AQ
Analytics Quarterly
Spring 2014

IN THIS QUARTER'S ISSUE: In the beginning was the word • It is on its way • Chips with Everything • Steak-holders! • Health and Efficiency • Sale of the Century • Starting your own Start-Up • A Spoonful of Data • All You Need is Data! • You want it WHEN? • Buy, Buy, Buy • We know where you are going • Predictive Analytics and BMW • Privacy and big data – good or bad? •

THE OR SOCIETY

the Analytics Network
a section of the OR Society

In the beginning was the word

Nigel Cummings

Extracting meaningful information from data held as free-form text has posed a major problem for any form of analysis.

Around 80% of the information we see before us, is dispersed in free-form which, prior to the development of text mining, had to be read by humans – a slow, labour-intensive and not always accurate process. A great deal of research has been expended on machine or artificial intelligence – indeed it was to computing what cold fusion was to physics for most of the second half of the 20th Century.

The University of California, Berkeley, for example, began in the 1990s to apply machine intelligence to untangle the complexities of data mining from text. Marty Hearst at the School of Information Management and Systems published a paper in 1999 entitled 'Untangling Text Data Mining'. The intent at that time was to draw contrasts and attention to what was seen as an exciting new kind of problem for computational linguists.

Text mining can be defined as the analysis of semi-structured or unstructured text data. There are several areas of speciality in text mining which are generally acknowledged, all of them have different objectives as their definers. These include: information retrieval; document clustering; document classification; information extraction and; concept extraction.

For the programmers amongst us, text mining may be a matter of a few hundred lines of programming and access to libraries of data mining algorithms. For those of us not so equipped, text mining can be demonstrated and utilised quite simply by accessing any web browser and using search engines to obtain information. For more significant text mining needs however, there is no shortage of vendors. Companies ready to supply text mining applications commercially, include: IBM, Statistica, SAS, and Lexalytics, to name but a few.

A book I found of interest on this subject is: **The Text Mining Handbook: Advanced Approaches in Analysing Unstructured Data**. Authors are: Ronen Feldman and James Sanger . Published by Cambridge University Press in August 2013. ISBN: 9780521836579. I would recommend this book to anyone who is interested in the field of unstructured data mining.

It is on its way

Nigel Cummings

A new approach to the application of data analytics can help retailers improve customer delivery.

A major e-grocer in the United Kingdom increased its profits by adopting a 'foresight approach' to customer home delivery. New research into this 'foresight approach' by scientists at Warwick Business School, Lancaster University Management School and the University of Southampton has resulted in a new way of applying analytics to the problems associated with home delivery services.

The foresight approach to home delivery predicts when people want their shopping delivered based on what delivery prices or incentives are being quoted for different delivery time slots. The advent of cloud computing which provides cheap, powerful analytics support is crucial in gaining efficiencies when mining and analysing customer shopping retail behaviour habits. The increase in online shopping has also been spurred on by the growth and adoption of smart phones and tablets.

According to Dr Arne Strauss, assistant professor of Operational Research at Warwick Business School. 'It is important to incentivise customers and steer them to particular delivery times. Maximising profits is a problem because the final set of orders for a given delivery day are not known until shortly beforehand, yet decisions on the pricing of delivery time 'slots' have to be made in advance based on an estimate. With our new approach we demonstrate that analysing the customer data already at retailers' fingertips and using it to predict the impact of future expected orders in the estimation of delivery costs produces higher profits than only using orders accepted to date in this estimation.'

Dr Arne Strauss can be contacted by email at: Arne.Strauss@wbs.ac.uk

AQ

Chips with Everything

Nigel Cummings

Five years ago, if a plant manager at McCain wanted to know how they were doing compared to the company's other 50 production facilities around the globe, they would have had to do what most manufacturing plant managers still do: put in a request to IT and wait.

Today, that same manager only needs to walk onto the production floor and look up at a giant LED screen. On that screen he/she can see exactly how well production is going based on meters that compare their production lines to every other line worldwide. This has been made possible due to the application of analytics.

According to McCain's CIO Roman Coba, the McCain analytics journey started around five years ago when the company decided that instead of moving through some 18,000 physically produced paper reports from access databases, which had a latency of around 24 hours, the company would be better served by applying analytics technologies to achieve real-time monitoring of its global output.

Recently however this aspect of analytics has been seen as an opportunity to give employees decision power no matter where they are or what they do. According to Coba, 'Now that the genie is out of the bottle people only want one thing: More.' Since introducing analytics into the company, production has gone up by 'double digits' and down time has been cut by an equal percentage.

Worker injuries have also been drastically reduced by giving safety managers the ability to access data not only from their own plants but from operations worldwide. This has seen their total incident rate change from a 7 rating (which is very bad) to a 1 (which is very good).

Safety is an interesting case in point where bringing together all of the analytics at McCains has really paid off. The IT organisation at Mc Cain has also benefitted. Not only is IT freed up from having to produce and manage so many paper reports, the analytics applications enhance IT's credibility within the organisation. IT is now viewed as an enabler of innovation, and fewer managers and directors bypass IT to get the tools they need to do their jobs.

AQ

Steak-holders!

Nigel Cummings

Keeping animals healthy and treating them when they are sick is a labour-intensive yet fundamental requirement of livestock farming.

Now GrowSafe Systems, a Canadian company, claims to have developed a system which utilises sensors, analysis and medications which can recognise, diagnose and treat unhealthy cows automatically (adding medication to their drinking water). The same system can also identify when the animals are ready for the market.

The result is healthier animals and lower costs, risks, and environmental impact for their owners. Over the past five years, the Samuel Roberts Noble Foundation and GrowSafe Systems have collaborated to test GrowSafe's analytics based approach to better understand feed efficiency in pasture-fed beef cattle and develop tools for sustainably improving production of grazing livestock.

The animal behaviour research behind it actually started out in the 1990s. It has, however, had to wait for advances in computer processing power, Internet accessibility, routers, and advantages provided by analytics and cloud computing to provided the power necessary to adopt predictive analytics with automated response systems for the farming industry.

GrowSafe recently won an 'Ingenious Award 2013' for how it collects and analyses big data. Its research and subsequent roll-out of technology now enables farmers to make better decisions about their cattle.

GrowSafe's beef solution works by strategically placing biometric sensors in the water troughs and feeding areas of the feedlot. These sensors constantly collect data on the weight of the animals, plus their movements, drinking behaviours, feeding behaviours, temperatures, and many different types of other data. The data is analysed overnight and results are displayed in a 'farmer-friendly' way in the morning identifying any cows that may need to be looked at more carefully.

The animal's radio-frequency identification tag is then used to pick out the right animal. The system also provides the farmer with the decision support necessary to decide whether the animal can be treated or has to be taken out of the food-chain.
http://www.growsafe.com/

AQ

Health and Efficiency

Nigel Cummings

A whole ecosystem of analytics solutions has emerged over the last couple of years or so which are helping to make healthcare providers more efficient.

In 2012 it was estimated by analytics-specific industry bodies that approximately 500 petabytes (PB) of digital healthcare data existed across the globe, that figure is expected to rise to 25,000 PB by the year 2020. That is a 50-fold increase in the amount of digital healthcare data compared to the 2012 estimate - a huge amount.

Having better data should lead to better healthcare outcomes; as to whether more data equates to better data is debatable. What it does allow is a move towards analytics driven 'evidence-based medicine ', which is more clearly tailored to the individual, based on their symptoms combined with medical and genetic history possibly even incorporating life-style data extracted from such unlikely sources as FaceBook and other social media. Data collected via loyalty cards can give a pretty accurate indication of what people in a given household eat and drink and possibly from what minor medical problems they suffer. Of course, there is always the possibility that there is one person in the household who eats all of the healthy food purchased and the others eat all the unhealthy stuff – no system can be 100% correct.

Data analysis can help us make some of the decisions necessary to provide efficiencies and higher levels of effective healthcare. Data, however, needs to be consistent, apolitical and compatible before it can be integrated or interchanged across and between systems and countries.

Data protection also creates problems with different laws in different countries. There is also the difficulty of achieving the required level of anonymity as data analysis systems become increasingly more sophisticated; if too much information of this type is stripped off, what remains may be of very limited value.

Avent claims it can integrate data from multiple sources to solve specific problems for clinical, quality / compliance and operational areas which involve Business Intelligence (BI) and analytics for visualisation in a unified view. The areas in which it claims it can provide insight include: readmissions management & tracking, financial performance, patient engagement, and regulatory compliance. More information can e located at: http://www.pepperweed.com/capabilities/healthcare-analytics.html

Data protection also creates problems with different laws in different countries.

IBM has invested a huge amount of time and money in the acquisition of analytics-based companies into their analytics business for healthcare services. More information on IBM's offerings can be found at: http://www-01.ibm.com/software/analytics/healthcare/

SAS too, provide a range of services relating to health analytics, the following web link will lead you to a page on their website which explains what health analytics is about from the SAS perspective, and what technologies and skills can be used to deliver business, clinical and programmatic insights into the complex interdependencies inherent in healthcare. http://www.sas.com/en_us/insights/health-analytics.html

Another company, Avenade, claims to be experts in enabling healthcare organisations to align with changing regulatory standards and shifting to predictive analytics. http://www.avanade.com/en-uk/offerings/pages/healthcare-analytics.aspx .

While McKinsey Advanced Healthcare Analytics specialise in providing local market strategies for what is considered to be one of the fastest-growing segments in the healthcare coverage landscape analytics! http://www.maha-us.com/

Sale of the Century

Nigel Cummings

Abercrombie & Fitch use predictive analytics to maximise profitability. The days of obtaining bargains in clothing sales may be numbered.

Traditionally end of season sales fulfil several purposes, they engender customer loyalty, clear slow-moving stock from the shelves and give retailers the opportunity to order more stock to furnish next seasons' requirements. With increasing focus on lean supply chains, companies are beginning to use predictive analytics to help them reduce stock levels, increase speed to market and raise profits. Leaner and keener supply chains stimulated by such drivers as these could mean the end of bargain sales as we know them.

Teen apparel retailer Abercrombie & Fitch has, for example, partnered with a new solutions provider First Insight as part of its strategy to increase speed to market. First Insights' consumer-driven predictive analytics package has helped reduce markdowns and mitigate risks associated with new product introductions. Predictive analytics now enables Abercrombie & Fitch to test new candidate products, in every product category, on a weekly basis, throughout the merchandising organisation and the product development lifecycle. This, in turn, is leading to a drop in promotions (fewer bargains for the consumer) and an increase in average unit retail prices.

Leaner and keener supply chains stimulated by such drivers as these could mean the end of bargain sales as we know them.

'By using First Insight to identify more winning products and price them appropriately, we are increasing speed to market with the right styles which will yield increases in sales and margin,' said Gillian Galner, GVP of Abercrombie & Fitch. 'After evaluating the solution for an extended period, we have confirmed the solution will deliver a strong return on investment.' More details on: http://www.firstinsight.com/

Clothing retailers are not the only companies who are using predictive analytics to give them the ability to start making smart decisions about their customer behaviour, impact on marketing and customer strategy. Chelsea Building Society for example have recently adopted Pitney Bowes' 'Portrait Foundation and Interaction Optimiser' to improve operational efficiency and deliver better customer service.

Chelsea Building Society chose analytics as a route to deliver further sales growth year on year and recognise the need to collect customer data from disparate sources, provide a consistent customer experience across all channels, identify cross sell and sell opportunities and eliminate manual processes which take time and cost money.

The software provides a 'single customer view' which includes complete product portfolio, contact and personal details, complaint information, related workflow tasks, significant events and marketing information. The system has vastly increased operational efficiency and is now live in all branches.

More details on: **www.portraitsoftware.com**

Starting your own Start-Up

Sayara Beg, Chair of the Analytics Network

This year the OR Society will be holding its 3rd Annual one day meeting on Advanced Analytics and Big Data at the BMA House in Euston, London, UK and already the speaker line up is generating intense interest – more details are given elsewhere in the Newsletter.

This year one focus of the meeting will be on Analytics as a specialist skill set that is singled out as a profession in its own right. Three of the speakers will talk about how the Analytics professional can be certified, chartered or even a CEO of their own highly successful Analytics start-up.

Alan Hambrook, Zoral Group, co-founder and Chief Executive Office, one of our speakers gives the following advice, 'It is important to keep in mind that venture capitalist firms receive hundreds of business plans and they do not have time to look at all of them let alone evaluate them, but essentially be aware that venture capitalist firms specialise and work in preferred markets.

'What is important is that before you approach a venture capitalist firm spend considerable time understanding them, be clear in your mind about why you have chosen them and what you are asking from them.'

Check them out then short list about five. Now you know who to approach but how should you approach them? With a business plan, of course!'

Hambrook recommends visiting the 'Insights' page of Highland Capital Partners {http://www.hcp.com/insights} for excellent advice, and considers them to be a top class venture capitalist firm, adding 'When you make your first approach, start with your least favourite on the short-list. That way you will get good feedback, practice, constructive criticism to hone your pitch/presentation, before you meet the more favoured.'

So, what is most important to a venture capitalist firm, you may ask?

Three things…

1. Quality of the management team and their experience
2. Quality of the management team and their experience
3. Quality of the management team and their experience.

'This must come over very quickly and succinctly in the Business Plan and PowerPoint pitch'

If you have caught a Venture Capitalist firm interest and they decide to evaluate a business plan, then they dig deep and focus on what is important to them.

Peter Bell of Highland Capital Partners says, 'On the team, I look for entrepreneurs that are passionate and have hunger and relevancy. For the size of the opportunity, I look for bold products with high-sustainable margins and significant value propositions.'

Almost every first time start-up enthusiast, falls into one of these mistake categories:

1. Selling/over-selling – the pitch to a VC firm is not a 'sale' - the VC firm is being asked to be a co-shareholder, so it is important to ask them good, intelligent questions.
2. Unsubstantiated claims. Have good, concise, reputable third party validation of any claim you make, (e.g. Market size, state of competition, etc)
3. Saying too much about the idea, not enough about the team. It is well known amongst VC firms that many can see an opportunity but very few can actually execute it with a solid team – so a VC firm will be constantly be thinking 'why does this team stand a good chance?'

So if you are thinking of starting your own start-up, my advice would be to begin networking and fostering partnerships with people who have relevant experience and expertise to help you make the right decisions. The ORS Analytics Network is the perfect place to meet like-minded individuals who could possibly help you start your own Start-Up.

Alan will be presenting at the Analytics Network Conference 'Developments in Big Data' on the 30th April 2014 on the subject of advanced behavioural and predictive analytics using unstructured, social and behavioural data.

A Spoonful of Data

Nigel Cummings

If the NHS were a human being, it would be classified as a pensioner now, for 2014 marks its 66th year of existence; 66 years of service in which, despite many criticisms levelled at it, it is still acknowledged as one of the leading health organisations in the world.

Last year there were repeated calls for the NHS to live up to the £20bn 'Nicholson Challenge' for driving efficiency and savings i.e. to find £20 billion in 'efficiency savings' by 2015.

If this challenge is to be met, it is reasonable to assume that much of the savings will come through technology. This is because the NHS is shifting from being an institution focussed on treating chronic conditions, to one that uses data-sharing, O.R., Analytics, collaboration and digital healthcare practices to deliver healthcare services predictively.

One example of the application of this technology has been the digitisation of patient records. A move which is seen as vital for accelerating patient care, cutting waiting lists, and enabling more straightforward information sharing between departments and providing timely insight for medical practitioners.

Digitisation in addition to reducing costs also has the potential to simplify the working lives of healthcare professionals by reducing the time spent on administrative tasks and paperwork. Some NHS Trusts have shown considerable initiative by building advanced systems to support paperless working practices which have proven so successful that they will eventually be rolled out into all NHS installations. Simplification of patient records systems and easier sharing of data between NHS installations is both laudable and economic, but there is another phenomenon which is having a transformative impact across the UK – big data!

The World Economic Forum recently highlighted the fact that chronic care for high blood pressure, vascular diseases, lung diseases and diabetes could account for up to 80% of healthcare budgets by 2030. Although it did emphasises that the use of big data technologies and techniques to create platforms capable of improving efficiencies in the management of chronic health conditions such as cancer, diabetes, pulmonary conditions and cardiovascular disease would help to keep costs down if they were to be adopted universally.

The technologies most likely to have an impact in efficiencies and cost benefits are predictive analytics set across multiple datasets in real-time, the delivery of stratified medical pathways, drawing on patient, environmental, social and genetic data to anticipate treatment pathways, and the correlation, analysis and interpretation of telehealth, telemetry and genomic data to pre-emptively treat disease.

Predictive modelling is not new in the NHS but the difference now is the availability of technology that allows big data to be mined quickly and rapidly displayed to a variety of stakeholders and providers along integrated care pathways and across health and social care.

Coming to a surgery near you, doctors will be able to reliably prescribe treatment for disease that isn't due to manifest itself for 20 years or more. He or she could pre-emptively place you on a prediction driven treatment pathway that keeps you fit and healthy to maintain a high quality of life for longer. O.R. and Analytics will play a big part in this.

So, instead of spending £80bn on caring for the sick, much of this money will be needed to provide the equipment and technology to identify those at risk and provide pre-emptive treatments. There is life in the old yet!

All You Need is Data!

Nigel Cummings

'Data scientists are the new rock stars' according to the likes of Olaf Swantee, CEO EE, Ken Rudin, head of analytics at Facebook and Jeff Magnusson, manager of data science platform architecture at Netflix.

These gentlemen believe the data sciences are now so popular that if you ask children what they aspire to be as adults, data scientist will be a choice mentioned in the same breath as fireman, doctor, rock star, rapper, or even astronaut. This is because these days being a data scientist is seen as being someone involved in a glamorous industry.

Just looking at the United States alone for a moment, there are almost 190,000 positions available for up and coming data scientists. Companies looking for success in the data sciences are taking in all sorts of science graduates; Bachelors, Masters and even Doctors. Stanford, North Carolina State and Northwestern universities are already experiencing a huge influx of students clamouring for degrees in data management and analysis. A similar situation exists in Europe and applications for data science placements are 'on the up' in the UK – you only have to attend one of our careers open days to see that!

A study by the Royal Academy of Engineering shows that British industry will need 1.25 million new STEM graduates between now and 2020 just to maintain current employment numbers. Even that figure might not be enough to satisfy the British data industries' requirements though.

This is because is 'big data' is as yet unquantifiable, just how big is it? A study at the end of 2012 by IDC predicted the 'digital universe' would reach 40 zettabytes (ZB) capacity by 2020, though in reality that figure could be much higher – 40 ZB is 4*1022 bytes or approx, 40*270 bytes)

This terrific surge of data is being created by many external forces which include: financial transactions, mobile phones and social media, the number of clicks that take place daily on the Internet to access information, and even from the updating and keeping of medical records.

According to IDC only 1% of the world's data is currently being analysed and the technology and tools for collecting and storing information has to date raced far ahead of our skills to understand it - data collection is outstripping our abilities to develop technologies to analyse it! Filling this 'big data gap' is not just a question of 'getting up to speed' with technologies though, filling the gap also means importing 'new talent' into data driven employment - these people will be in demand as much as the software developers of the dot-com boom were.

Eric Siegel, author of Predictive Analytics, summed up the value of data when he said. '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 $1,200 per year.'

Data scientists Reid Hoffman and Konstantin Guericke created LinkedIn, in December 2002 to help build individuals' networks for them. The 'people you may know' feature has, through its ability to target individuals for marketing purposes, raised the 'value' of the company to approximately $7.5 billion (£4.5 billion). LinkedIn is one company that shows analytics can be a route to massive revenues!

AQ

You want it WHEN?

Nigel Cummings

Ocado's vans are a familiar sight on suburban streets. But how is Ocado able to compete with such giants as Tesco, Asda and Sainsbury's, to name but a few?

The answer appears to be in its software. Ocado employs some 340 software developers and other IT specialists within its technology division - more than half of the head office headcount.

According to Paul Clarke. Director of Technology, 'We are not a typical retailer - technology is woven into everything we do. We are much more like Google or Amazon than our traditional competitors.' Ocado's main selling point is the ability to designate 'one-hour delivery windows' to customers.

Tesco, currently only offers this feature in 'some' locations, and provides instead a far less convenient 'two-hour window' option everywhere else. To honour its commitment, Ocado has developed its own route optimisation software, to direct its delivery vans to customers, accounting for predicted traffic and accommodating other drop-off points, thereby saving both fuel and time.

The company needed complete control in terms of the algorithms it uses and the optimisation processes to achieve those one hour slots. In the early days of the Ocado business, standard commercial routing software was utilised, but it was not good enough, it was not capable of meeting the demands of a large scale home delivery retail commitment. So the company recruited analysts and programmers and started to build its own routing applications.

Now, almost all of the technology that powers Ocado - from the web shop and mobile apps, middleware, the real-time control systems in its warehouses and delivery - is built internally. This has given the company complete control and ownership of the intellectual property it possesses, and this means that if there are things needed to change or fix, as a 24/7 operation, the company can usually achieve that change by 3am!

Ocado is very much a 'green' operation, this is evident to its customers too as the one hour slots are scored and colour coded to show its customers which are the optimum times for energy efficient deliveries. Less efficient delivery slots are usually flagged with a small service charge too, this helps the company offsets its carbon deficits and encourages its customers to select the 'greener' delivery options. Whilst Ocado seems to have found its solution to providing timely, greener deliveries, other retailers have yet to follow suit.

Arne Strauss, Assistant Professor of Operational Research, Warwick Business School, along with researchers from Lancaster University Management School and the University of Southampton are using a yield management approach to similarly encourage on-line shoppers to choose delivery slots. A combination of delivery prices, offers and loyalty points are used to encourage customers to pick times which will help the retailers minimise failed delivery attempts and fuel usage.

Strauss recently said: 'Traditionally online retailers would collect orders including delivery time requests until a certain cut-off time and plan their delivery schedule accordingly. Therefore, maximising profits is a problem because the final set of orders for a given delivery day are not known until shortly beforehand, yet decisions on the pricing of delivery time 'slots' have to be made in advance based on an estimate.

'Analysing the customer data to predict the impact of future expected orders produces higher profits than only using orders accepted to date in this estimation. Our model can outperform the static two-tier delivery pricing policies that are often found in practice by around 4% in profit. In an industry that operates on very small margins, this profit potential is significant.'

Buy, Buy, Buy

Nigel Cummings

If you are a LinkedIn user and you suddenly find engagement with it has become much more interesting lately, it could be due to the fact that LinkedIn now incorporates Mintigo Predictive Analytics.

The best marketing campaigns quickly get beyond basic company demographics, combining analytics with integration to social networking platforms to get to know prospects as people. LinkedIn has over 277 million members worldwide (you may be one of them) and millions more will join up every year. The sheer number of users makes it one of the most valuable social networking sites for B2B marketing today.

Mintigo, a company specialising in predictive analytics technologies has been collaborating with analysts at LinkedIn to integrate, what they call, Customer DNA technology to enable LinkedIn advertisements to be more precisely targeted to prospects that have the highest probability of buying, the goal being to avoid displaying adverts to non-responsive LinkedIn users and concentrate on those more likely to buy. The application of Mintigo technology apparently allows LinkedIn's analysts to make recommendations based on the level of engagement of users in terms of content and quality.

A quick glance at the graphic provided here will give you some idea of what the analysts at LinkedIn actually see when they examine you and your viewing habits of their online presence. The dashboard view divides members into 'warm', 'need nurturing' and those to whom cross selling and up selling can easily be facilitated.

Mintigo builds a Customer DNA profile of any company's ideal target customer. It then then searches the Web, social networks, and 3rd party databases to find contacts that fit the unique profiles generated. It can even score any leads a company like LinkedIn has already acquired, and it can supply 'net-new' contacts that are likely to buy products.

Customer DNA has also been designed to continue to learn over time by monitoring marketing and sales campaigns. 'Every time you email or call, it learns more about which campaign works best on you!'

Without being too cynical about this, I think it is easy to see from the graphic provided, that even the most acute and agile minds subscribing to what is, regarded as an 'upmarket and professional' social networking system can be manipulated by the power of analytics, and in some cases manipulated extensively! Have we all, now entered the age of Manipulytics?

AQ

The sheer number of users makes it one of the most valuable social networking sites for B2B marketing today.

We know where you are going

Nigel Cummings

Of the top 500 internet retailers, 98% use some form of web analytics. It has become an essential tool for understanding online customer behaviours and driving improvement.

Despite the success of analytics online though, relatively few companies with physical venues employ advanced analytics solutions that track customer behaviours in their physical spaces – until now that is!

Growing numbers of companies are engaging in the emerging field of location analytics. One of the reasons for take up of location analytics is the availability of data from mobile devices, ranging from smart phones to tablet PCs, in analytics terms each of these devices can be seen as a 'honey pot' of commercially usable data just waiting to be raided by retailers.

Potential customers' tracking data is typically sent to the location analytics vendor from mobile devices and the data is analysed and accessed via online dashboards that provide actionable data tailored to the needs of specific employees. The scale of data collected from such sources is staggering.

Location analytics firm RetailNext for example currently tracks more than 500 million shoppers per year by collecting data from more than 65,000 sensors installed in thousands of retail stores. A single customer visit alone can result in over 10,000 unique data points, not including the data gathered at the point of sale.

Another analytics reliant company, Euclid, collects six billion customer measurements each day across thousands of locations, and many location analytics firms say they are adding hundreds of new venues each month. Location analytics firms are even pushing beyond company-owned venues. Locately and Placed for example, routinely use opt-in apps on customer devices to track everywhere customers go.

The format of such analytics ventures is rather similar to conventional web analytics, it can be categorised similarly into four distinct sub divisions:

Design. Removing obstructions, changing marketing messages and other more subtle methods can be used to encourage customers to visit certain areas.

Marketing. One well known restaurant chain wanted to understand the pros and cons of sponsoring local music festivals, would it have a measurable effect or no effect at all on customer visits? Analysis of data collected before, during and after a festival revealed that the festivals resulted in over 1,300 new customer visits.

Operations. One large grocer used location analytics to understand customer wait times in various departments and check-out tills. Data analysis is used to identify staffing needs for each department throughout the day and optimal times to perform disruptive tasks such as restocking shelves and resetting displays.

Strategy. Data analysis showed that only 2% of main store customers were attracted to a newly opened retail park outlet store indicating that the outlets gave them access to an entirely new customer base with minimal impact to existing store sales.

It looks like location analytics will become the 'must-have' for offline retailers wishing to maximise sales. The ability to identify, track, and target customers in physical locations will enable companies to extend preferential status and rewards to customers based on their behaviours, rewarding them based on the number and frequency of visits, where they go in venues, and their 'exclusive' loyalty.

Predictive Analytics and BMW

Nigel Cummings

Car manufacturer BMW is using big data and analytics technology to help detect and fix vulnerabilities in their cars before new models go into full-scale production.

BMW is using IBM analytics technology to foresee early stage production problems long before they might cause problems in series production. The analyses apparently include product and development data as well as warranty, diagnostics and repair information and the sources of data used are truly immense – this BMW analytics project derives input from dealer and service facilities worldwide.

BMW is using powerful analytical installations to obtain insights from the large amounts of data generated from the product design and production process. Analyses that used to take several months prior to the adoption of analytics and cloud computing are available within a few days now, so that potential issues are detected and fixed swiftly to avoid recurring faults.

The IBM SPSS predictive analytics software used can combine and analyse data from test drives of prototypes too. The results of the analyses are immediately directed back into BMW's operational processes – this helps reduce error rates and creates cost benefits.

The application of big data and analytics technology has also given BMW the ability to discover patterns and anomalies which help in the prediction and anticipation of maintenance. This allows for targeted evaluation of product, maintenance and repair data, and for the issue of repair instructions on a timely basis.

Another benefit is the automation of some analyses, and due to the fact that different BMW business divisions and subsidiaries often have similar analytics queries, this information can be quickly shared across the enterprise to best effect.

The market for data and analytics is estimated at $187 billion worldwide by 2015,and two-thirds of IBM Research's work is now devoted to data, analytics and cognitive computing in response to this, and IBM has filed 4,000 analytics patents to date.

The company's data and analytics portfolio includes decision management, content analytics, planning and forecasting, discovery and exploration, business intelligence, predictive analytics, data and content management, stream computing, data warehousing,

AQ

Privacy and big data – good or bad?

Nigel Cummings

With the surge in organisations developing big data strategies and businesses keen to mine any and all information they collect, our privacy is at risk.

There is little we can do to curtail NSA, MI6 and, no doubt, most of the other world's 'Security/Secret services' but how far should we let this go? Are we happy about everyman and his dog using online data about us?

At a Big Data Summit recently, Martyn Croft, CIO at The Salvation Army, said that, 'The fact that organisations are examining ways to exploit big data without thinking about the privacy implications of collecting and storing it, could prove to be a huge problem.' Using the operation of his organisation, the Salvation Army as an example, Croft explained that information today is cheap, and that leads organisations to attach little value to it. 'The danger is that if the information isn't valued by the firm, then it will have little incentive to protect it properly'.

Now, while an organisation might not value an individual dataset, the individual to whom that information pertains usually will. 'For my organisation, much of that data is about an individual and I think if it's about an individual, you have a duty to take care of that data and take care over how you try to transform it and use it'.

Unfortunately few companies observe and practice that duty of care over data collection, and Robert Bond, Head of Data Protection and Information Security at law firm Speechly Bircham says such lack of recognition of the privacy of data is a risky business. 'The idea that we could do so much with this data so we'll just suck up and keep as much as we can for as long as we can, because you never know when you might need to use it, is highly dangerous. Because the more data you've got the more likely you are to lose it or to be hacked. The law in various countries, including the States, says,

'You can't keep data longer than is necessary. If you don't have a plan as to why you are keeping that data for that period, when something goes wrong you are going to get hammered that much harder by the regulators.'

Bond suggested one remedy for this is to encrypt data and have a data governance policy would also be helpful, as the regulators are likely to be less harsh if there is some sort of policy in place. Unfortunately, the law, at least in the UK, is very weak. The maximum fine that can be imposed is £500,000 and very few companies ever get as far as being prosecuted.

Many of us are involved with formulating the data collection structure, analysis and regulation of the data stored by companies for whom we provide our expertise. Perhaps from a moral perspective and to inhibit possible future ramifications in our career progression, we should emphasise 'as a matter of record', when we engage with data collection for analytical processes, our concerns for the security and protection of the individuals' data we analyse?