CRO FORUM

Big Data & Analytics: the algorithm of modern business

CROs in a changing environment

Table of Contents

Executive Summary

“Big Data is a crucial risk management tool in an ever evolving and fast changing world”

Renzo G. Avesani
CRO, Unipol

The smart phone checking your blood pressure, your car’s navigation system, the websites you visit daily - just three examples of datasets you create that are available here and now. They make your daily life easier but they also do much more: they model your health & motor liability risk profile.

The term “Big Data” is often cited but what does it mean? According to Gartner, “Big Data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making”[1]. It can be summarized by the “5V” definition: Volume, Variety, Velocity, Veracity and Value, where the analytics applied to this mountain of data has the capacity to provide the insights for both risks and opportunities. According to IBM, Volume refers to the vast amount of data generated every second; Velocity refers to the speed at which new data is generated and the speed at which data moves around; Variety refers to the different types of data we can now use; Veracity refers to the messiness or trustworthiness of the

[1] http://www.gartner.com/it-glossary/big-data

data, and Value refers to our ability to turn our data into value.

Big Data and its use will be transformative, and not just for insurers. For instance it can enable insurers to provide more risk prevention focused services as well as to offer more tailored insurance products.

There are over

google searches every single day. When it was first launched in 1998, it served 10,000 searches a day (ABI)

But each opportunity also carries risks. What are these risks and what are the best ways to manage them? These were the questions we asked ourselves when we set out to explore the new risk landscape ahead of us. The following pages will tell you what we found. We will also outline ideas on how to manage new risks based on the observations we made. The concepts we present to you are not final because the field is developing fast and disruptive change is always just around the corner.

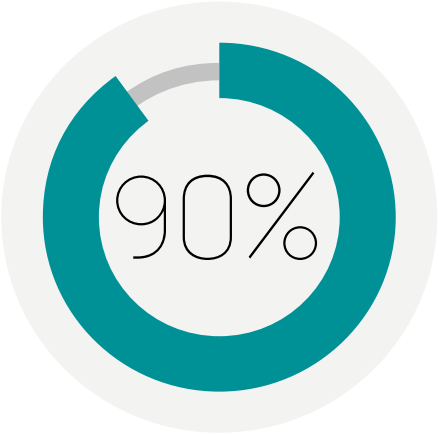
Nearly daily analyses announce how many billion devices connect via the "Internet of Things", how the number of tweets shared increased or by how many terabytes data globally increased. All this data is the foundation for new technologies from autonomous cars to health apps telling you to move more. Artificial intelligence using it creates smarter machines, starting to understand human-like trains of thought. Talking to your smart phone to ask for directions? You are already there. The examples show that Big Data has already become daily business and companies are heavily investing in it. This move into technology and data creates new partnerships allowing companies with completely different risk appetites and skill sets to work together to create something new.

While these are early days for Big Data & Analytics in insurance, there is the potential for it to become disruptive in certain lines of business. For example, the question of what happens to motor liability insurance in the age of autonomous cars is one that comes up immediately. For the policyholder, motor insurance using GPS data or where there is use of wearable health devices, more personalized products based on customer behaviour can become a reality. On the other side of the disruptive potential there is also opportunity. Big Data & Analytics enables insurers to glean a deeper, better understanding of risks, improved risk selection and potentially new ways of loss prevention.

These up and downside examples highlight the need that we must think about what Big Data & Analytics adoption means for the insurance industry. Moreover, this is not restricted to how our business develops. Big Data & Analytics also brings operational, regulatory and strategic challenges. The availability of data on a global scale and often real time will force us to be able to adapt fast. How do we deal with the data we collect in real time? The increased proximity to policyholders generates more visibility to them. Society and regulators have expectations on how the insurance industry handles data in tune with the technological developments we see. **This paper intends to provide a roadmap and practical tools for Chief Risk Officers to adapt to this changing risk management landscape.** Indeed, these developments

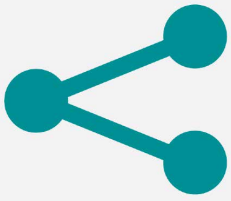
3,6
billion

global unique mobile subscribers in 2014 (GSMA)



of all the world's data has been generated over the last two years (ABI)

30
billion

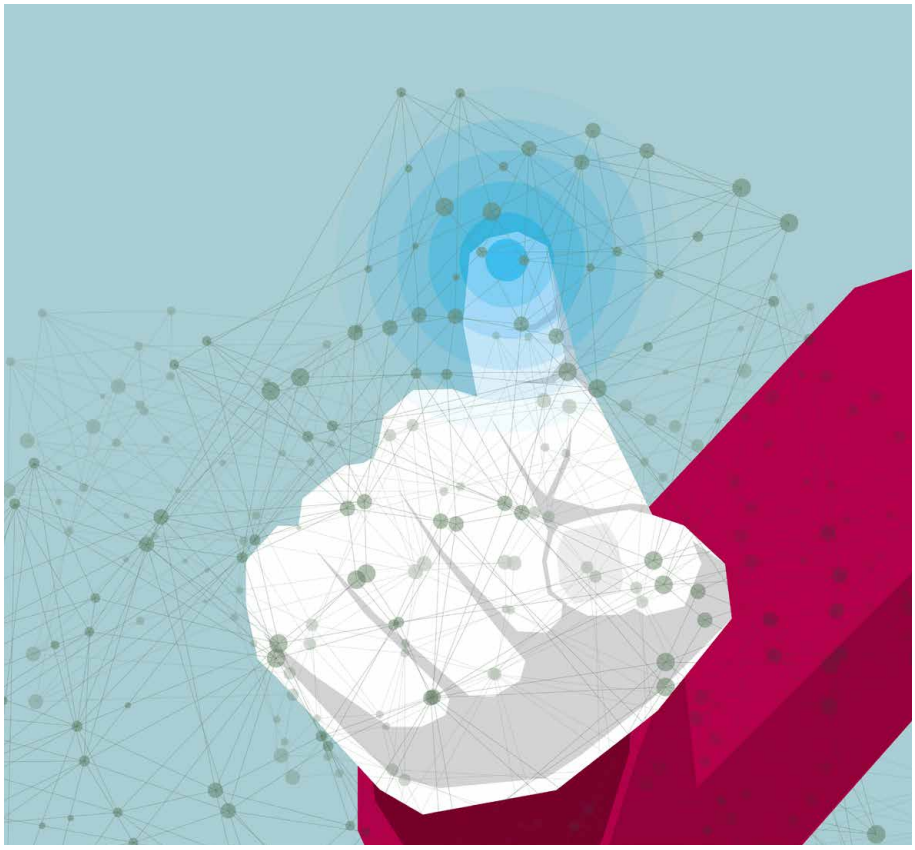


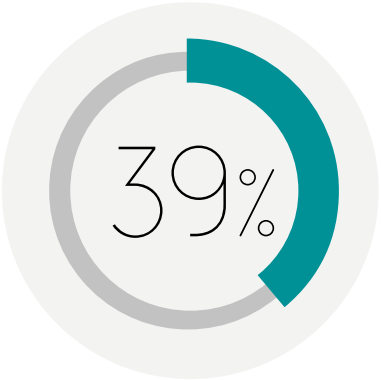
pieces of content shared on Facebook every month (McKinsey)

will also deeply affect the role of the Chief Risk Officer (CRO). With risk analysis strongly reinforced by the use of Big Data, CROs will need to manage key issues such as early new risks detection, risk profile, exposure control and risk quantification in a new, more real-time environment. Data miners new to the insurance industry will sit together with actuaries and underwriters. The new multi-disciplinary skilled people and partners along with the evolving technologies will require a new risk culture.
The CRO as a key partner in assuring that company policies are in line with the evolving expectations of regulators and society has a key role here. Attention to privacy and data protection is just one example. Managing partnerships with technology companies who have a different risk culture is another one.

As these examples show, Big Data has the potential to deeply affect our industry. From product design to operational execution and regulatory compliance, big changes will come with Big Data.

Each opportunity also carries risks. What are these risks and what are the best ways to manage them?





of the US insurers with written premiums of over 1 Billion $ use Big Data & Analytics

Introduction

Insurers have realized the importance of Big Data and are determined to investigate its benefits. And it is much more than an academic approach: the number of US insurers using Big Data & Analytics, for example, has already more than doubled between 2013 and 2014, reaching 39% of the players with written premiums of over $1Bn[2]. At same time, in the general technology spending the **Big Data market is growing seven times faster than the general IT market[3].** Along with opportunities, however, there are also new risks associated with Big Data.

The paper is composed of five main parts (plus annex, glossary and references).
This introduction outlines how Big Data & Analytics is going to influence the insurance business and its core processes.
Two cases focused on health insurance and telematics will follow as practical examples of Big Data & Analytics adoption.
The paragraph about the Role of the CRO will then discuss the risk management implications of Big Data and how CROs can help business executives implement successful Big Data & Analytics strategies, also in the light of a highly dynamic regulatory environment.
The fifth chapter (Conclusion) is where recommendations for senior executives are summarized.

[2]SMA White Paper, 2014 (Big Data in Insurance - Beyond Experimentation to Innovation).

[3]Communication from the European Commission issued on 6 May 2015 (A Digital Single Market Strategy for Europe).

Big Data started with science in high-energy physics, astronomy and genomics. Very famous digital giants have then added new dimensions to Big Data production and processing[4]. Sophisticated algorithms aim now at analysing complex networks and find relational and semantic interpretations of on-line activity, documents and images. In recent times, connected devices and sensors have started to form the so-called Internet of Things and are enriching available information with geo-spatial and other data.

Certain insurance-adjacent sectors are reasonably clear about the benefits of Big Data & Analytics.

In healthcare, even though data digitization and the integration of different sources are only in their infancy, personal and remote monitoring devices are adding lifestyle and behaviour data as well as disease-specific parameters to keep chronic patients monitored. Hence, lifestyle and behaviour-induced diseases become easier to prevent and manage, while clinical, claims and policy data can be analysed together to investigate treatment effectiveness.

In transportation, thanks to the data collected through sensors, GPS and social media, Big Data & Analytics supports the optimization of multimodal transport and traffic flow management.

The picture on page 8 outlines the role of Big Data in the insurance value chain. In essence, Big Data technologies help to improve the insurer's relationship with policyholders through more direct communications.

Big Data & Analytics improves the insurer's relationship with policyholders and builds upon additional information on covered risks.

Data from digital interaction and sensors allows a refined profiling for customers giving their consent during both the marketing and the servicing phase, thus leading to personalized solutions that better suit customer needs. Tailored services based on a higher number of touch points can increase customer engagement and retention improving risk assessment and loss prevention at the same time. The holder of a travel policy could for example revise their travel plan and so receive advise and alerts, or the holder of a motor policy could digitally communicate the maintenance status of their vehicle and so receive price incentives. In this context, it is also easier to encourage and reward persistently safe or healthy behaviour.

The health insurance case and telematics example discussed later in the paper will further help put these statements in context.

Finally, the availability of direct and complete information about accidents makes settlements quicker and frauds easier to detect.

Claims processing, for example, becomes faster and more reliable thanks to the integration of telematic and cartographic data while fraud cases are more easily identified detecting fraud rings based on Internet, social media and company data.

Direct digital information from smart devices and external data enrichment sources improve risk identification and risk monitoring, too. In this respect, for example, a semantic engine can be used to continuously identify relevant news on selected corporations from the Web and so enable Fire and Theft automated underwriting, while smartphone and social media data are the heart of research projects for loss prevention by predicting the spreading of infective diseases (for example the MoBS laboratory of Northeastern University in Boston; www.mobs-lab.org).

Progressively, pricing is being driven less by demographic data, and more by the risk factors themselves, so segmentation becomes more granular, pricing more risk-based, and risk selection more effective. In this context, satellite images for example are much better than postcodes to assess the actual location of real estate properties and their exposure to flooding.

Although there are multiple promising Big Data & Analytics use cases, the insurance sector faces a number of challenges related to the implementation and adoption of Big Data.

Compliance & Data Privacy

Compliance will require strong efforts as regulation evolves, also due to the nature of data: use of personal data, if not sensitive data, potentially leads to reputational risks, with people unwilling to disclose relevant data that may feel discriminated with respect to undiscounted premiums or limited services. Overall, since public sentiment varies with regard to privacy, depending for example on age and culture, it is important that insurers enforce transparency and accountability measures to prevent incidents and misunderstandings.

Customer Interaction

Search engines, online retailers and social media adopted Big Data & Analytics early on and have been processing large amounts of customer data for several years. These and other digital operators can therefore leverage customer knowledge and pervasive interaction to design and deliver insurance in innovative ways. To manage this strategic risk, insurers must consider the new competitive landscape and evaluate appropriate measures such as distinctive customer interaction and leveraging unique physical networks that deliver tangible services.

Technology and the Ecosystem

The amount of technology involved in Big Data & Analytics initiatives and operations triggers specific business risks, too. Traditional information systems are generally inadequate for Big Data and the required investments may have an uncertain payback. Rapid technology evolution alters customer needs and can suddenly reduce the attractiveness of sophisticated solutions. Partnerships with external technology companies with possibly different risk appetites require some work to implement best practices.

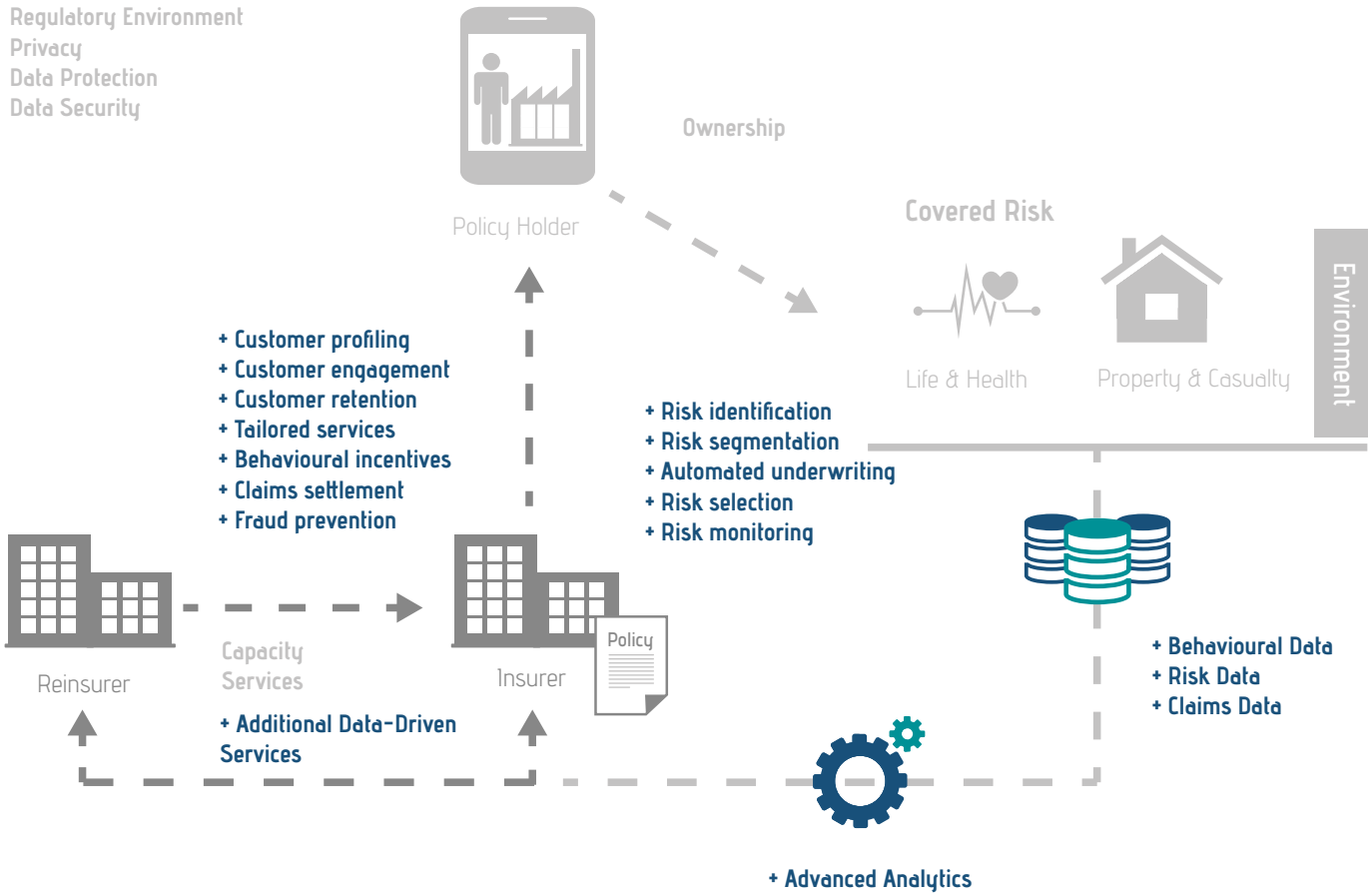
Analytical modelling capabilities & Data Quality

If insurance information collection relies on evolving technologies, analytical models might require adjustments to match the new data, too. Data quality frameworks to avoid erroneous conclusions will also be key.

The following sections will specifically further address challenges pertaining to Big Data & Analytics implementation. These challenges include changes to the insurers' risk profile, data quality, the need to acquire new capabilities in the analytical domain, and the regulatory environment.

> **Data is insurers' raw material"**
>
> Mohamed Baccouche
> Life CRO, Axa

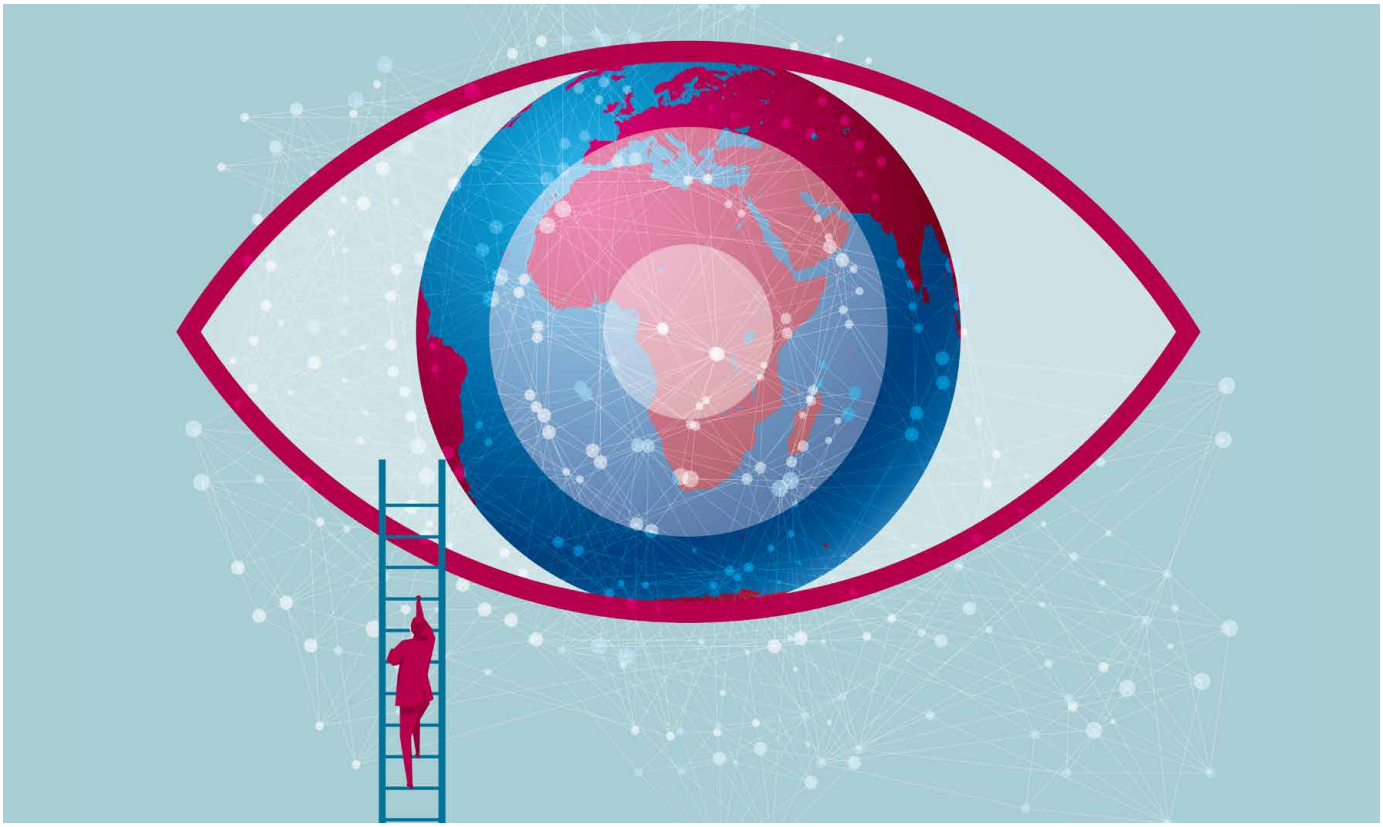
Big Data & Analytics in the insurance value chain



Frontiers in Massive Data Analysis according to the National Academy of Sciences (NAS, USA, 2013)

"Big Science" (e.g. energy physics, astronomy and genomics) has been generating massive amounts of data for quite some time, however more sectors are now confronting the challenges of massive data. ECommerce, search engines and social media, for example, have many million users. The search for relational and semantic interpretations of the phenomena underlying the data that these users generate requires sophisticated techniques. New challenges in terms of data generation, preparation and analysis are emerging. On one side, all statistical tools are based on assumptions about characteristics of the data set and the way it was sampled, and those assumptions may be violated in the process of assembling massive data sets. Second, but not less important, tools for assessing errors in procedures, and for diagnostics, are themselves computational procedures that may become infeasible as data sets move into the massive scale. A key point of the NAS report is related to skill need: computer scientists involved in building Big Data & Analytics systems must develop a deeper awareness of inferential issues, while statisticians must concern themselves with scalability, algorithmic issues, and real-time decision-making. Domain scientists have also an essential role to play because of the explosion of design decisions and possible directions that analyses can follow. Specific forms of data create particularly interesting challenges, i.e. human language, images, geospatial and time-related data, and networks and graphs. Data sources operating in real-time also provide an interesting challenge, since statistical research has rarely taken computational complexity into account. Finally, the report emphasizes that massive data analysis is not one problem or one methodology.

Therefore, massive data analysis cannot, in general, be reduced to turnkey procedures.



Business cases

The number of practical Big Data & Analytics use cases in the insurance industry is rapidly growing. Success is not guaranteed but the scenario looks highly interesting: if not all of these use cases have proven to have a positive business impact, there are already several examples from different lines of business of how Big Data & Analytics are successfully adopted.

In this section, we will use two cases from health insurance and telematics to demonstrate that creating business value out of the combination of Big Data and Analytics is no longer a theoretical goal.

Appropriate analytical techniques may represent a strong basis for new business ideas and CROs can support setting up similar use cases as well as advise the business with regard to ethical feasibility, an additional powerful defence against reputational risk, as we'll see.

Creating business value out of Big Data and Analytics is no longer a theoretical goal.

a. Big Data & Analytics in health insurance

The growing availability of health-relevant data and methods to turn the numbers into knowledge hold vast potential for improving the quality of disease management programs. At the same time, any successful implementation of Big Data & Analytics calls for significant investments. And, as personal data are involved, the legal and reputational risks are also significant.

In Europe, many insurance companies operate disease management programs as a well-accepted customer service and as an effective measure to control the increasing costs of chronic diseases, especially diabetes, adiposities, chronic heart failure and back pain.

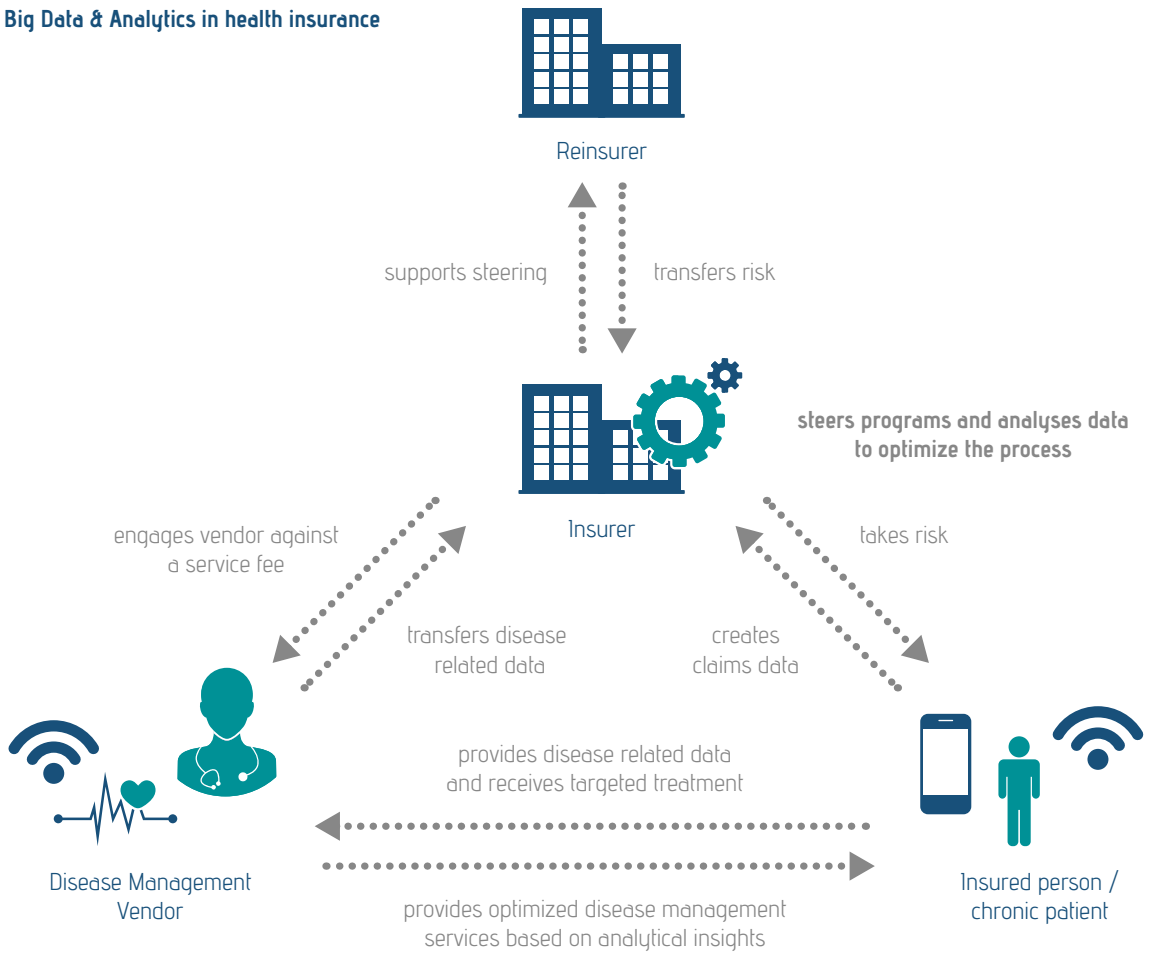
The programs bring the insurance company closer to the insured, help to position the company as a reliable partner in health questions and prevent individual health problems by early detection and targeted intervention. Big Data & Analytics improve the efficiency of the program in various ways and is gaining in importance with the increasing availability of external data.

Who is involved in disease management programs

In health insurance, Big Data & Analytics can facilitate the prevention of adverse cost trends.

Offered as a service to insured persons with chronic diseases, disease management programs generally focus on lifestyle-related syndromes, such as diabetes and chronic heart failure. Measures go far beyond the classic insurance coverage and may include education, consulting and/or medical monitoring of patients with the purpose of supporting a healthier lifestyle, increasing quality of life and reducing healthcare costs. Connected objects related to well-being could play a key role.

In most cases, specialised vendors engaged by the insurance company carry out the program measures while the insurance company itself usually steers the process and has a coordinating role. Reinsurers may also be involved in a consulting function, related to medical content or data analytics, and as risk carriers (e.g. with performance guarantee solutions for programs).



Risk identification and customer satisfaction: two goals of Big Data & Analytics in health insurance

Observations from patient monitoring are a highly valuable information asset to define reliable long-term cost trends. For example, risk loadings based on medical expert-systems can be adapted to reflect market and product-specific conditions. A higher degree of market specification based on Big Data & Analytics insights contributes to a more reliable medical underwriting decision and an identification of additional insurability.

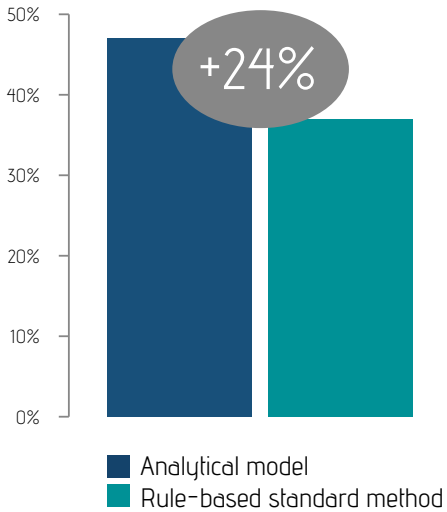
At the same time **Big Data & Analytics can improve customer retention.** Individualized care offers help patients to better live with their disease and experience shows that satisfaction with the insurance company is strongly increased by such programs. Analyses prove that the cancellation rates are clearly lower for participating insureds than for other chronic patients in the portfolio. This increase in customer retention, especially in the group of younger chronic patients, contributes to a better portfolio mix and a buffered ageing of the portfolio.

Advantages and opportunities never walk alone; a danger for insurance companies may come from the discrimination of insured persons caused by advantages for compliant disease management participants who are willing to provide their data. Examples of such benefits are premium reductions, access to other plans or products and access to more exclusive provider networks. In this regard, **it is crucial for the insurance company to adhere to the legal framework and to carefully assess an ethic code of behaviour** in order to limit as much as possible the risk of reputational damage due to unjustified unequal treatment.

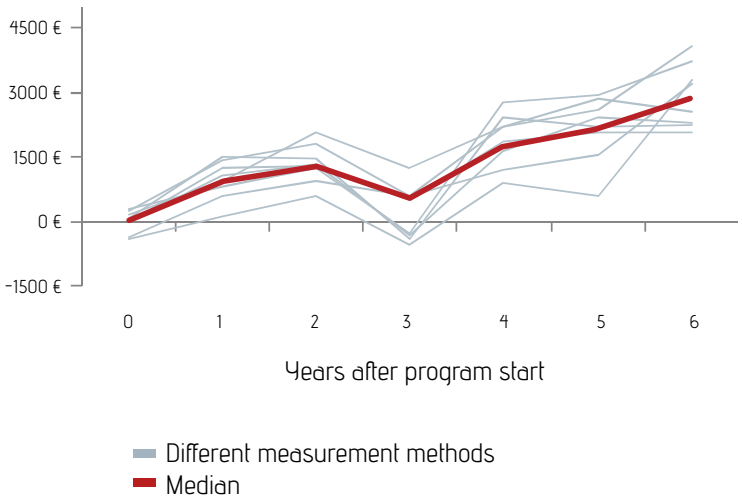
Analytical techniques support identification of candidates and economic measurement of disease management programs

Historically, the identification of patients for disease management programs had been based on a static set of rules drawn from medical experience. Inclusion (and exclusion) criteria were based on the occurrence of certain diagnoses or age. The availability of big data information on risk factors such as education level and physical activity, allows for example, in a diabetes program, enhanced identification approaches. In the diabetes program of a German private health insurer, a 24 percent increase in the identification rate of future high-risk patients was measured through the analytical approach (see the left chart). Assuming that these patients offer the highest intervention potential, the analytical identification approach increases the effectiveness of the program. In the example shown, no clinical data, e.g. glucose level or blood lipid levels, were used in the identification model. Including such parameters should further enhance the identification process.

Identification rate of high-risk diabetes patients



Savings potential of a chronic heart failure program



Measuring the economic impact of a program is an important question from a business perspective. In many disease management programs, however, the evaluation is impeded by a self-selection bias arising from the free decision to participate. The optimal measurement technique in this situation is subject to controversial debate in scientific literature, and reported results are very heterogeneous, as they depend on the technique applied. As the example of a chronic heart failure program of a German private health insurer shows (see chart on the right), an overall increase in savings over time is identifiable. However, different analytical measurement techniques lead to a certain range in terms of savings potential even though all risk factors measurable from claims and policy data are controlled. Ideally, the program effect is determined by a comparison between participants and a non-participating control group, which is similar with respect to all disease-related risk factors. Consequently, the consideration of additional risk factors for chronic heart failure, e.g. nutrition and smoking habits, captured by electronic devices, is assumed to further reduce the variation of measurement results.

How can the process of health insurance be supported by Big Data & Analytics?

Every stage of disease management programs, from design and execution to result measurement, can use Big Data & Analytics techniques.

	Content	Analytics	Big Data
1 Design	Identification of care gaps and definition of bespoke intervention	Identification and segmentation of participants (e.g. clustering, cost prediction)	Behavioral and socio-economic data (e.g. physical activity, education level, ...)
2 Execution	Education, support and medical monitoring of participants	--	Clinical data from smart devices (e.g. bluetooth scales, glucose level meter, ...)
3 Evaluation	Financial and medical outcome measurement	Analytical measurement techniques (e.g. matched-pair)	Behavioral and socio-economic data (e.g. bluetooth scales, glucose level meter, ...)

1 Design phase:

- A crucial aspect of the design phase is the identification of care gaps and disease-related risk factors in order to allow the profiling of the affected population into risk segments, with tailored intervention plans for each segment.
- While some risk factors (age, diagnoses and utilisation of medical services, only to mention few examples) are directly accessible to the insurer from claims and policy data, important factors as behavioural information are usually not captured by health insurance companies. To fill this gap, external Big Data sources can be tapped, and combined with internal information, always acting within a solid legal and ethical framework. Behavioural information from wearables or social media on an individual's smoking status or physical activity, for example, would be valuable for risk segmentation, but in many markets it would not be permissible due to privacy laws.
- Clinical data, for example patients' blood pressure or lab results, may also be important. In some markets, it is possible for insurance companies to capture this information for the claims adjudication process. In order to make use of all available information, the segmentation and identification process needs to be supported by analytical techniques such as individual cost prediction models for patients or clustering algorithms identifying homogeneous groups of patients.

2 Execution phase:

- In most cases, specialised vendors are responsible for execution. They provide disease-specific education and consulting by medical experts - nurses, physicians or psychologists - with the aim of increasing patients' knowledge of their diseases. This in turn is to enhance self-management and self-motivation for behavioural changes, to improve compliance with therapies and support in organisational questions, such as recommending specialist physicians.
- The consulting approach is often supported by medical monitoring of clinical/medical parameters, such as weight, blood pressure or glucose level. The monitoring process can be significantly improved through easy-to-use electronic devices that regularly measure disease-relevant parameters and transfer the related information to the disease management vendor in real time. For example, a strong increase in weight, glucose level or blood pressure may directly alert the disease manager and trigger a call to the patient. This tele-monitoring facilitates early detection of critical developments, such as heart attacks, and allows for timely intervention. In this way, emergency situations, hospitalisations as well as high rehabilitation and long-term costs can be avoided.

3 Evaluation phase:

- Insurers usually focus on two aspects when evaluating disease management programs: medical cost savings and improved patient outcomes. Reliable measurement of these factors depends on the availability of a valid comparison group, which can be a challenge because of non-measurable differences between the patients who accepted and those who rejected, e.g. the motivation to change the health behaviour.
- The analytical challenge is now to identify a subgroup that is as similar as possible to the group of participants with regard to all risk factors. Based on different analytical techniques, e.g. matched-pair algorithms, that allocate controls to each participant based on similarity in risk factors, a more valid comparison is possible. However this approach requires a comprehensive picture of the individual risk structure, which can be sharpened based on external big data information.

b. Big Data & Analytics in motor telematics

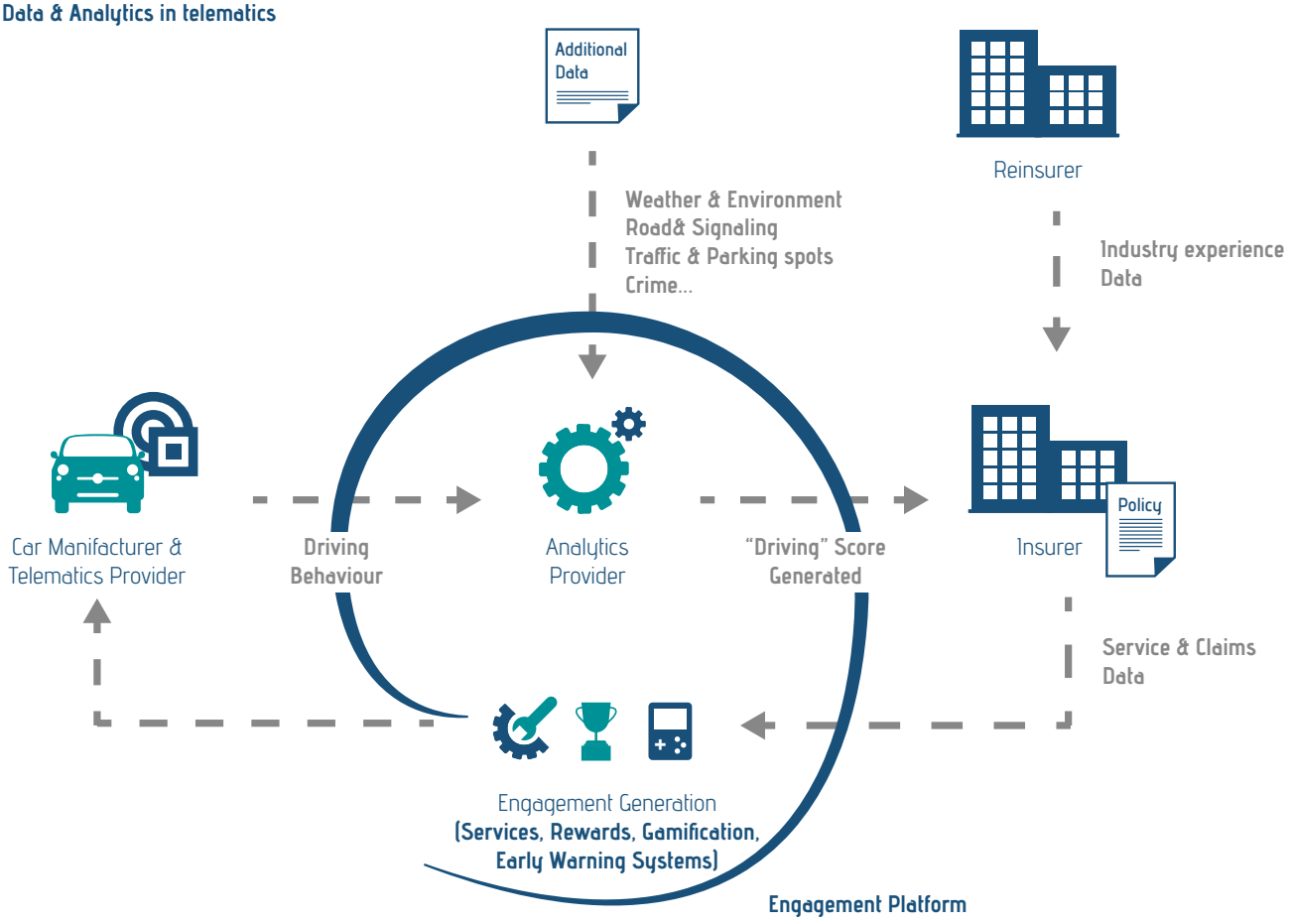
Transportation offers a wide territory to leverage Big Data & Analytics and set up a safer environment with many solutions, eventually including driverless cars. A new frontier for insurers but one where new players could enter the market.
The traditional motor insurance risk segmentation has been mainly based on data related to the person (demographics, behaviour, experience, even sex gender not so long ago) and the vehicle (mainly related to its power/ speed).
With adoption of Big Data & Analytics the insurance industry started to enhance their knowledge base and offerings with valuable data. Information gathered on drivers' behaviours are, for instance, able to provide the insurer with an effective prevention against fraud, at the same time also providing the policyholder with useful services and tailored/competitive prices. This type of data are however under strict data privacy controls.

The insurance industry started to enhance their knowledge base and offerings with valuable data.

Who is involved in programs based on motor telematics?

Motor telematics programs involve several players operating alongside the insured and insurance company. In details, car drivers accept that their data is collected by service providers, while the insurance company, in addition to its traditional role of covering the specific risk, also plays the part of an overall coordinator. Vendors specialising in motor telematics and analytics can provide the insurer with scores based on data analytics through end-to-end motor telematics solutions including data collection, data submission and data analytics.Also car manufacturers have a role, hosting in their new cars optional motor telematics devices gathering all desired data (example: seat belts, number of passengers, braking, speed, GPS information…).

Big Data & Analytics in telematics



Last but not least, regulators and lawmakers respond to existing and potential privacy concerns relating to the use of personal data and provide insurers with a related compliance framework.

Big Data & Analytics can lead to mutual advantage for both, the insurer and the policyholder.

Big Data & Analytics: what's worthwhile in motor telematics?

Big Data & Analytics allow the collection of a variety of information that can lead to mutual advantage for both, the insurer and the policyholder.

Phone apps, telematics boxes or other digital devices may be set up to collect data on driver behaviour such as acceleration, braking or the way corners are rounded, comparing them to the road signalling, road type and setting. The assessment of driving style may then be based on the pace, smoothness, calmness, anticipation, and consistency of the driver in his or her manoeuvres.

Customer experience can be highly enjoyable. While insured drivers can in fact learn a lot about their driving behaviour (having fun at the same time) through a set of scores, points, and other distinctions, all this data can turn into a lower policy price. Customer satisfaction and retention are improved not only through a more favourable and tailored pricing (including pay-as-you-drive solutions) but also through additional services, such as gamification, tips how to reduce fuel consumption and emergency services, in case of accident and theft.

On the insurer side, several business goals can be reached thanks to motor telematics adoption:

- get closer to the insured driver, increase customer engagement
- offer new products and services with a dynamic pricing thanks to gathered data
- improve loss prevention influencing the driving behaviour through gamification (example: fuel efficiency challenge) and direct feedback
- offer additional services (online/ mobile claim handling, emergency services, maintenance offers, life-style oriented services, loss prevention oriented services, etc.) based on transport and external data (example weather conditions)
- create new cross-selling opportunities as well as third party service offering (such as car maintenance) by building a community and gamification-based engagement concept.

Example: Motor telematics cases

Unipol - The data collected by the service provider via an on-board unit (OBU) include trip data, type of road used, time of the day and crash data (timestamp, multiple location records, latitudinal and longitudinal acceleration, braking…). This data enables Unipol to design a pay-as-you-drive model, an automatic call system which contacts emergency services in case of an accident, and a car theft location service. A set of incremental rewards can also be given when the driver adopts good habits. This business model enables the reduction of accidents among drivers subscribing to the platform and to increase Unipol's profitability even though customers receive discounts up to 50% (for fire and theft covers).

Axa - Axa launched its Drive Coach app in April 2015. Once launched on a smartphone or the Apple watch, the app analyses accelerations, brakes, and corners performed during the journey. It provides users with a fun way to learn about their driving behaviour through a set of scores, points, and other distinctions such as medals. The more points the driver gets, the more balanced their driving behaviour is. Then the app allows comparing scores with other drivers using the app, sharing it on Facebook with friends, and getting tips on becoming a safer driver and further improving their driving style.

c. Conclusions

In general, Big Data & Analytics have a strong potential to improve the quality of life of chronic patients by improved disease management programs as well as the driving behaviour of insureds by motor telematics.
A greater client retention is a key result due to personal benefits, additional services and price advantages.

In both business cases, analytics based on the combination of internal and external data facilitate a better understanding and mitigation of risks. Strong analytical capabilities and access to appropriate data are key requirements for successful implementations. Furthermore, both examples show that customers should understand the processing of their own data and should be able to give open consent to its use. Insurers, accepting their data controller role, must work within sure legal borders and adopt a strong ethical code.

Big Data will enable new tailored and dynamic insurance solutions”

Cristoph Flückiger
CRO Group Operations and Technology Risk, Zurich

The CRO's role in a new light

As we have seen above, with Big Data & Analytics new opportunities walk hand in hand with risks. While the opportunities may allow insurers to achieve measurable business results thanks to an improved customer experience and a more reliable risk identification, at the same time risk management teams have challenging questions to address.

Be it a business revolution or not, **Big Data adoption is clearly bringing a strong evolution in CRO's role as a second opinion provider for the company's executives.**

Risk analysis is strongly reinforced by the use of Big Data and advanced techniques, allowing CROs to manage key issues such as early new risk detection, risk profile and exposure control, and risk quantification, even working in a management environment characterized by faster execution time, asymmetric information and rapidly changing technologies.

Managing new risks in the Big Data era is a matter of techniques but also of culture. Risk management now requires multi-disciplinary skills, with actuarial competence and data science skills working together: computer scientists and machine-learning developers work side-by-side and leverage the knowledge of experts from the natural risk disciplines as well as from social or behavioural sciences. Skill set evolution needs to take into account new technical tools development, and to be supported by a new risk and organisation culture, also allowing the insurers to partner with companies that have a totally different risk appetite.

Alongside with need for a wider mix of competences, Big Data adoption requires CROs to handle risks also in light of the impact of appropriate data governance, evolving legal/regulatory environment, and ethics. The uncertain legal environment, with a higher sensitiveness for data privacy emerged at global level, requires companies to behave very carefully, even beyond the strict legal requirements: a solid and transparent ethical approach will be crucial to ensure that customer experience does not translate into a negative feeling with high reputational risk.

As seen in the above business cases, Big Data & Analytics adoption may have a strong and positive influence on customer experience, while it allows the insurance companies to provide their customers with high value services and set highly customized and competitive prices, leveraging a more granular knowledge of individual risk profiles. But opportunities bring risks, and reputational risk linked to the potential impact of customers feeling discriminated is then a key area for CROs.
Data reliability (essential in defining the most appropriate business propositions) also remains crucial for CROs, together with data governance: data integrity needs to be protected against a rapidly growing cyber risk and, at same time, the use of personal data requires accurate management processes securing that it is always authorized by aware customers.

Analytics based on the combination of internal and external data facilitate a better understanding and mitigation of risks.

Opportunities in risk analysis

- Early risk detection, risk identification, risk understanding
- Risk profile, exposure control
- Risk quantification, models verification and calibration

Adapting skill sets and risk culture

- New complementary skills in multi-disciplinary teams
- Partnerships and technology

Developing control frameworks that adapt to Big Data & Analytics velocity and dynamic regulatory environments

- Second opinion
- Regulation
- Data governance

Overall, two broad objectives for risk management may be expressed:

- Address direct risk management needs in terms of early risk detection, risk monitoring, risk quantification.
- understand the use that the business makes of Big Data & Analytics techniques and the assumptions it makes when defining algorithms in order to continue formulating accurate second opinions.

The pursuit of these objectives should go hand in hand with a strong risk culture adapted to the first lines of defense and with control frameworks that ensure data quality and regulatory compliance.

Opportunities in risk analysis

CROs' action may be supported by the adoption of Big Data & Analytics techniques related to:

- early new risks detection
- risk profile and exposure control
- risk quantification.

While adopting these new techniques, data integrity and cyber risk management will be key.

Early risk detection, risk identification, risk understanding

Semantic analysis and cross-referencing of key-word occurrences can reveal signals of emergence or potential for litigation even when they are still weak. Therefore it may offer valuable support to the action of CROs for early identification of new emerging risks. This will become increasingly important in a fast-changing environment characterized by flexible business models and partnerships. Big Data & Analytics can also be effective in providing valuable information on risks that are difficult to quantify and require a continuous monitoring, such as reputational risks. As an easy example, we have already seen how customer experience and satisfaction in health insurance can be positively influenced by Big Data & Analytics. When applied to the improvement of disease management programs, content analytics can play the role of mitigating the risk of insufficient service quality: service quality can be monitored by analysing client feedback on the website of an insurance company or medical treatment vendor or in social media platforms, such as disease-related forums.

"Big Data is about the improvement in identifying and understanding risks, including the complexity of risks and their correlations"

Bernhard Kaufmann
Group CRO, Munich Re

Risk profile, exposure control

Due to a more intense customer engagement and the collaboration with digital native partners (with different culture and appetite for risk), company risk profiles may change and a shift towards product liability is expected.

Usage-based insurance (where price is directly linked to actual use) will also have a large impact on the way exposure and capacity usage is monitored.
Furthermore, accumulation control will be impacted as underlying data might become widespread, with partners potentially requiring different technologies and different models to interpret them.

Big Data & Analytics will also allow CROs to obtain more granular knowledge of risk profiles and, as seen in the case of telematics, with benefits for policyholders. In summary CROs will help their companies in understanding which behavioural characteristics lead to insurance-relevant claims, how to measure them and how to find incentives to avoid them.
CROs should also encourage leveraging Big Data for new exposure control,

Big Data & Analytics will allow CROs to obtain more granular knowledge of risk profiles.

including accumulation measurement. This will be enabled by techniques such as real-time control and measurement, geographical or time-based accumulation.

Geocoding for example can enable a better analysis of concentration risk by using satellite data on weather patterns and their impact. More precise and granular data will also support risk reporting and improved aggregation.

Risk quantification, models verification and calibration

Predictive modelling can use inputs from digital underwriting data, large-scale, automated claims analysis, and analyse patterns of past events to model future risk outcomes with greater granularity. The improvement of forecasting possibilities will either require the CROs to develop new models or to enhance existing models based on external inputs. CROs should play a strong role in determining areas where it is relevant to turn to Big Data analysis in order to complement current existing databases and risk models.

The preliminary work on these complementary models will be essential to provide the most useful risk outcomes. On the one hand, Big Data & Analytics will contribute to filling current risk knowledge gaps. Moreover, traditional data sets will provide Big Data analytics with the required verification and calibration (e.g. co-variance analysis). In that sense, it could be similar to the process followed by Google Flu trends complementing data sets with CDC data[4].

Potential pitfalls that CROs should help avoid include issues arising from the erroneous interpretation of correlations coming from big data findings that are not necessarily representative of causality links. Reliance on Big Data should be balanced with verification and calibration possibilities to validate models and findings. Erroneous codes or incorrect calibration of underlying risk scoring could have long-lasting consequences and should be avoided, as **insurers move into a context of faster execution time, complexity and rapid risk model changes,** while maintaining their business model of long-term commitments towards their policyholders.

If on one hand the collection and analysis of external data, allowed by advanced analysis techniques, offers concrete chances to develop new forecasting models or to enhance existing models, on the other protection of internal data remains a crucial area for CROs.

New complementary skills in multi-disciplinary teams

Managing new risks in the Big Data era is as much a matter of techniques as it is of culture.
In other words, in order to formulate accurate second opinions also on pricing, underwriting and claims management processes, CROs will have to define risk management needs in terms of skill sets, technical tools (analytics tools, data integration needs between Big Data and traditional databases, user interfaces etc), risk and organisation culture.

With the appropriate competence available, Big Data & Analytics can be an enabler of more integrated interactions on risks with the business, as it allows for quicker risk analyses and faster interactions. It will likely further strengthen data-driven decision-making and allow for appropriation of risk appetite by the business.

In this new risk culture, the first line of defense will also be data scientists developing algorithms that will do the calibration of risk profiles, pricing, and reserves. Awareness should be strong within this community on the long-term consequences especially for some lines of business such as life and savings. Maintaining the ability to interpret algorithms should be part of the objectives of the first line of defense.

[4]Using Networks to Combine "Big Data" and Traditional Surveillance to Improve Influenza Predictions, Nature

Getting the right skills and building up the right culture in the right time is an art”

Elena Rasa
CDO, Generali

There will be a rapidly increasing need for risk management teams to become even more multi-disciplinary.

To elaborate, for example interpreting results of Big Data based models (e.g. predictive), existing actuarial teams will have to be complemented by data scientists and other related skills. On the other hand, actuaries could receive training to gain data scientist skills. But not only: additional expertise from other disciplines such as social or behavioural physics will also be required to interpret data.

Also the day-to-day implementation of risk analysis and governance will not only require the involvement of more experts, but also faster execution time will be required (e.g. expertise in specific health sectors, meteorological science, etc). Basic legal and compliance knowledge on data-specific aspects could be added to risk management skills, especially in view of possible geographical discrepancies in data privacy laws, as we will see later in this paper.

In order to actively support this development in the modern risk management organisation, functions such as data auditor or chief data analytics officers could be established. It is also important to provide these new job families a perspective and therefore actively work together with human resources in order to develop a carrier path particularly for recruitment targeting capabilities.

CROs will have to define risk management needs in terms of skill sets, technical tools, and organisational culture.

Partnerships and technology

Partnerships along the whole and value chain are expected. Those partners can be start-ups, digital natives, partners from different industries with a totally different risk appetite and risk culture or even awareness. As Big Data & Analytics goes alongside technological partnerships with companies outside the insurance industry, efforts should be made by insurers to ensure there is a sufficient level of risk control measures, awareness, culture and maturity. Maintaining the ability to control and test these measures will be a central part of business continuity management. Risk could even be higher in case of direct investment by the insurers in the partner company, when CROs would need to manage, in addition to direct credit risk, also an extended reputational risk linked to the partner's behaviour.

In this environment, the CRO's role is to ensure that the coordination framework of those partner networks (due diligence, on-boarding and on-going oversight) is appropriate, so that the different risks (outsourcing, operational, compliance, counterparty, dependency on partner expertise etc.) are correctly managed. Depending on the partner's involvement in the distribution of insurance products, emphasis should be put on mis-selling risk management, carefully thinking about how to oversee this risk without limiting the partners entrepreneurial spirit and still meeting the regulatory requirements.
There can be various business models, starting from a very simple collaboration up to a strong integration, with legal questions on data ownership and control arising. And in case of additional service linked to an insurance product (such as in the above-mentioned cases, for example), the CROs need to take care of improving the oversight model.

Depending on the business models that develop in the field of Big Data & Analytics, the need to purchase external data and to invest in analytical capabilities may also emerge, and the use of widespread data storage would necessitate to manage adequately risks to availability of data, reliability of the business model and data protection.

Technology is a key point in developing adequate Big Data & Analytics capabilities, and investing in appropriate analytics tools will be required. But the fast changing technological environment is a challenge.

Ensuring adequate cyber security is also key: preventive and preparedness measures against data loss and unauthorized access must be taken (see the CRO Forum paper, The Cyber Risk Challenge and the Role of Insurance[5]). As data volumes grow and a wide array of technologies is integrated into value chains, the patterns of exposure to cyber risk can vary rapidly.

Second opinion

Big Data & Analytics will bring acceleration and greater interconnectivity, which will make it easier to avoid having a segmented view on risks. Risk taking activities will be challenging to oversee as the distribution channels, market and customer base may substantially change.

Diversified skill teams, faster execution time and asymmetric information will characterize the risk management environment more and more; CROs will be called to provide their executives with high quality second opinions.

To be able to develop a sound understanding of the business and the use it makes of Big Data & Analytics techniques, risk managers should be involved early on in the design phase of the process and tools.

> **"Partnerships with companies with a totally different risk appetite, culture, or even awareness are expected"**
>
> Patrick Raaflaub
> Group CRO, Swiss Re

[5]CRO Forum, Cyber resilience: The cyber risk challenge and the role of insurance, December 2014.

Where digital underwriting is developed particularly for large and complex transactions, risk management should be more involved in overseeing the underwriting framework.

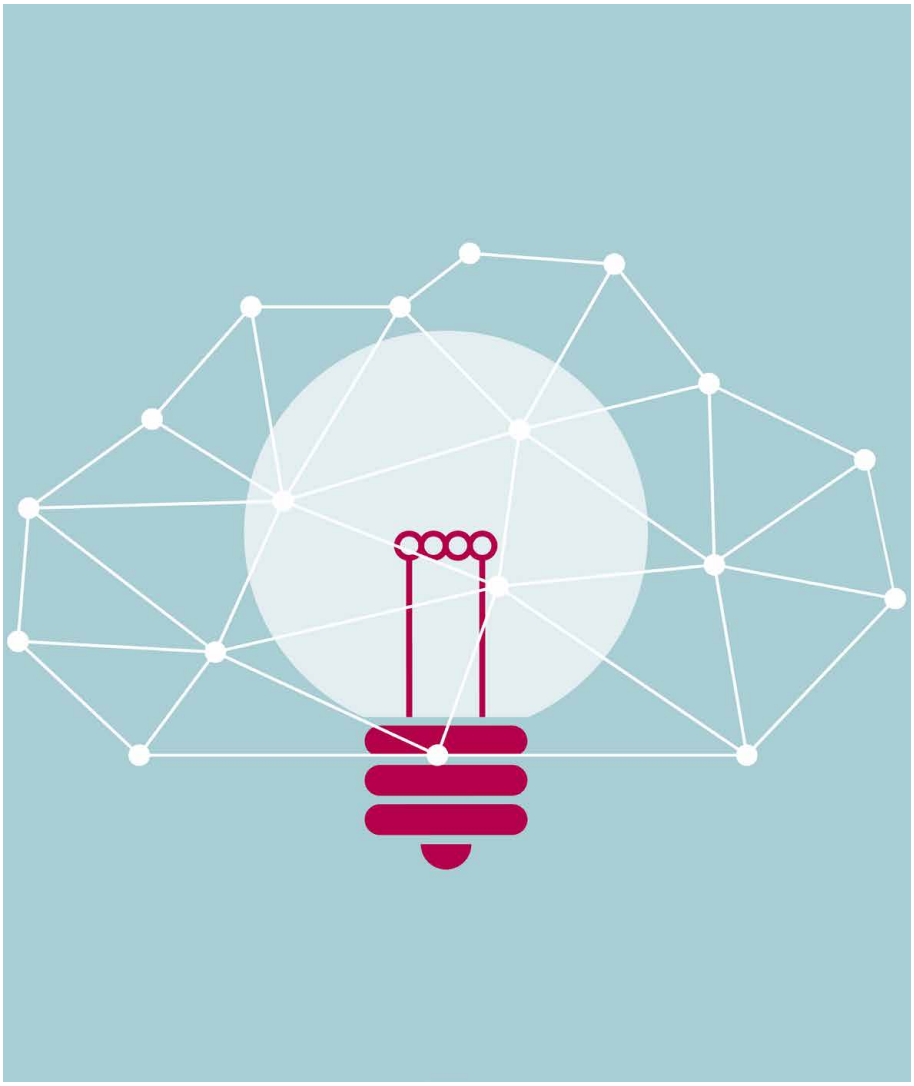
Risk managers should address the challenges that arise from the fact that product pricing becomes more dynamic and is directly influenced by policyholders' behaviours. Models could also encompass some kind of self-learning features which will require the model to adapt and assumption testing approaches to maintain oversight. Self learning models could result in additional risk. What happens is that the pricing models might learn the wrong things (inappropriate pricing, model results).

Heightened complexity and sophistication have the potential to mask some risks and can be misleading. In this context, the role of risk management will be to ensure that risks related to this increased complexity are clearly understood. Maintaining an understanding of algorithms developed by the first line of defense should be an important objective of the second line of defense.

There is a need for developing control frameworks that adapt to Big Data & Analytics velocity and dynamic regulatory environments.

Regulation

In the insurance and reinsurance business, regulatory changes occur regularly for all types of risks, and adherence to them is a key requirement. On top of that any discussion of risks arising out of Big Data must consider increasing levels of privacy concerns from the general public. In the same way, regulators and enforcement bodies are today increasingly interested in Big Data and will continue to identify the new ways in which Big Data presents a risk to privacy and data protection. It is therefore essential that the CRO is able to provide an accurate and up to date picture of the relevant legal requirements, regulatory outlook and public perceptions.
Even while regulation allows some room



for companies to set up appropriate processes to leverage Big Data & Analytics, the CRO needs to be able to identify areas where innovation is far ahead of regulatory thinking in order to anticipate potential concerns. In addition, legislation and regulatory approaches are not consistent across borders so the CRO's role is also to provide an accurate picture of these differing requirements.

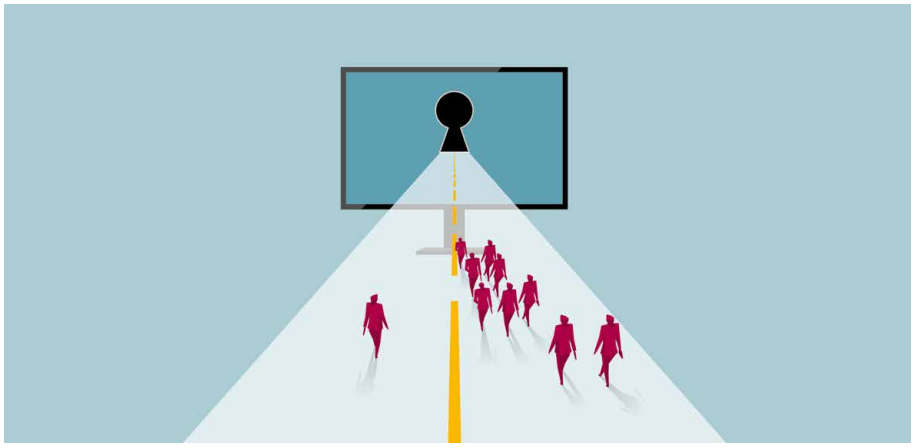
In this environment, there is a risk that Big Data processing which is, in fact, compliant with the law, still attracts regulatory scrutiny; therefore a clearly defined set of ethical standards may represent a powerful tool for the organisations to ensure that trust is maintained and to avoid reputational damage.

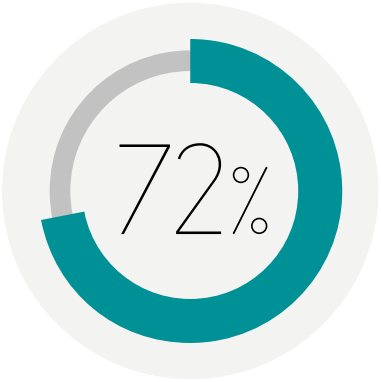
Having said that, it is important to take into consideration that the message from regulators (specifically in Europe) is that customer needs to clearly understand the use of his data, the company has to explain it in an easy and fair way, and the customer has to give his consent.

The CRO plays a key role in ensuring that the company's resources are properly aligned in order to identify how these rules affect Big Data. Key drivers should be which safeguards and controls are necessary to meet compliance standards, consumer protection laws and growing public awareness of the sensitiveness of personal information.

In particular, carelessness with regard to sensitive personal data such as health information or disclosures to third parties (for example when combining datasets to gain new insights) can quickly damage reputation or lead to legal actions.

The current negotiations over the future of data protection regulation in the European Union provide a useful example of how the CRO can contribute in the definition of a Big Data strategy which is not only innovative in the way it uses data to generate benefits, but also compliant with the law and respectful of public concerns about privacy.
The outcome of the EU negotiations is expected to be a General Data Protection Regulation that will apply directly in all EU member states.





of Internet users worry that they are being asked for too much personal data online

(European Commission)

The Regulation is likely to be in force for a number of years, during which time the capabilities of Big Data will likely expand and develop in ways which the legislators cannot be expected to foresee.

The CRO plays a crucial role in:

- Coordinating relevant resources such as Big Data practitioners, data protection experts and legal teams to make sure there is a common understanding of privacy risks and how they can be mitigated
- Developing key messages that ensure there is no disconnect between business objectives and regulatory perceptions

As part of the European Commission Digital Single Market Strategy, on-going consultations on geo-blocking, platforms and standards will also be key determinants of the regulatory framework for big data.

A clearly defined set of ethical standards may represent for the organisations a powerful tool to ensure that trust is maintained and to avoid reputational damage.

Data governance

As we have seen in the above paragraph, beyond regulatory requirements, the increased volume of data and the use that is made of it will require ethical choices from (re)insurers in order to avoid reputational risks. CROs can provide advise on the definition of a company's own code of conduct that would de facto exclude some practices, based on their knowledge of key insurability conditions and mutuality principles, the use of big data by the business and inputs in risk models.

Commitment to manage data responsibly with clear accountability is key for building consumer trust.

But it is not a mere matter of volume. **The reliability and quality of external data will also be a key element at stake.** Low reliability and quality of data may lead to erroneous conclusions and non-repeatable findings from Big Data.

To enhance data-driven decision making and offer products more tailored to clients' risk profiles and needs, data oversight and governance will be essential to ensure their reliability. In health insurance for example, capturing unbiased, quality clinical data will ensure that patients are allocated to the right risk segments, guarantee that their access to health programs is fair and that the outcomes of programs can be correctly evaluated. Doing otherwise could lead to patient dissatisfaction and inefficient programs.
CROs should play a leading role in both sides: advising on a company's code of conduct and setting up a strengthened data quality, data management and data governance framework within their organisations.
To successfully address the big data challenges, a data governance framework should encompass the following aspects:

- data sourcing and extraction
- data cleansing
- data quality
- algorithms quality and ethics
- discrimination between data that can legally and ethically be used or not
- results back testing
- data archiving requirements for insurance needs
- data deletion for operational or legal purposes (e.g. "Right to be Forgotten")

The quality of data

Since Big Data & Analytics is about information processing, the quality of the information on which the analyses are performed is key. Although most data quality issues in Big Data & Analytics initiatives are not necessarily different from those pertaining to traditional information processing, Big Data has some specificities. Among these are the challenges stemming from the heterogeneity of data, range of sources, sampling bias (especially when sampling from datasets that are too large to be used as a whole) and use of the so-called data lakes. According to Gartner, data lakes can be broadly defined as enterprise-wide data management platforms to analyse disparate sources of data in their native format.[7] These solutions should eliminate the up front costs of data ingestion and resolve the problem of information silos. Accepting data without oversight or governance, however, carries risks and can make data scientists' findings non repeatable or, worse, incorrect. A basic illustration of not using metadata comes from car telematics. Certain algorithms estimate car accident probability based on driving style, e.g. analysing braking and cornering. If the link with acceleration data does not track device installation methods or vehicle types, for example, it is likely that data scientists reach erroneous conclusions in ranking drivers' riskiness. Plain device rotations among installation batches could cause confusion on which event is braking and which is cornering, while stiff suspensions might be interpreted as harsh driving. The amount of data should not ease caution towards bad bias and spurious correlations, either, nor give the illusion that rare events become easier to investigate, particularly because of error rates associated with searching over large classes of hypotheses. Thinking again of car telematics, the number of factors that can cause accidents is hard to pin down, with causes ranging from weather to distraction and from road conditions to vehicle maintenance. Data quality for Big Data deserves therefore no less consideration than that for traditional data, while data management should extend to the full data lifecycle, involving processes, technology and people responsible for data acquisition.

[7] http://www.gartner.com/newsroom/id/2809117

The CRO checklist for sound second opinions in a world of Big Data & Analytics

- How are the following aspects governed?
  - ◊ data sourcing and extraction
  - ◊ data cleansing
  - ◊ data and algorithm quality
  - ◊ discrimination between data that can legally and ethically be used or not
  - ◊ results back testing
  - ◊ data archiving requirements for insurance needs
  - ◊ data destruction for operational or legal purposes (e.g. "Right to be Forgotten")
- How is the calibration of underlying risk scoring established?
- Who is responsible for the accuracy of codes and algorithms a product pricing relies upon? Does this function have sufficient awareness of its role as a first line of defense? Does the second line of defense have sufficient understanding of the codes and algorithms used in this project?
- Is the pricing for this product dynamic and directly influenced by the policyholder's behaviour? Does the pricing and risk model have a self-learning feature?
- Are external and internal data sets complementary and is this complementarity used for verification and calibration?
- What are the measures in place to mitigate cyber risk, prevent loss of or unauthorized access to data? Is there a process in place to identify potential vulnerabilities in data storage and in interactions with third party vendors? Is there a business continuity plan in place?
- Is there a process in place to ensure that the use made of personal data is ethical and compliant with customer protection and data privacy laws?

Conclusions

The examples and practical use cases on opportunities and risks stemming from Big Data & Analytics show that to make the most of Big Data the following actions should be taken:

- Upgrade data quality and data governance frameworks to avoid pitfalls and negative business implications because the nature and lifecycle of Big Data is different from that of traditional data.
- Create the conditions for data scientists' findings to be incorporated into products and services quickly because Big Data & Analytics solutions pertain to a world where agility and speed matter.
- Discuss specific privacy solutions with Data Protection regulators to prevent ex-ante restrictions that limit innovation because regulation is a key factor in Big Data & Analytics adoption. Also, engage policymakers on Open Data to identify steps for its potential usage for society.
- Look beyond regulation to promote fairness, transparency and accountability and implement robust data security, because customers' distrust has the potential to be an impediment for Big Data-related products and services.
- Source capabilities and plan skills development or acquisition because data engineering and science require specific expertise.
- Define best practices to establish partnerships with start-ups, digital natives and entities from different industries.

Looking ahead, the growing pervasiveness and importance of information will evolve business models further. For example, telecom operators, among others, are investigating data monetization and data brokers are already a reality. Many advocate an informed personal data market. This means that in the future, individuals might start attributing a price to their data. Some may even develop specific economic expectations when releasing it.

The new envisioned economic approach to personal data is not entirely new to insurance. The directionally inverse correlation between information disclosure and premium levels is well known. However, new questions arise when considering how different sectors might interact in a data market. For example, one might wonder what operators in the digital space will attach to data prices that insurers can make sense of. What will customers' expectations be in terms of premium reductions when disclosing information to insurers? For which lines of business and to what extent can insurance pricing mechanisms accommodate the choice of not releasing personal information?

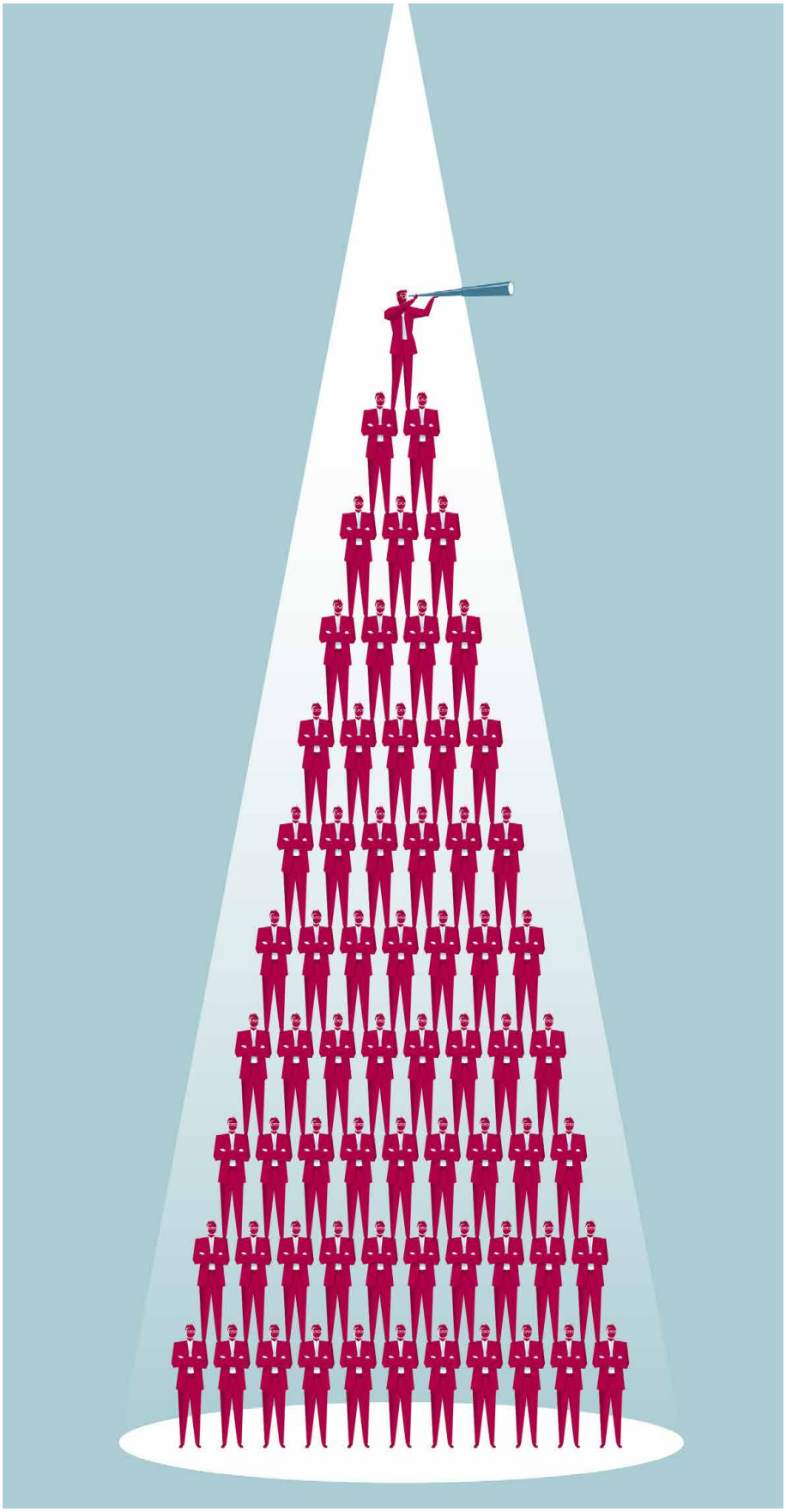
CROs play an important role in helping organisations apply risk management to these and many other questions. Big Data & Analytics integrates cross functional capabilities that aim to strike a balance between growth and related risks, CROs have a role here.

The growing pervasiveness and importance of information will evolve business models further.

In particular, CROs can offer their operational risk management and risk modelling experience to help:

- Define policies on data acquisition and use.
- Design and implement data security and data quality controls.
- Design and validate new analytical models.
- Identify gaps in organisations capabilities and competences.

Big Data & Analytics offers unique opportunities to the insurers who know how to win customers' trust while leveraging their own long-established ability to handle information. More importantly, with Big Data & Analytics the insureds will enjoy more specific prices, customized services and benefit from improved loss prevention.



Big Data & Analytics offers unique opportunities to the insurers who know how to win customers' trust while leveraging their own long-established ability to handle information.

Annex

The definition of Big Data

The definition of Big Data has been debated extensively. In the context of this paper, two definitions from respected sources will suffice. According to Gartner, Big Data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making[1]. This definition is in line with large part of the debate, which focuses on data itself. Another interesting definition is available from Forrester: Big Data is the frontier of a firm's ability to store, process, and access all the data it needs to operate effectively, make decisions, reduce risks, and serve customers[2].This definition is slightly more business oriented and, together with the previous one, emphasizes that Big Data is really two related topics. One is the combination of different information sources and the other is massive data analysis. The former is mainly about the application of established analytical techniques to new data sets arising from data enrichment and data brokerage. These possibilities, associated with legal and regulatory aspects, pertain to most of the use cases typically discussed in business contexts as well as in this paper. Massive data analysis, on the other hand, is about finding interpretations of phenomena underlying truly massive or heterogeneous data. This type of analysis is basically associated with advanced scalable data-centric technologies, statistics, and machine learning.

[1] http://www.gartner.com/it-glossary/big-data

[2] http://blogs.forrester.com/mike_gualtieri/12-12-05-the_pragmatic_definition_of_big_data

References

36th International Conference of Data Protection and Privacy Commisioners, **Resolution Big Data,** 2014

Association of British Insurers, **How Data makes insurance work better for you,** September 2015

Center for Data Innovation, **Open Data in the G8,** March 2015

Communication from the European Commission, **A Digital Single Market Strategy for Europe,** 6 May 2015

CRO Forum, **Cyber resilience: The cyber risk challenge and the role of insurance,** December 2014

Davidson M, Haim DA, Radin JM.(Nature), **Using Networks to Combine Big Data and Traditional Surveillance to Improve Influenza Predictions**, 2015

Enric Junqué de Fortuny, David Martens, and Foster Provost, **Predictive modeling with Big Data,** December 2013

European Data Protection Supervisor, **Preliminary Opinion of the European Data Protection Supervisor,** March 2014

European Data Protection Supervisor, **Europe's big opportunity,** 27 July 2015

Forrester, **http://blogs.forrester.com/mike_gualtieri/12-12-05-the_pragmatic_definition_of_big_data ,** 5 December 2012

Gartner, **Critical Capabilities for Advanced Analytics Platforms,** 24 February 2015

Gartner, **http://www.gartner.com/it-glossary/big-data ,** December 2015

Gartner, **http://www.gartner.com/newsroom/id/2809117 ,** 28 July 2014

GSMA, **The Mobile Economy,** 2015

Insurance Insider, **How smart analytics helps build true knowledge champions,** October 2014

James Guszcza et Al. (Deloitte), **Too Big to Ignore,** 2013

McKinsey&Company, **Big data: The next frontier for innovation, competition, and productivity,** June 2011

McKinsey&Company, **Open data: Unlocking innovation and performance with liquid information,** October 2013

Munich RE, **Business analytics provides structure when dealing with big data,** 2014

SMA, **Big Data in Insurance,** June 2014

The Actuary, **Black swans and red herrings,** 2009

The National Academies, **Frontiers in Massive Data Analysis,** September 2013

Third Nature, **The Challenges of Big Data & Approaches to Data Quality,** 2013

Disclaimer

The CRO Forum is supported by a Secretariat that is run by
KPMG Advisory N.V.

Laan van Langerhuize 1, 1186 DS Amstelveen, or
PO Box 74500, 1070 DB Amsterdam
The Netherlands
www.thecroforum.org