Telling Tales with Inspector PredPol

Xavier Raufer

1. Predictive Policing: Myth and Reality

It is always the same story: a journalist, who is either gullible or has taken a backhander, is in seventh heaven as he or she reports the news we have all been waiting for. In the fall of 2015 it was the turn of a certain Kevin, author of a piece on the French website *Science Post* boasting: "Crime-Predicting Artificial Intelligence is in the Pipeline." Kevin's article featured the predictable reference to the film *Minority Report*, followed by a spate of technical terms intended to mesmerize and daze the reader. We also learn that we are soon going to be "arresting criminals … before they commit a crime." Is this the stuff of science fiction? No, it is tomorrow's world.

Except for one thing: it is no easier to model uncertainty today than it was when Aristotle was alive (and our conception of time is still underpinned by ancient Greece). Let us put it another way: we would not have been able to predict September 11 even if we had had a document database containing everything we know on terrorism from the year dot to September 10, 2001. It is a bit embarrassing to have to keep on saying it, but if yesterday's known automatically solved tomorrow's unknown, we would all be lottery winners…

The gullible journalists who peddle these tall tales on behalf of software dealers seem to be ignorant about how predictive policing systems work. So, let us give them a hand: algorithms sift through data on past crimes with the aim of anticipating future offenses. But—and this is something we really need to stress—this is absolutely impossible! So, what does predictive software actually do? It works on the following assumption: since a crime was committed in such-and-such a place yesterday, another felony might be carried out in the same spot tomorrow. But that is not predictive reasoning; it is wishful thinking.

And just how credulous are the individuals who spread these fairy tales? In August 2015, the French newspaper *Le Parisien* made the bold claim that predictive policing software in Munich had "cut burglaries by 30%." But what do we find when we check the official statistics supplied by the *Bundeskriminalamt* (the Federal Criminal Police Office)? Burglaries increased by 16.9% in Germany between 2003 and 2012. How about in 2014, when Munich's famous miracle software was introduced? The burglary rate did not even level out, rising once more, this time by 1.8%. The "predictive" magic wand simply triggered a displacement effect: once burglars felt they were being watched, they began plying their trade elsewhere, going

[A] CNAM Paris

to neighboring towns to rob and steal. Over Germany as a whole, there was zero effect.

The miracle software can at best help the police organize their work better and respond more often in the right place at the right time—although, of course, there is nothing "predictive" about this. What is more, the positive effect can only ever be temporary, because the way we react as human beings has not changed since we still lived in caves: if we think we are being watched, we alter our behavior. And it makes no difference whether we are being observed with the naked eye or electronically. So you have paid tens of thousands of euros for game-changing software that, 6 months later, does not tell you anything worth knowing. There is a name for that: it is called a scam.

These observations are backed up by the latest research from the US. IARPA—the Intelligence Advanced Research Projects Activity—quietly began looking at an IT tool in March 2015 designed to "model and predict rare events." IARPA, do not forget, is the high-tech laboratory run by the US intelligence service, and if IARPA's on the lookout for a forecasting tool, it can only mean one thing: it does not already have one. But if Kevin's report that "Crime-Predicting Artificial Intelligence is in the Pipeline" were really true, then IARPA would have already gone 95% of the way, since sporadic serious crimes are exactly the "rare events" that US intelligence would like to be able to predict.

There is only one possible conclusion: current predictive policing software, which some people in the media are gullible enough to big up, is considered garbage by the elite scientists and researchers who work for the US intelligence community, QED.

2. Can Super Technology Replace Human Intelligence More Globally?

There are people out there, it is true, who worship machine intelligence, who eulogize the capacity of computers to solve any problem at far greater speed than our tiny human brains can. But the real problem when dealing with these zealots is not with technological progress per se and whether it is a wonderful or worrying thing. The technology is already out there, and that is the way it is. None of us, after all, wants to go back to the Stone Age, still less this author—delighted as he is by the world of computers, these desktops or laptops that are like an extension of the human brain, and which make my working conditions so much easier!

Nor does the problem lie with Silicon Valley and its mystical glorification of its cyber creations, or the idea that Google almighty will one day rule the world, finally taming the *bête humaine* and keeping us from reigning supreme as has been our custom. Technology is like a religion for Californians, and we should just leave them to get on and play with their cyber cults in their Singularity University.

No, the nub of the problem is the way our press barons (whose deep pockets bought our newspapers) treat the whole business. Once-serious journalists are now the lackeys of the new economy billionaires. They devote their time to PR, extending

a rapturous welcome to every technological breakthrough (including e-commerce) and—by an amusing coincidence—to everything that lines the pockets of the tycoons who employ them.

Here is one of many examples of this phenomenon: the media gives a huge amount of publicity to predictive policing when, in fact, the software is about as effective as the technology behind on-line dating sites—they both use the same algorithms. But, yet again, no one dares to criticize! It is a high-tech frenzy, the one and only road to progress. Anyone who harbors any doubts is nothing more than a fuddy-duddy.

For further enlightenment, just read what the press barons have had to say about Malaysia Airlines Flight 370 since it disappeared without trace in March 2014. We have had live coverage of the staggering failure of the "eye in the sky", with Washington reduced to begging internet users to go check out the world's oceans. More than 2 years after the tragedy, we are none the wiser about the fate of the aircraft, and the so-called "news" media has not got a word to say about it.

All the talk at the outset was about the miracle of predictive technology. Then disaster strikes and everyone looks the other way. There is no denying that we have heard a lot about Flight 370 and its mysterious disappearance. Yet there has not been a word about the fact that these amazing state-of-the-art systems have failed miserably to find a 300-ton airliner with a wingspan and length of some 70 m. But is not this the selfsame spy technology that can (allegedly) pick out the brand name on a pack of cigarettes that has been discarded on the sidewalk? This is where the real scandal is to be found: the media has buried this glaring high-tech failure in silence when they normally shower it with praise on a daily basis.

3. Media Ecstasy and Predictive Policing

If you liked the subprime real-estate rip-off, then you are going to love the predictive policing scam. It is a security shakedown that, naturally enough, has electrified the media. The French paper *Journal du Dimanche* expressed its amazement at "the machine that can detect crime," while *Le Monde Magazine* showcased "the software that predicts offenses before they take place." But, if we look a little more closely, the media excitement is not quite as spontaneous as it seems: all the articles are almost identical, and the whiff of a free plug can be smelt a mile off.

Most of the newspaper articles on predictive policing mention *Minority Report* (the film based on a novel by Philip K. Dick) as proof of how serious the technology is. But our enthusiastic journalists clearly do not know anything about the movie, which does not have anything to do with predictive policing! In Dick's work, three psychics or "precogs" can anticipate homicides before they occur, and they alert the police—which is not the same thing at all.

Let us take a look at those cyber "Bernie Madoffs" who claim they can "predict" crimes and terrorist attacks so we can emphasize the irrefutable evidence (developed below) at the outset. Both today and in the long term, certain irreducibly

complex phenomena are and will continue to be impossible to predict—earthquakes are a good example. But will it be possible to predict other risk events? In the best-case scenario, only very slightly. Let us have a look.

In fields where we think we can make predictions, and where we have been doing so for quite some time—the economy and finance, the weather, seismology, etc.—the results of the forecasts have been pretty catastrophic, especially when it comes to evaluating risk. Here is the proof: before the subprime crisis, all the rating agencies (Standard & Poor's, Moody's, and so forth) had quantitative estimates in their possession on the risk of house buyers defaulting on their loans.

These risks were assessed by means of quantitative analysis, a discipline that is derived from the physics of probabilities, and that is supposed to "scientifically" control the risks of trading. But when the crisis broke, the real risk turned out to be two hundred times worse than the agency predictions! As one expert chuckled, thinking you are risk-free based on the estimates of rating agencies is tantamount to smearing yourself with sunscreen to protect yourself from a nuclear blast.

None of this prevented the cyber Madoffs from conjuring up a new trick, which in the US is called crime prediction or predictive policing—PredPol if you want to be cool. This predictive analytics software is supposed to be able to forecast crimes, and even, why not, political crises, revolutions and enemy attacks.

Let us read the articles that the press devotes to predictive policing. According to one paper, it is "the software that can predict crimes … XXX (the name of the software) has arrived in the UK". We also learn that it can "forecast where and when criminals might strike" or "predict where burglaries, robberies, and assaults will take place in the future … with convincing results."

Now, the cyber Madoffs (who, according to the gullible media, claim to "predict crimes") use the same type of quantitative analysis and "predictive algorithms" that brought Wall Street crashing down! And how is this crime-predicting software powered? It is powered by "an algorithm designed to forecast where and when a crime will occur using a database of past offenses" or by "using historical statistics" and "criminal databases from the 1960s." Without exception, each of these different types of software derives its reference material from the past; they are all powered by data that is strictly retrospective.

The software sucks up anything it finds lying around on the internet, and the ensuing mass—which is operated by guesswork—is dressed up in high-sounding, high-tech names, such as big data or data mining, before the raw material is fed through an algorithm cruncher. And the results are plausible: there is likely to be a temporary improvement in police performance until the villains turn the displacement effect to their advantage (as is their wont). But this has nothing to do with predictive power—and here is why.

The crucial question—the question that our media in their excitement ignore (knowingly or not)—is: Does our knowledge of the past mean we can predict the future? Take an example: Does the weather yesterday guarantee what the weather will be like tomorrow? Obviously not, because there is an element of uncertainty,

the immutable and non-negotiable part of any prediction—which the cyber Madoffs and their devotees are careful not to mention. It is unlikely, too, that things will get any better in the future as systems such as PredPol are gradually upgraded. Quite the opposite, in fact: the situation can only get worse. Experts are of the opinion that so-called predictive software and algorithms have a high capacity for "pre-formatting" and influencing reality. This boosts their apparent validity and persuades customers that everything is working fine. Meanwhile, out there in the field and in actual life, where the real thugs and drug dealers are on the make, nothing much has changed.

This technological illusion, this crime-prediction software, is based on nothing more than guesswork, and is rooted in a simple idea: what happened yesterday will happen again tomorrow. A gang was at work in such-and-such a neighborhood on Monday. So, on Tuesday we will set up an ambush where the software tells us to, and—hey, presto, we'll bust 'em for sure! But this has not got anything to do with predicting and everything to do with circular reasoning.

PredPol-type algorithms are also used for an entirely different type of service: on-line dating. Here is what an expert has to say: "People pay for these dating services but, after taking a closer look at them, the algorithms that are supposed to help you find your ideal partner almost certainly do not work." Now, after on-line lovers, the cyber Madoffs are targeting the police and politicians. Welcome to the club for suckers.

4. Is it Possible, in a Wider Sense, To Predict Everything?

The field of prediction is one of the worst excesses of this cyber propaganda. Mainframe computers, proper algorithms, big data—wrap it all up in the right system and you will be able to predict anything and everything. So, thanks to Wikipedia for instance, and the right mathematical model, we can (it is claimed) accurately predict the success of a film at the box office 1 month in advance. And the same recipe can also be used in the food industry for new sodas and sandwiches.

Some gurus have even declared that by 2030 we will have global precognition machines that will rid the world of all its negative clutter. According to the software sales reps, specialists already have the power to predict whether you are going to "click, buy, lie or die." These reps argue that anticipating human behavior means we can take better decisions, dispel financial risk, strengthen the public health system, destroy spam, stimulate sales and, naturally enough, lead the charge against crime. (No details yet about an anti-worm treatment or the ability to cure scrofula. But do not give up hope.)

Let us shift focus to crime prediction in particular. Since around 2010, a burgeoning number of articles in the English-speaking print and electronic media have been spinning the same story: soon we will be able to predict crime in the same way that we can forecast the weather. Specialized software such as PredPol will enable the police to predict offenses. In fact, the system is already up and running in California! The papers tell us that, when the software was tested, recorded crime dropped by 12% and robberies by 27%. And now everyone's rushing to jump on the bandwagon: not

just Silicon Valley but also the military, our universities, the world of publishing and even Hollywood!

Let us start with some of the books that have been published on the subject. *Predictive Analytics for Dummies* is fairly typical, the authors introducing us to the holy grail that is predictive technology with great enthusiasm and cut-and-dried assertions. But what is actually between the covers? The book, which is full of business school marketing banalities, is based on nothing more than extrapolation. Neither the exhaustive table of contents (which runs to seven pages) nor the index (which is twice as long) features a single reference to the idea of time or temporality—which are clearly crucial concepts in the field of forecasting, whether it is predicting, anticipating, or making assumptions.

Here is another book that is typical: *Predictive Analytics*, a work by a gimmick merchant who never even defines the term "prediction". Observing behaviors or habits, using your savvy, optimizing, extrapolating, and estimating probabilities—is that all forecasting or predicting?

Take an example: when a store sees a customer buying a maternity dress, do you really need a super computer to ask her if she also wants a baby bottle and diapers? Likewise, when a reader orders a detective novel on a website, do you need to be an internet genius to suggest similar titles? Is not it the same thing as the commonplace: "Customers who bought this item also bought…" that we see on commercial websites? But that is how this rather crude arrangement works, endeavoring to make sense of data that is diffuse, chaotic, and collected in bulk. It is a system that has everything to do with optimization, or common-sense marketing, but nothing to do with predicting. A case in point: the speed and power of computers is (in this instance) undeniably superior to that of the human brain, and software sorts e-commerce customers into four categories:

> A—People who buy a product but ignore the advertising (put them on the back burner).
>
> B—People who only buy a product when there is no advertising (on the back burner).
>
> C—People who surf the internet without buying anything on-line (on the back burner).
>
> D—People who do not buy anything without advertising but do buy if they see it:these "definite receptive customers" are put in a specific database then bombarded with advertising.

The repeated use of the word "predictive" in *Predictive Analytics* will exert a hypnotic effect on the reader. But nothing in the book really helps to predict anything. And seeing as book series bearing this title are all the rage, what we have here is more

akin to *Behaviorism for Dummies*.

Let us leave the books to one side, and focus on strategy. We discovered in November 2013 that the US military in Afghanistan was trying out a "new predictive model" called the Global Events Database, designed by a political science professor from Pennsylvania State University. This software, we were told, "collects news on the internet" and "catalogs all sorts of events, from local elections to genocides." It then extracts "short and long-term predictions" from the data that might be useful for "managing crises" and "predicting conflict levels in Afghanistan."

TV was quick to get in on the act: in the US, the series *Person of Interest* (broadcast in France on TF1) had 14 million viewers on average on CBS TV in 2012. The show tells the story of a computer genius who invents a machine that can thwart terrorist attacks and predict heinous crimes. What a brilliant contraption! It is an intelligent computer that even has feelings—artificial intelligence with a heart. Our hero battles with the government, the mafia, and corrupt New York police officers, and (of course) prevents crimes before they are committed. *Person of Interest*, say the critics, has something of *Minority Report* about it in yet another reference to the celebrated film, which is the immutable marker of the predictive rip-off. Even US universities are now rushing to offer students diplomas or masters in predictive analytics.

Predicting Crime: the Tales the Media are Quick to Tell

The articles in the press all claim that predictive policing "reduces the crime rate by analyzing data on criminal offenses and where they are committed" or that "the predictive method can be effective for people at risk." In similar vein, they report that: "Maryland is generating electronic predictions on criminals on bail to see who will kill and who will be killed"; "Scientists and the police have designed predictive systems that can tell which convicted killers will strike again"; "It is now possible to use sophisticated computer analysis to predict where and when crimes will be committed"; and finally: "Predictive policing programs based on algorithms and historical data can guess the location and nature of future crimes."

These articles have a strong whiff of marketing about them rather than critical analysis—so much so that you would say they are more like editorial advertisements. Take a look at the following examples: "Predictive software used by the police is twice as effective as a human analyst working with the same data"; "A predictive policing tool has reduced burglary by a third in 5 months in a Los Angeles neighborhood"; and: "Throughout the US, dozens of police departments have already bought similar systems." For *The Police Chief*, the mouthpiece of the powerful International Association of Chiefs of Police, "predictive policing marks the beginning of a new era."

On the Up and Up: Predictive Justice

Predictive technology truly is the tool for all occasions, even spreading to the justice system. The judiciary quite rightly asks the following questions: Will this inmate, who is awaiting release, commit a serious crime in the years ahead? Will this first-time offender turn out to be a repeat offender? Under normal circumstances, judges only have their experience or intuition to fall back on. But what if predictive software could help them make the right choices about detainees? And guide their decisions about who should be given a suspended sentence or parole without any risk to society?

Is it possible, in short, to devise an objective system that would increase the number of detainees given parole and reduce the number of repeat offenders? To select which inmates should be sent on reintegration programs during or after their sentences? To decide (in a common law system) on who should remain in prison, and for how long? And to identify easily-influenced young prisoners who should under no circumstances be detained alongside hardened criminals?

In the US, four-fifths of the bodies that review parole releases now use predictive software. What data do they use for assessing inmates? Age, sex, background, and education, together with the date of the first arrest, prior offenses, and the inmate's behavior in prison. The criminal record of the detainee's closest friends and the results of psychological tests are also included, and even whether his or her mother drank to excess while pregnant. All this information is then compared with similar profiles.
The fact that decisions are taken on the basis of this type of calculation may seem worrying. However, as a criminologist is duty bound to point out, it is a little less disturbing if you have read "Extraneous Factoring in Judicial Decisions" (Princeton, February 2011). This frightening report, which compares hundreds of criminal cases, demonstrated the following: judges in the US tend to hand down more severe sentences before lunchtime, when they are hungry, than a couple of hours later, when they have been fed and watered.

Whether it is the vagaries of computer software or the ups and downs of a judge's digestive system, it is always a lottery. What is constant is this: everything to do with so-called policing or justice is based solely on retrospective information. We are told that "predictive models analyze *historical* data", and that "we collect all this data on past events"—and it is here that the conceptual problems begin.

5. How does Predictive Software Actually Work?

Considering the price—$73,000 to buy the predictive policing software plus an annual subscription of over $45,000—it is important to understand how everything works. But once again, the articles published on the subject all give the same explanations, starting with entering "all the data on crimes committed from the year X together with their location." Or: "We analyze the records of around 1.5 million crimes committed from 2003 to 2012." Why? Because "future offenses will often be

committed at the scene of an earlier crime." As a criminologist, this makes me sit up and take notice.

And so to the algorithms. The docile journalists or ad execs who write articles on predictive policing parrot the same story: "The algorithms understand (i.e. 'assimilate') criminal patterns and generate a prediction." The criminologist in me sits up again.

All this proves (claim the predictive devotees) that prediction is a serious business, since these famous "algorithms are based on earthquake prediction models" or (an alternative version) "on models that predict aftershocks." This so-called "proof" of the software's effectiveness is repeated time and again, article after article. And now, as a criminologist, I am falling off my chair because I know (and will demonstrate below) that at present it is absolutely impossible to predict earthquakes, something that none of the aforementioned fairy-tale writers took the trouble to check.

Of course, none of this has stopped some South American countries from embarking on the predictive odyssey. Engineers in Chile, we learn, have been "combining criminology and mathematical modeling", and have begun predicting the hot spots along the country's land border (which, at 6,170 km, is enormous) where crime and illegal migration will occur. In Brazil, the city of São Paulo acquired Microsoft's Detecta system in spring 2014, which helps fight crime by "aggregating data" and "creating automatic associations." The police are kitted out with laptops, tablets, and smartphones so they can access the system and organize preventive-based patrols.

But why should preventive tactics be confined to the internet? Patrolling the social networks may also be a way of predicting crime. How? The Predictive Technology Lab at the University of Virginia claims that Twitter can "predict" certain types of offense. In the March 2014 edition of the science journal Decision Support Systems, the laboratories' research team explains that geo-located Tweets (whose location can be clearly identified) are capable of predicting between 19 and 25 types of offense, including harassment, robberies and assaults. How can that be possible? "If enough Twitter users says they want to get drunk in the same neighborhood, we can 'predict' the alcohol-related offenses", say the authors. These Twitter analyses are then compared with "historically high concentrations of criminal acts." Does this type of operation have anything whatsoever to do with prediction? Not at all, as we will see below.

In summary, predictive analysis operates as follows: data mining (looking for data on the internet that is often hidden) plus statistics plus sophisticated algorithms and special software (mining tools) results in modeling and (it is claimed) predictions.

- Is the Software for Real?

Answering this question properly means providing the reader with some relevant background information:

- First, and most importantly: today, as for the foreseeable future, uncertainty, entropy, randomness, and chaos (in the scientific sense) belong to the realm of the unforeseeable and unpredictable. In other words, anything that at any given time may be possible, will not necessarily take place. Or, as Donald Rumsfeld, US defense minister during the Iraq war, once famously reported: "There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns; the ones we do not know we do not know."

To claim that algorithmic observation can be used to model or standardize unknown unknowns would clearly be intellectual fraud because what is retrospective can never be predictive. Otherwise, having a list of all the previous lottery draws would mean you would win the next one. If all the data entered into an algorithmic cruncher was derived from past events, the result would be nothing more than a probability, an often clumsy probability, a rough and ready extrapolation: criminals were in such-and-such a spot yesterday at a certain time, so they will be in the same place tomorrow.

- Since the first computers were invented, scientists have been obsessed by the idea of prediction. At the outbreak of World War II, Norbert Wiener, the father of cybernetics, tried to design a model that would predict the movements of German fighter planes so it would be easier to shoot them down. He failed in his attempts, it is said, because there was insufficient computing capacity.

- Over the last 30 years and more, various military research institutes—including DARPA in the US—have spent millions of dollars on discovering how to aggregate and combine masses of seemingly disconnected and disparate data. Their aim has been to uncover correlations so they can carry out analyses or make predictions on, for instance, future riots or attacks.

DARPA launched the Data to Decisions program in 2010 with a budget of $92 million. The project was designed to develop an algorithm to connect, exploit, interpret, and anticipate events using the mass of information stored, sold, and exchanged on the internet—and always motived by the same old dream: to predict social unrest, terrorist attacks and strategically significant events. But, judging by the turmoil US foreign policy is experiencing from Afghanistan to Iraq, DARPA does not yet seem to have found the predictive equivalent of the philosopher's stone.

There's Nothing Less Neutral than an Algorithm…

Computers and computer science are not neutral. The tools that collect and analyze data are not neutral either. Far from being the gold standard, algorithms reflect the bias of the people who devise them, fallible human beings, in other words, who might infuse their work with wishful thinking rather than pure science.

These algorithms, these tools for assembling data on the world around us that the gullible envelop in quasi-theological adoration, may have been rigged by their designers (for their own benefit), or by mischief-making or hired hackers. Does this give us a glimpse of what happens to predictive policing when it is "worked over" by hacktivists lying in wait on the dark web? Will police patrols be sent to locations where there is nothing happening while, on the other side of the same neighborhood, burglars are going about their business undisturbed? It would be harsh to belabor this point.

More broadly, it is difficult to verify whether algorithms really fulfill their mission: not just because they are capable of influencing and pre-formatting reality but also due to their sheer volume and power. If the algorithms are used on a sufficiently large scale, they generate their own validity and exert a "flocking" effect on the material facts. This is something the media should know all about, since there have been numerous recent cases of rip-off algorithms:

- In the 1970s, the Black–Scholes Model was said to be able to predict the future value of shares. But in 1998, in spite of the apparently awesome algorithms, the Long Term Capital Management hedge fund collapsed, leaving the global credit market staring into the abyss.
- In 2001, a rigged model—based, you have guessed, on the most fascinating algorithms—enabled Enron to assign an astronomical value to vanishing assets. The company then buckled and its directors were put away for a lengthy period.
- During the subprime crisis, it was discovered that rating agencies were "adapting" their software (based—no surprises here—on esoteric algorithms) to the desired outcome.
- After the aforementioned economic crisis, JP Morgan was obliged to "apologize" for using "unsuitable" software, based on advanced algorithms, that resulted in the banking giant losing $6 billion.

In 2008, three major hedge funds suffered huge losses due to "unpredictable market movements"—movements that the magic algorithms were supposed to predict. Algorithms can also lead to downright juicy scams. Take the 2005 case of the Texan businessman calling himself a "former military intelligence officer" and university professor. This individual claimed to have invented an algorithm that could make a fortune on the foreign currency markets. He swindled $33 million from his unsuspecting clients before being put away for 20 years.

But let us get back to the heart of the matter. Can predictive policing really be taken seriously?

First off, let us try to solve the earthquake prediction affair. Here is what geophysicist Bill Ellsworth, a researcher at the US Geological Survey, had to say in a recent interview: "No one knows how to predict earthquakes … If earthquakes are predictable, we do not yet know how to do it. And there is a good chance that they are not predictable … We do not even really know the equations that govern the way that earthquakes work…" As the interviewer then observes: "A scientist like Bill Ellsworth, even with all the data at his fingertips and all the computer modeling *money can buy* (our italics), admits that predicting earthquakes to some degree … escapes him."

In similar fashion, the criminal milieu clearly forms part of the real world, where complex human interactions cannot always be understood even by the most sophisticated models, and where statistical analysis tools tend to produce meaningless results. In addition, we cannot rely on laboratory experiments for a causal analysis.

Man, Machines, and Crime

Let us start with an observation, one that is commonplace for any criminologist but that eludes IT professionals and big data devotees. Human activity is not a natural resource that can simply be extracted from the cloud and monetized. It is not like coal or oil that can be mined at will. Because—and this has been the reality for a thousand years!—human-beings, regardless of whether they are criminals or not, never let themselves be watched without reacting. Our reptile-like brains and our genes are stamped with the imprint of millions of years spent avoiding predators in the fight for survival. And these predators were armed with lethal claws when we, mankind, had no such threatening appendages: no horns, no hooves, no armor. Our sole weapon was our huge brain: we adapted, we hid, we used trickery, we cheated. In short, mankind *reacts*.

Mankind is constantly trying to resist being observed and watched by computers. We are active. So, exploiting big human data is not some mundane mining activity like extracting coal or oil; it is a game of chess or, if you like, a boxing match. When it comes to illegal activities, here is an eloquent example of this phenomenon. How do drivers react in France when the authorities increase the number of radars lining the roads and motorways? They adapt by cheating. The practice of cloning vehicle registration plates (which get caught by the radars) has gone through the roof: in 2010 this type of violation jumped by 98%, and in 2011 it was up by 73%. From 2010 to 2012 the number of clonings rocketed from 5,079 to 17,479.

The Usefulness of Big Data and its Prediction Capabilities

Predictive policing is clearly unrealistic in its current form, being neither more nor less effective than on-line dating (and they both make use of similar algorithms). Nevertheless, there is a way forward for predictive technology:

- Data mining tools can be useful in the vast field of known knowns: nothing is better equipped to sort, classify and order information than a computer—in other words, computers excel at high-speed optimizing.
- Specialized software can improve medical diagnoses and therapeutic effectiveness; ditto for targeted advertising and estimating insurance premiums. And, if computer scientists and criminologists could work together, there is no doubt they could generate some extremely valuable preventive tools for the police.
- Exploring big data would also make it possible to foresee new trends, uncover emerging tendencies and glimpse hidden dynamics, all of which would obviously be very useful when it comes to security.

But, while waiting for these untold technological breakthroughs, it would be risky to take things any further.