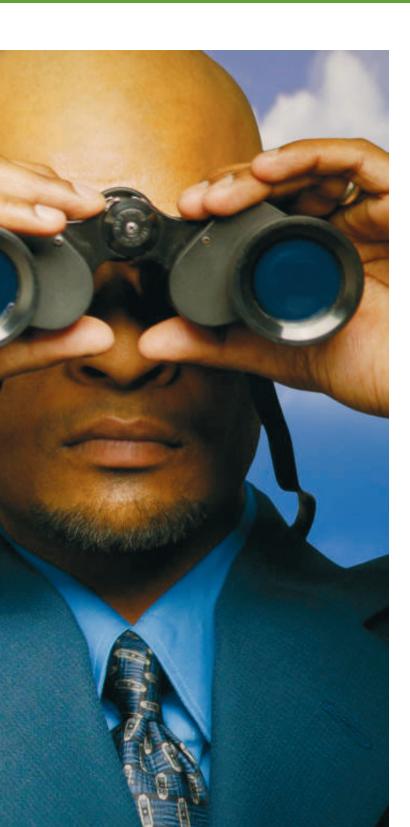ARTICLE

DATA MINE OR DATA YOURS? INFO WARS AND THE ESCALATING ARMS RACE

Eric Siegel, author *Predictive Analytics*

"Data is such a unique thing to ship, we have a special word for its telekinetic mode of transport. We call it telecommunication."

DATA MATTERS. IT'S THE VERY ESSENCE OF WHAT WE CARE ABOUT.

Personal data is not equivalent to a real person—it's much better. It takes no space, costs almost nothing to maintain, lasts forever, and is far easier to replicate and transport. Data is worth more than its weight in gold—certainly so, since data weighs nothing; it has no mass.

Data about a person is not as valuable as the person, but since the data is so much cheaper to manage, it's a far better investment. Alexis Madrigal, senior editor at *The Atlantic*, points out that a user's data can be purchased for about half a cent, but the average user's value to the Internet advertising ecosystem is estimated at US$1,200 per year.

Data's value—its power, its meaning—is the very thing that also makes it sensitive. The more data, the more power. The more powerful the data, the more sensitive. So the tension we feel around data governance is inevitable. If nobody cared about some piece of data, nobody would try to protect it, and nobody would want to access it or even bother to retain it in the first place. Data mining industry leader John Elder reflects, "The fact that it's perceived as dangerous speaks to its power; if it were weak, it wouldn't be a threat."

Ever since the advent of paper and pen, this has been the story. A doctor scribbled a note, and the battle to establish and enforce access policies began.

ARTICLE

But now, digital data travels so far, so fast, between people, organisations, and nations. Combine this ability of data to go anywhere at almost no cost with the intrinsic value of the stuff that's traveling, and you have the makings of a very fickle beast, a swarm of gremlins impressively tough to control. It's like trying to incarcerate the X-Men's superhero Nightcrawler, who has the ability to teleport. It's not confined to our normal three dimensions of movement, so you just can't lock it up.

Data is such a unique thing to ship, we have a special word for its telekinetic mode of transport. We call it telecommunication.

Data wants to spread like wildfire. As privacy advocate David Sobel put it, "Once information exists, it's virtually impossible to limit its use. You have all this great data lying around, and sooner or later, somebody will say, 'What else can I do with it?'"

This new, powerful currency proves tough to police. A shady deal to share consumer records is completed with no more than the press of a button—no covert physical shipment of goods required.

THE POWER STRUGGLE OVER DATA

 `The Internet of free platforms, free services, and free content is wholly subsidized by targeted advertising, the efficacy (and thus profitability) of which relies on collecting and mining user data.´
-Alexander Furnas, writer for *The Atlantic*.
The stakes increase and the opponents' resolve hardens like cooling lava.
In one corner we have privacy advocates, often loath to trust organisations, racing to squeeze shut data's ebb and flow: contain it, delete it, or prevent it from being recorded in the first place.

In the other corner we have the data hustlers, salivating: the hoarders and opportunists. This colorful group ranges from entrepreneurs to managers, techies, and board members.

As a proponent and practitioner of predictive analytics, I am a data hustler. I want your data—not because I want to know more about you individually, but because I can learn from it to predict human behaviour in general.

A SURPRISING TURN OF EVENTS

We predictive data hustlers have been surprised and stymied by recent incidents. We are learning that it's not about the input—precious personal data—but rather the output – newly inferred predictions about one's future behaviour. What we techies sometimes fail to consider is that your intentions for tomorrow can amount to undisclosed, private truths. Privacy advocates lodge concerns as Hewlett-Packard infers an employee's intent to resign, retailer Target deduces a customer's pregnancy, and law enforcement in Oregon and Pennsylvania foretells a convict's future repeat offense.

Google nimbly sidestepped around this landmine, sacrificing a significant boon in the name of privacy by halting its work on the automatic recognition of faces within photographs. When he was Google's CEO, Eric Schmidt stated his concern that face recognition—the application of predictive analytics' core methods to 'predict' (detect) who is in a photograph—could be misused by organizations that identify people in a crowd. This could, among other things, ascertain people's locations without

> " embracing predictive analytics also enacts a less visible but universal change that effects us all: the 'predicted masses' as a whole are unwittingly yielding as new power shifts to organisations. "

their consent. Schmidt acknowledges that other organisations will continue to develop such technology, but Google chooses not to be behind it.

Beyond threatening the privacy of some individuals, embracing predictive analytics also enacts a less visible but universal change that effects us all: the `predicted masses' as a whole are unwittingly yielding as new power shifts to organisations. "We grant private entities—with no interest in the public good and no public accountability—greater powers of persuasion than anyone has ever had before and in exchange we get free email," writes Alexander Furnas. The Obama campaign's use of predictive analytics to help win the 2012 presidential election is a recent example of predictive power wielding tremendous influence.

Technology advancement promises great strides for the world, and yet brings threats to civil liberties. As mere mortals, are we consumers, patients, and voters overly susceptible to the invisible powers of advanced mathematics? With so much to gain from data, how can we adeptly add caution to the manner in which we proceed? Litigators, activists, and citizens have some catching up to do. Only by learning more about how data exerts power can the societal issues be identified and addressed.