -- and, yes, ourselves.

The world is moving into a new age of numbers. Partnerships between mathematicians and computer scientists are bulling into whole new domains of business and imposing the efficiencies of math. This has happened before. In past decades, the marriage of higher math and computer modeling transformed science and engineering. Quants turned finance upside down a generation ago. And data miners plucked useful nuggets from vast consumer and business databases. But just look at where the mathematicians are now. They're helping to map out advertising campaigns, they're changing the nature of research in newsrooms and in biology labs, and they're enabling marketers to forge new one-on-one relationships with customers. As this occurs, more of the economy falls into the realm of numbers. Says James R. Schatz, chief of the mathematics research group at the National Security Agency: "There has never been a better time to be a mathematician."

From fledglings like Inform to tech powerhouses such as IBM (), companies are hitching mathematics to business in ways that would have seemed fanciful even a few years ago. In the past decade, a sizable chunk of humanity has moved its work, play, chat, and shopping online. We feed networks gobs of digital data that once would have languished on scraps of paper -- or vanished as forgotten conversations. These slices of our lives now sit in databases, many of them in the public domain. From a business point of view, they're just begging to be analyzed. But even with the most powerful computers and abundant, cheap storage, companies can't sort out their swelling oceans of data, much less build businesses on them, without enlisting skilled mathematicians and computer scientists.

The rise of mathematics is heating up the job market for luminary quants, especially at the Internet powerhouses where new math grads land with six-figure salaries and rich stock deals. Tom Leighton, an entrepreneur and applied math professor at Massachusetts Institute of Technology, says: "All of my students have standing offers at Yahoo! () and Google ()." Top mathematicians are becoming a new global elite. It's a force of barely 5,000, by some guesstimates, but every bit as powerful as the armies of Harvard University MBAs who shook up corner suites a generation ago.

Math entrepreneurs, meanwhile, are raking in bonanzas. Fifteen months ago, Neal Goldman of Inform sold his previous math-based startup, a financial analysis company called CapitalIQ, for \$225 million to Standard & Poor's () (like *BusinessWeek*, a division of The McGraw-Hill Companies). And last May two brothers, Amit and Balraj Singh, sold Perabit Networks -- a company that developed algorithms for genetic research -- to Juniper Networks () for \$337 million.

In a world teeming with data, we ourselves become the math nerds' most prized specimens. Researchers at Aetna Health Care, Amazon.com (), and many other companies are piecing together mathematical models of customers and employees. Some models predict what music we'll buy, others figure out which worker is best equipped for a particular job. For now, these models are crude, the digital equivalent of stick figures. But over the coming decade, each of us will give birth to far more fleshed out simulations of ourselves. We'll be modeled as workers, shoppers, voters, and patients. Some of the simulations will have our names and credit cards attached, perhaps a few genetic details. In others, our identities will be shielded. Many of these models will be eerily accurate and others laughably off mark. But companies and governments will use them all the same to predict how to sell us things, steer us clear of diseases, and ramp up our productivity. And yes, they'll try to use them to keep us from hijacking airplanes or detonating bombs.

This mathematical modeling of humanity promises to be one of the great undertakings of the 21st century. It will grow in scope to include much of the physical world as mathematicians get their hands on new flows of data, from 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. "We turn the world of content into math, and we turn you into math," says Howard Kaushansky, CEO of Boulder (Colo.)-based Umbria Inc., a company that uses math to analyze marketing trends online.

The Dark Side

This industrial metamorphosis also has a dark side. The power of mathematicians to make sense of personal data and to model the behavior of individuals will inevitably continue to erode privacy. Merchants will be in a position to track many of our most intimate purchases, and employers will be able to rank us not only by productivity, but by wasted minutes. What's more, the rise of math can contribute to a sense that individuals are powerless, a foreboding that mathematics, from our credit rating to our genomic map, spells out our destiny.

Debates over these issues have flared up many times in the past decade. And they are sure to rear up again as the U.S. Congress investigates the Bush Administration's mining of phone and Internet traffic in its effort to sniff out terrorists. But the merger of sophisticated data mining and higher math has tremendous power to conquer mankind's scourges as well. As Jack Einhorn, chief technical officer of Inform, puts it: "The next Jonas Salk will be a mathematician, not a doctor."

The clearest example of math's disruptive power is in advertising. There Google and other search companies built on math are turning an industry that grew on ideas, hunches, and personal relationships into a series of calculations. They can pull it off because, quite simply, they know where their prospective customers are browsing, what they click on, and often, what they buy. Internet companies use this data not only to profile customers but also to pitch for more contracts. Some 18 months ago, 30 blue-chip companies, from Procter & Gamble Co. () to Walt Disney Co. (), underwent a series of tests promoted by the Interactive Advertising Bureau, an industry group. These studies crunched consumer data to measure the effectiveness of advertising in a host of media. The results came back in hard numbers. They indicated, for example, that Ford Motor Co. () could have sold an additional \$625 million worth of trucks if it had lifted its online ad budget from 2.5% to 6% of the total. Ford responded vigorously: Last August it announced plans to move up to 30% of its \$1 billion ad budget into media targeted to individual customers, half of it through online advertising. Such moves are sure to generate even more data, giving greater clout to the numbers people.

Just ask Imran Khan, the director of search advertising at E-Loan, an online lender. An accountant by training, Khan has turned the advertising operation into an enormous statistical laboratory. Like most others in the industry, he started three years ago by bidding on keywords on the major search engines. Over time, Khan's team has amassed a portfolio of 250,000 key words and phrases. Each time a Web surfer types one of those words in a search engine, an E-Loan ad appears next to the results, and Khan's team pays the price bid for each click. But running search-based ads is hardly a static process. Working with Efficient Frontier Inc., an analytics startup in Silicon Valley, Khan crunches his stash of words, calculating the return on investment for each one and tweaking thousands of bids hour by hour. He spends \$15 million a year -- half of E-Loan's ad budget -- and he accumulates massive feedback from customers.

As data mavens gather more information about customers, they gain muscle to demand changes inside companies. Take media. With banks of consumer data continuing to swell, quants on the marketing side will be able to provide editors and program managers with increasingly sophisticated statistical models, telling them which types of TV scenes or articles appeal most to certain demographic groups. As publishers seek to optimize profits and performance, data analysis will grow in importance. The risk: It gives math-based analysts, not to mention advertisers, a growing role in editorial decisions. "It puts a question mark around the classic church-state divide in the media," says Rex Briggs, founder of Marketing Evolution, the San Francisco company that conducted the 30 advertising studies.

Rising flows of data give companies the intelligence to home in on the individual customer. Internet marketers are the natural leaders, but traditional businesses are following suit. Gary W. Loveman, CEO of casino giant Harrah's Entertainment Inc. () and a former Harvard B-school professor, has led the company to build individual profiles of millions of Harrah's customers. The models include gamblers' ages, gender, and Zip codes, as well as the amount of time they spent gambling and how much they won or lost. These data enable Harrah's to study gambling through a host of variables and to target individuals with offers, from getaway weekends to gourmet dining, calculated to maximize returns. In the last five years, Harrah's has averaged 22% annual growth, and its stock has nearly tripled.

Pi in the Sky

Math is also positioned to shake up investigations. Whether in law, journalism, or criminal detective work, sleuths have relied for centuries on the human brain to pick through strands of disparate evidence and to find patterns. Sherlock Holmes sometimes looked for them in plumes of pipe smoke. And why not? Even today, no machine could sift through the photos, names, words, geographical coordinates, snippets of video -- that towering mountain of information that computer scientists call "unstructured data."

But some companies are making inroads. Colorado's Umbria has built a system to sift through millions of blogs in real time, looking for market intelligence. Umbria breaks down English messages into the smallest components -- words, phrases, grammar, even emotions -- and turns them into math. Then it analyzes the content, looking for trends. It can give cell-phone companies or fast-food restaurants the latest buzz on an ad campaign or a new sandwich.

Sometimes it uncovers trends researchers weren't even looking for. A recent search for Gatorade (), for example, showed that large numbers of young men look to it as a cocktail mixer in hopes that the electrolytes in the sports drink will ease hangovers. In the future, similar insights could uncover countless other patterns. They could help bankers spot entrepreneurs careening toward bankruptcy or point police toward sociopaths planning terrorist acts.

At the Sunnyvale (Calif.) campus of Yahoo, chief researcher Prabhakar Raghavan heads a team of 100 mathematicians and computer scientists. Scribbling on a white board covered with equations, Raghavan describes Yahoo's immense pool of data, featuring the online activity of 200 million registered customers, as Yahoo's most precious resource. There is a whole world of uninvented businesses, he believes. They'll come into being as Yahoo discovers new ways to satisfy the urges, curiosities, and desires of this customer base. The hints of these future businesses float in the oceans of Yahoo's data. Raghavan's mandate is to sift through that data and form new connections among consumers, e-marketers, and advertisers. Better algorithms, he says, "are critical to survival."

As companies continue to receive ever more data about their own processes and their workers, many will use math to boost productivity and shake up the workplace. This doesn't have to be limited to one company. Vast globe-spanning projects can be modeled, then cut into tiny pieces, with each task going to the best-qualified person. Pierre Haren, CEO of Paris-based ILOG, a company that turns customers' raw data into visual displays, foresees virtual assembly lines. "We'll have systems that tap our knowledge by the minute," he says. "Productivity could rise by a factor of 10."

That may sound like more digital pi in the sky. It's actually an extension of mathematical modeling that's been going on for half a century at companies like IBM. Following World War II, researchers at Big Blue constructed a mathematical model of the company's supply chain. It featured raw materials, trucking schedules, and manufacturing plants. Once the company had a working model, it put it through a mathematical analysis called optimization. The results suggested specific improvements, and the rejiggering sped up IBM's operations and cut costs. Decades later, IBM turned optimization into a leg of its services business. Today, IBM consultants are implementing math-based blueprints to upgrade steel mills in China and revamp operations at the U.S. Postal Service.

If you look back at those old supply-chain programs, there's one important element nearly absent: the human being. People were represented by numbers and were largely interchangeable. The mathematicians' systems lacked the data to provide more detail. And even if they had amassed a huge pile of it, the primitive computers of the time would have choked on it.

Now, though, at an IBM research center a half-hour's drive north of New York City, a 40-member team of researchers is scrutinizing people. The team combines data miners, statisticians, and experts in operations research. The current project is to refocus the supply-chain programs on 50,000 of the consultants in IBM's services division. That means that instead of modeling machines, furnaces, and schedules, they're building models of their colleagues.

A leader in this effort is Syrian-born Samer Takriti, who came from the math shop at Enron Corp. Years before the accounting mess brought the company down, Enron pioneered advanced math to create new financial markets. IBM hired Takriti for a second stint in 2000, a year before Enron's collapse. Big Blue named him senior manager of stochastic analysis. That's the science of incorporating random behavior, including the meanderings of humans, into math models.

The first step in modeling IBM's workforce, says Takriti, is to harvest all sorts of data from company records. To date, these professionals are divided into 200 categories. But the math team is hunting for richer personal details. A survey of company e-mail, Takriti says, could highlight communication links between employees and the informal social networks that they create. Workers who e-mail each other a lot are more likely to work well together. Calendar data could show which consultants have more free time. Eventually, by tracking mobile devices, the system will know exactly where the consultants are. And when a contract comes through for, say, a new call center in Manila, IBM's optimization program will cull through its global database and put together the perfect team.

Calculus Ahead

The program will take years to implement. "People are complicated," says Takriti. "If you have a system, they figure out how to game it. Machines never do." This means the researchers will have to factor in a certain amount of human behavior, from lowballing sales targets to "accidentally" deleting a rival's snazzy report. This threatens to make the models fuzzier. Still, if IBM's operation yields fruit, you can bet that Big Blue will be offering similar workforce modeling services to its customers.

Eventually IBM-like programs will reach us. And it doesn't take much imagination to see where that can lead. Managers will operate tools not only to monitor employees' performance but also to follow their movements and drive up productivity. Perhaps, like Internet marketers, they'll even have the tools to link these initiatives to revenue or return on investment. On the other side, consumers will be armed with ever more data, from predictive models of real estate markets to patient mortality charts for comparing different oncologists.

It adds up to an era chock-full of numbers. Outfitting students with the right quantitative skills is a crucial test facing school boards and education ministries worldwide. This is especially true in America. The U.S. has long leaned on foreigners to provide math talent in universities and corporate research labs. Even in the post-September 11 world, where it is harder for foreigners to get student visas, an estimated half of the 20,000 math grad students now in the U.S. are foreign-born. A similar pattern holds for many other math-based professions, from computer science to engineering.

The challenge facing the U.S. now is twofold. On one hand, the country must breed more top-notch mathematicians at home, especially as foreigners find greater opportunities abroad. This will require revamping education, engaging more girls and ethnic minorities in math, and boosting the number of students who make it through calculus, the gateway for math-based disciplines. "It's critical to the future of our technological society," says Michael Sipser, head of the mathematics department at Massachusetts Institute of Technology. At the same time, school districts must cultivate greater math savvy among the broader population to prepare it for a business world in which numbers will pop up continuously. This may well involve extending the math curriculum to include more applied subjects such as statistics.

Private Lives

One significant challenge to the math revolution is to build new businesses from data without sacrificing privacy. If customers, patients, and workers have reason to fear that the intimate details of their lives are floating around in databases, they'll likely work to lock up their information or move it off network. This could disrupt efforts to use math and data mining to fight disease and to battle terrorism. The goal now is to create systems that share group information while shielding the individual. This way, researchers working with a database of HIV or breast cancer patients, for example, could study them by age, race, income, medication, education, and neighborhood without zeroing in on one person.

Mathematicians are at the heart of the privacy battle -- on both sides. In Microsoft Corp.'s () laboratories near San Francisco, Cynthia Dwork, a cryptographer, is working on a system to shield individuals while making use of the data. Dwork and her team are encasing each person's records in a camouflage of numbers that she calls "noise." Think of looking at a picture of a crowd. As soon as you zoom in on an individual face, it becomes pixelated. It's a promising approach, but even Dwork admits that mathematically gifted hackers can continue to pry open doors that she and her team slam shut. "As cryptographers, we know the power of the adversary," she says.

Math's other problem? Sometimes it's just not as smart as advertised. As mathematicians expand their domain into the humanities, they're working with new data, much of it untested. "It's very possible for people to misplace faith in numbers," says Craig Silverstein, director of technology at Google. The antidote at Google and elsewhere is to put mathematicians on teams with specialists from other disciplines, including the social sciences.

Just as mathematicians need to grapple with human quirks and mysteries, managers and entrepreneurs must bone up on mathematics. Midcareer managers can delegate much of this work to their staffers. But they still must understand enough about math to question the assumptions behind the numbers. "Now it's easier for people to bamboozle someone by having analysis based on lots of data and graphs," says Paul C. Pfleiderer, a finance professor at the Stanford Graduate School of Business. "We have to train people in business to spot a bogus argument."

And to spot opportunities. As more of the world's information is pooled into mathematics, the realm of numbers becomes an ever larger meeting ground. It's a percolating laboratory full of surprising connections, and a birthplace for new industries. Yes, it's a magnificent time to know math.

By Stephen Baker, with Bremen Leak in New York

Copyright January 23, 2006 by The McGraw-Hill Companies.

This implies no endorsement, either tactic or expressed, of any company, product, service or investment opportunity.

Electronic redistribution of this article is strictly prohibited. To obtain rights to make black-and-white photocopies of this article please contact the Copyright Clearance Center at 978-750-8400.