BIG DATA

IN THIS ISSUE'S SPECIAL REPORT WE EXPLORE:

- THE NEW TECHNOLOGIES CAPTURING A GLUT OF USEFUL INFORMATION
- HOW REAL-TIME ANALYSIS IS BEING APPLIED TO RISK EVENTS
- WHAT INSURERS NEED TO THINK ABOUT WHEN IMPLEMENTING BIG DATA PROJECTS

WORDS BY EDWARD MURRAY

ILLUSTRATION BY STUDIO TONNE

In 2008, the number of devices connected to the internet surpassed the 6.7 billion people living on the planet. Depending on who you believe there could be as many as one trillion devices connected to the internet by next year.

The information recorded by these devices – the Internet of Things – creates a titanic treasure trove of big data for insurers and they're pushing hard to analyse it effectively and apply the findings to every aspect of their businesses.

As every aspect of life and business becomes digitised, so too does our ability to understand how individual events relate to each other and this will change the way we live.

Peter Hartwell, Senior Researcher at HP Labs, put it very powerfully when he said: "With a trillion sensors embedded in the environment – all connected by computing systems, software, and services – it will be possible to hear the heartbeat of the Earth, impacting human interaction with the globe as profoundly as the internet has revolutionised communication."

CAPITALISE ON WHAT YOU'VE GOT

Many insurers are already focused on extracting fresh insights from the vast amounts of data they currently hold, as the industry rushes to catch up with the rest of the world.

"Their history has been to exploit their internal databases. They are maturing in their understanding of the old data they have and looking for value," says George Marcotte, Managing Director, Financial Services Analytics for UK and Ireland at Accenture.

However, insurers seeking to put their finger on the pulse of the Earth's heartbeat and get an insight into the direct and indirect correlation between world events and commercial losses will probably need to make significant investments in data collection, generation, conversion and storage to fully unlock this treasure.

SYNC WITH OTHER SOURCES

In the area of risk management alone, IT company Celent estimates that financial firms will spend US$470m on big data projects this year, and predicts this will rise to US$730m by 2016.

These projects fall into four main areas: risk assessment and measurement, risk control and monitoring, risk reporting and governance and front office and risk operations. Drilling down further, they cover everything from risk modelling and scenario analysis, to fraud detection, risk-based pricing and regulatory reporting.

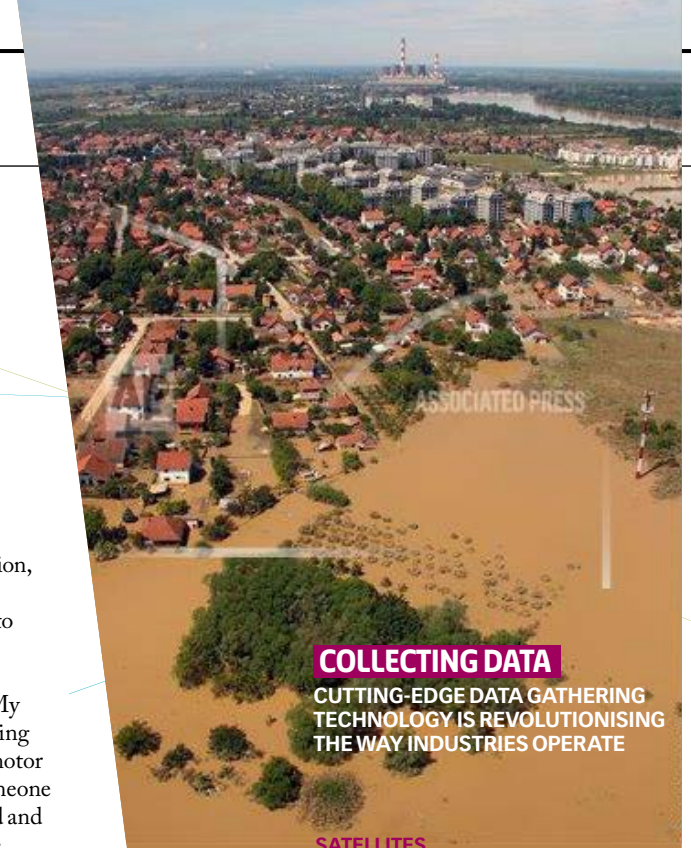
Fraud detection is a big driver when it comes to big data investment in the insurance industry. Through pan-industry initiatives such as the Claims and Underwriting Exchange and the My Licence project, the Driver and Vehicle Licencing Agency (DVLA) will open up its database to motor insurers so that all information given when someone applies for cover can be instantly cross-checked and forms pre-populated with accurate, up-to-date details of car and driver.

Joining up multiple sources of data can also have significant benefits when handling claims, as was seen in the recent prolonged severe flooding in parts of the UK. Insurers were able to use digital terrain data from Ordnance Survey and Google to predict how inland and coastal flooding might develop and monitor its spread, alerting claims teams and loss adjusters as well as spotting potentially fraudulent claims that fell outside the affected areas.

DO IT THE WAL-MART WAY

One thing underwriters may find difficult to come to terms with though – say Kenneth Cukier and Victor Mayer-Schönberger, authors of *Big Data* – is the lack of immediate connection between different databases: "Society will need to shed some of its obsessions for causality in exchange for simple correlations: not knowing why, but only what."

A simple example of the breakthroughs that can be made once such inhibitions are abandoned comes from Wal-Mart. They reviewed years' worth of till receipts, matched that data against a wide range of other databases, and noticed surprising correlations between hurricane warnings and the types of food people purchased. It didn't matter that they couldn't explain why that happened – they had extracted vital sales intelligence and now change the stock in the shops as soon as hurricane warnings are issued.



COLLECTING DATA

CUTTING-EDGE DATA GATHERING TECHNOLOGY IS REVOLUTIONISING THE WAY INDUSTRIES OPERATE

SATELLITES

A satellite's ability to gather serious detail from great distances is exciting and frightening at the same time. Currently the highest detailed satellite data available has a magnitude of 50cm to a pixel, which means cars, and houses are easily identifiable.

In April this year, the European Space Agency's Copernicus Earth Observation project fleet of Sentinel-1A satellites took to the skies. It's aim is to collect data for the earth's big problems, including natural and manmade disasters that cost the European economy €15 billion a year. Offering global coverage day and night to a ground resolution of 5 x 20 metres, the fleet has already provided imagery of the recent Bosnia-Herzegovina floods. In combination with geospatial insight, the data and imagery gathered helped government and relief agencies get a clear picture of the situation and create damage assessment maps, allowing insurers to estimate flood damage and assess exposure.

DRONES

With no need for manpower, their ability to cover long distances at high altitudes, and to stay airborne for 30-plus hours, the data gathering capabilities of drones makes the business of surveillance a lot less risky and more economically viable.

The Australian government is a recent convert. Its new fleet of Triton drones can cruise at 20,000 metres, sweeping a distance greater than Sydney to London with a 360-degree radar. A costly investment maybe – however, more than 80% of China's oil is imported through the Indian Ocean on Australia's West Coast. Japan, India and South Korea also depend on this route. And with Somali piracy costing $5.3-$5.5bn in 2011, with $635m in insurance claims, the price tag suddenly gets put into context.

Beyond security, commercial drones are being used to gather data from aerial surveys of crops to improve yields and farming methods, remotely inspect oil pipelines, and provide detailed data to create 3D mapping.

Beyond security, commercial drones are being used to gather data from aerial surveys of crops to improve yields, and to remotely inspect oil pipelines

REMOTE SENSORS

Networks of physical objects that contain embedded tech capable of communicating with the outside environment will provide real time data on everything from cars to medical implants.

Airliner easyJet has said it will install a prognostic tool in all of its aircraft, which can feed back live updates to operations and engineering staff in real time – enabling ground staff to begin investigating an issue while a plane is still in the air and ensure the right resources or parts are available when the flight lands. These measures should improve efficiency and operational performance, and reduce delays – and generate valuable information on effective risk management for aircraft fleets the world over, allowing insurers to refine the covers they write.

Elsewhere, sensors embedded in refrigeration units for transport monitor their internal temperature and send alerts when it rises. This lets operators take action and prevents cargoes, including everything from seafood to medical vaccines, from being written off.

DEPLOYING DATA

SOME OF THE BIGGEST WINS FOR INSURERS WILL COME FROM HARNESSING THE POWER OF BIG DATA AND APPLYING IT TO NATURAL OR MAN-MADE EVENTS THAT CAUSE CRIPPLING LOSSES

GOOGLE CRISIS MAP

WHAT IS IT: **First used in 2012, Google Crisis Map layered data about the size, shape, speed and direction of Superstorm Sandy onto an interactive map. It also overlaid traffic data, evacuation routes, flood zones and shelter locations, and updated the map in real time.**

HOW IS IT BEING USED: **Google has launched a similar initiative for the 2014 US wildfire season. And recently, it has partnered with the Colombian Institute of Hydrology, Meteorology and Environmental Studies and the National Unit for Disaster Risk Management to provide people with access to useful information before, during and after a natural disaster, such as a tropical storm or earthquake.**

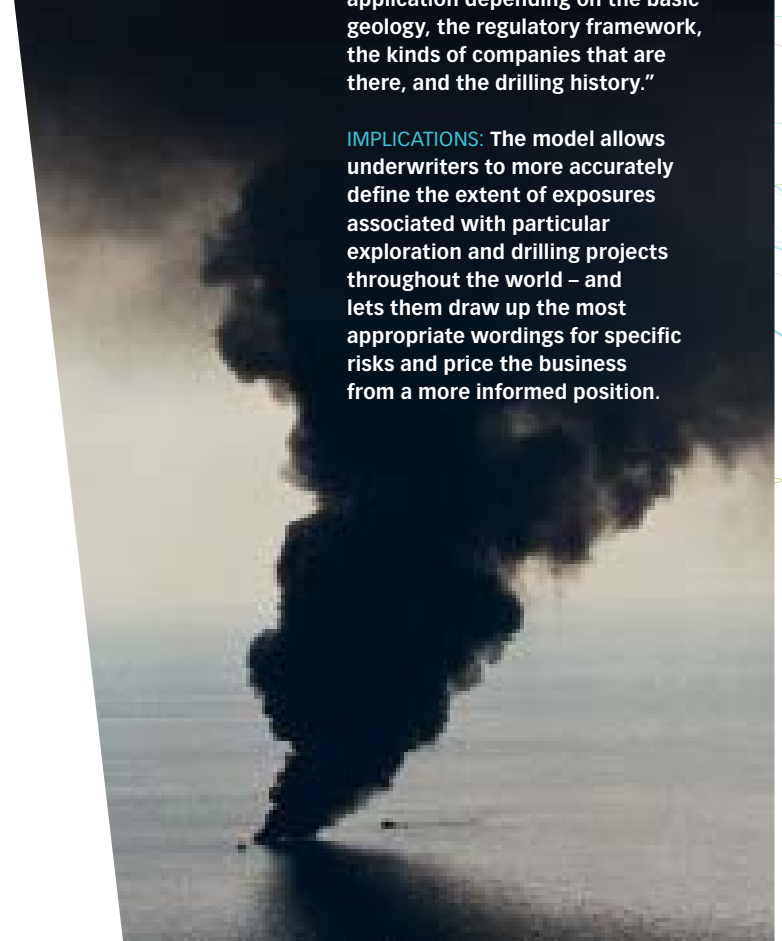
IMPLICATIONS: **The data comes into its own when responding to property claims after an event, especially when linked to satellite images of storm and flood damage. Insurers can get a very good idea of the scale of the damage, the losses and the on-the-ground response required.**

WILLIS OIL SPILL MODEL

WHAT IS IT: **A big data model to help clients understand potential losses from oil drilling operations.**

HOW IS IT BEING USED: **After the Gulf of Mexico oil spill, Willis worked with the State of Florida to help it understand if it should permit drilling off its coast. It began to build an oil spill model to estimate the frequency, severity and cost of oil spills. Gathering data from sources such as the US Government and the US Geological Survey, the broker now has a tool that can be used for projects around the world. Phil Ellis, Chief Executive Officer of Willis Global Solutions, comments: "We can adjust its application depending on the basic geology, the regulatory framework, the kinds of companies that are there, and the drilling history."**

IMPLICATIONS: **The model allows underwriters to more accurately define the extent of exposures associated with particular exploration and drilling projects throughout the world – and lets them draw up the most appropriate wordings for specific risks and price the business from a more informed position.**



PREDPOL

WHAT IS IT: **PredPol helps to put police officers in the right place at the right time, giving them the best chance of preventing crime. The technology is broadly similar to that used to predict the aftershocks following a major earthquake. Once a crime has been committed, it attempts to anticipate where the next one will happen.**

HOW IS IT BEING USED: **Deployed by the Los Angeles Police Department, PredPol has analysed data on 13 million arrests dating back over 80 years. In the Foothill Division where it was first implemented, crimes were down 13% in the four months following the rollout – compared to an increase of 0.4% in the rest of the city.**

IMPLICATIONS: **A reduction in crime has implications in many different lines of business, from property and motor to medical and travel. But its real potential would be if insurers can persuade police authorities to share the data so they can overlay it on their existing underwriting information to get a more accurate picture of theft and fraud risks in an area.**

MODELLING DATA

FROM TERROR ATTACKS TO PANDEMICS, BIG DATA IS HELPING TO REFINE THREAT MODELS

THREAT #1

PANDEMICS

Transmissibility and virulence are the two key factors in modelling a pandemic, and big data is providing new insights into both of these.

In regard to transmissibility, it is difficult to locate the first person that a pathogen has infected and then track every other person they have passed it on to. Without an accurate understanding of the total number of people a pathogen has infected, it is almost impossible to effectively estimate just how deadly it is.

However, big data is making it easier to tackle these problems. "What we're hoping," says Dr Andrew Coburn, Senior Vice President at RMS, "is that with big data we will get a much better idea of infection rates. One experiment that has been very interesting is Google Flu, which essentially tracks search terms. People who feel ill type in 'flu symptoms', or such like, and they've found that the wave of search terms is correlates highly with the outbreak of flu."

The earlier insurers can get an indication of a developing pandemic, the faster they can provide advice to insureds around everything from the drugs that hospitals will need to order and the resources local authorities will require, to what businesses and individuals can do to avoid infection.

In addition, Coburn says the focus for pandemic modelling has moved beyond the losses that insurers would bear from the life market: "What models are increasingly focused on is the business interruption and macro-economic impact that a pandemic can have. They are looking at things such as companies going out of business, trade credit insurance, contingent business interruption losses and all of that secondary activity that could be fairly crippling. A pandemic would also have an impact on an insurer's investment portfolio on the back of the downturn in the stock market that would occur. That is perhaps more significant in terms of balance sheet impact than the claims themselves."

Again, big data has an important part to play in modelling the changes that a pandemic would create in investment markets, and being able to monitor and understand the way investment decisions are made, when people are massing into crowded trades and where asset bubbles are starting to rise.

Forewarned is forearmed, and as big data gets better at delivering these insights, insurers and the policyholders they cover will get better at understanding where losses might occur and avoiding them.

THREAT #2

TERRORIST ATTACKS

Terrorist attacks present a number of major problems for insurers, particularly in regard to probabilistic modelling and risk accumulation. The good news is that terrorist risk modelling has become a lot more sophisticated in the last decade, and big data has played a significant part in this evolution.

A recent white paper from RMS reveals that its models for terror attacks are plugged into simulations of over 90,000 possible large-scale attacks against 9,800 different targets worldwide, with 35 potential methods of attack – ranging from improvised explosive devices to biological and nuclear attacks.

Explaining how the RMS models work, Chris Folkman, Director of Model Project Management at RMS, says they rely on understanding where terrorist plots take place; how often they take place; how successful they are and what the counter terrorism capabilities are within a given region.

"That basically gets us to the number of successful attacks that can be expected in a particular place each year and we use that as the starting point for the probability of an attack."

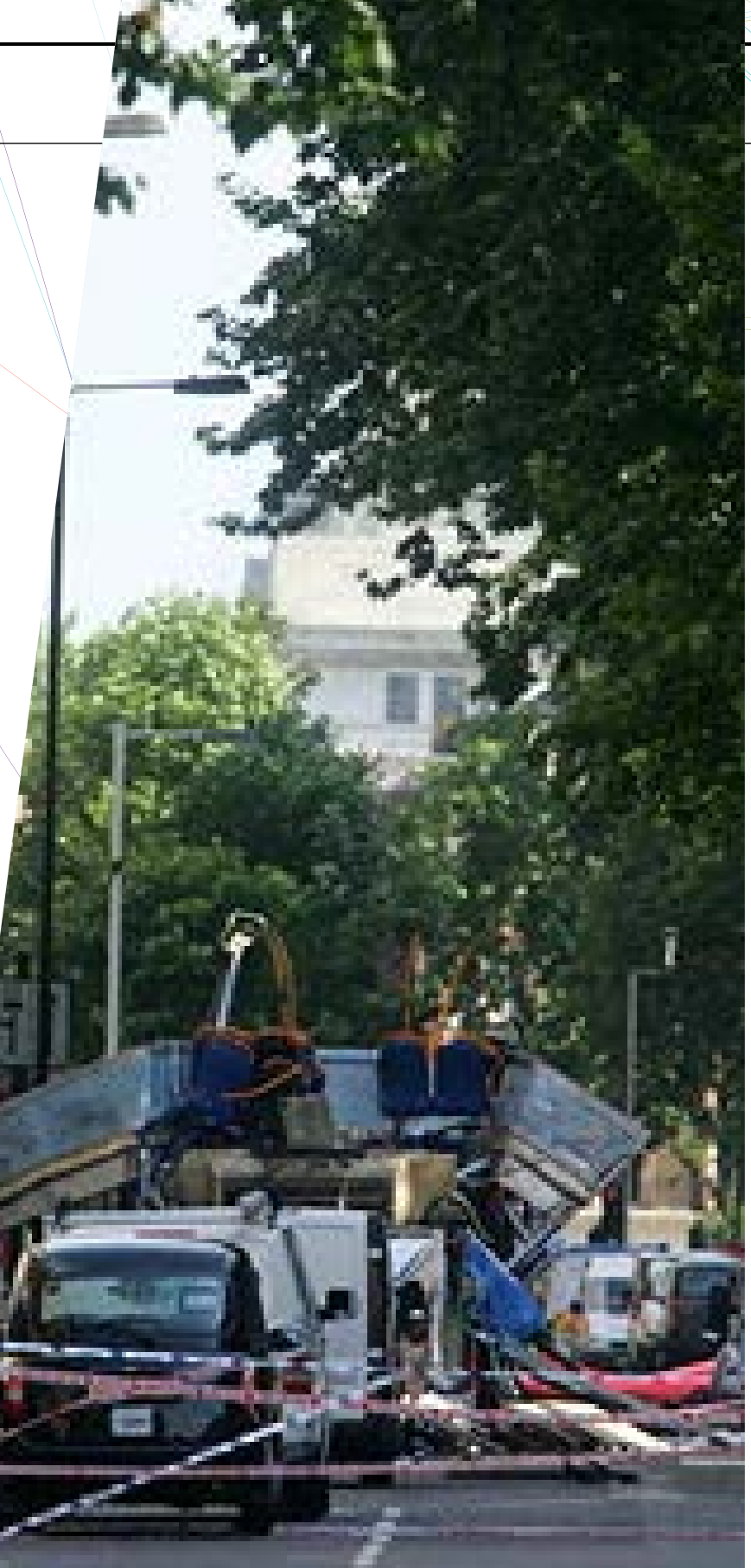
Looking beyond probability calculations, the other major problem for insurers is the localised nature of most terrorism events. The World Trade Centre was situated on a 16-acre site and yet the losses ran to tens of billions of dollars. "That means two things," says Folkman. "First, it is very difficult to diversify because the risk is concentrated and it is often in urban areas where there is a lot of insurable exposure. Second, it is important that the data that goes into the model in terms of the location is accurate – preferably to building level."

The RMS models can calculate exposure in relation to not only the property risks, but also those associated with lines of business – including life insurance, workers' compensation and personal accident.

For insurers, this sort of modelling brings a number of benefits, says Folkman: "The models helps insurers find hotspots of exposure. They can use them to quantify a worst-case scenario attack at a location of interest. And they can use them to create underwriting rules, and at a corporate level to create risk appetite guidelines."

The models help insurers find hotspots of exposure. They can use them to quantify a worst-case scenario attack at a location of interest

CHRIS FOLKMAN, DIRECTOR OF MODEL PROJECT MANAGEMENT, RMS



TALKING DATA

THE SCOPE FOR USING BIG DATA TO UNLOCK NEW BUSINESS IS HUGE. BUT FIRST, YOU HAVE TO MANAGE THE IMPLEMENTATION. WE TALK TO TONY BOOBIER, EUROPEAN INDUSTRY LEADER–INSURANCE, IBM

MARKET: **How influential is Big Data in the world of insurance?**

TONY BOOBIER: There are a lot of projects wrapped up under the big data banner. Whether insurers are looking at risk, customer analytics, or operational analytics, the big data agenda is all pervasive. To date, the applications have been limited predominantly to the personal lines market. One area that has experienced significant change is user-based insurance in the motor sector. But we're now seeing big data pervade in the commercial lines market too, and many of its specialised areas. Data is coming from space satellites, drones and devices embedded in everything from cars to cargo ships. And all of this information is being used to generate insight into the risks insurers underwrite.

MARKET: **How should insurers implement big data projects?**

TB: We think of implementation in terms of three core components. The first is the challenge presented by the data and how you get a solid information infrastructure. Then there are the skills and software solutions you need, and how you embed these tools into your organisation. Third is the issue of culture. Businesses that adopt a big data and analytics agenda need to have a fact-driven leadership. The board does not need to be technologically savvy, but it does need to provide leadership to ensure a data-driven culture.

MARKET: **How important is data accuracy?**

TB: There's a misconception that every piece of data needs to be irreproachable to be able to extract useful insight. We recommend that firms first identify the outcomes they're looking for and then work backwards to prioritise what data cleansing and organisation is required. If you do it the other way around and try and get all of the data spot on, it can feel like you're trying to tame a tidal wave.

In an environment where so much data is unstructured, you're never going to get it perfect. We tend to look at the world through UK or European glasses. But as insurers focuses on growth markets like China or Latin America, the quality of information we've traditionally had is not quite as good. There needs to be some pragmatism.