Institute and Faculty of Actuaries

Data Science in Insurance: Opportunities and Risks for Consumers

Policy briefing

Executive Summary

The world is now producing enormous amounts of data, with exponential growth in data generation seen in recent years. Much of this new data is being recorded in new ways, and advances in technology mean that it can be stored and analysed much more quickly than in the past. Against this backdrop, there has been much recent commentary on Big Data and Data Science: the ability to process and draw useful conclusions from larger quantities of data, from variable sources, much faster than ever before.

Data Science is already starting to transform many aspects of modern life, and has significant potential to promote innovation in the insurance industry. Insurers have long gathered data to understand the nature of the risks they collect, and Data Science in conjunction with advances in computing power offers a step-change in risk analysis by being able to see these risks in much more detail, and on a continuous basis. This offers potential benefits for insurers and policyholders alike, with scope for innovation in how products are offered and priced and claims managed. As consumers' expectations of all things digital, speed of response and ways of doing business increase, their expectations of the insurance industry are also likely to increase.

However, as it delves ever deeper into our lives, Data Science raises questions over ethics and the public interest. As insurers are able to see risks in finer detail, the level of cross-subsidy between policyholders could decline. It is also possible that some policyholders could find insurance harder or more expensive to obtain. Data access and consent to use it is also a key concern, as is data security. With personal data being gathered in increasing volumes, there is also a risk that insurers could be perceived as being overly intrusive, and acting in a 'Big Brother' fashion.

Compliance with existing and impending data protection regulation is highly relevant to the application of Data Science, including the General Data Protection Regulation (GDPR), which will apply from May 2018. However, insurers will also need to consider whether their use of Data Science is putting consumers' needs first, and in the general public interest. Given these public interest and regulatory issues, appropriate regulation of professionals working in Data Science will be very important. The Institute and Faculty of Actuaries (IFoA) will ensure that as these public interest issues evolve, the regulation of actuaries remains fit for purpose.

Evolving risks and the future of insurance

Informing the debate

September 2017

Contents

1. Background: the rise of Data Science

Global explosion in data

The world is now producing enormous amounts of data, with exponential growth in data generation seen in recent years. In 2015, it was estimated that no less than 2.5 quintillion bytes of data (i.e. 2,500 million billion bytes) were created every single day. To illustrate the rate of growth, ninety percent of the total data by then in existence was estimated to have been created in the previous two years alone.[1]

We are not just seeing greater volumes of data, but also data being recorded in new ways:

- huge amounts of consumer data is collected via internet search engines such as Google
- the growth in data generated via social media such as Facebook and Twitter
- data collected by mobile devices including smartphones[2], wearable gadgets and telematics.

We are now increasingly living in what is known as the Internet of Things, or 'IoT'. This is the rise in the number of home gadgets connected through networks to each other, all generating data and communicating with us. In the UK, the IoT is gaining traction in home heating systems, making efficient use of energy within 'smart' homes. Technology is also transforming fitness regimes with the increasing popularity of wearable devices that monitor exercise regimes and provide data on wellbeing. Insurers are already seeing opportunities from these particular examples applied to home and health insurance.

Advances in technology/ computing power also mean that data can be stored, transformed and analysed much more quickly and cheaply than in the past. New technology can now handle both structured data - which can be read/ understood by humans and computers - and unstructured data - such as images, tweets and text, which cannot be handled by conventional databases. Cloud technology is a further example of the breakthrough in cheap scalable computing power that can be applied to handle these analytical investigations.

Big Data and Data Science

Against the backdrop of this global explosion in data generation, there has been much commentary on 'Big Data' and Data Science. Big Data relates to datasets too large to be handled by conventional data processing software. There is no single definition of Big Data, but it can be thought of as relating to the 'five Vs'; the volume, variety, velocity, veracity and value (extracted) from data:

- volume: the vast amounts of data now being generated and gathered
- variety: many different types of data brought together from different sources; with it being unstructured/ structured, static/ real-time and in-house/ external
- velocity: the increasing speed at which new data is processed and stored, making analysis in real-time possible
- veracity: the need for data to be relevant, accurate, appropriate and consistent
- value extracted: the value from the additional insight in seeing people/ objects/ processes in far greater detail.

Data Science then relates not only to the collection and use of Big Data, but also to the subsequent analytical tools, methodology and technology used to benefit from the extra insight offered by the corresponding data sets.

2. Data Science in insurance: benefiting from the digital experience

Data Science offers insurers a step-change in risk analysis by being able to see their risk exposures in much greater detail

Data Science: potential for innovation in insurance

Many aspects of modern life are already being transformed by Data Science - health care, science and research, politics and even sport. Moreover, Data Science has the potential to promote substantial innovation in the insurance industry, including how products are offered, and to whom. Risk management is central to insurance, and to understand the nature of the risks they underwrite, insurers are also gatherers of data. Data Science offers insurers a step-change in risk analysis by being able to see risks of the things they insure - such as vehicles, people or property - in much more detail than before. It also offers the ability to identify relationships which would otherwise have remained hidden within data, or to make better predictions based on continual updating of far larger data sets.

Actuaries are already applying Data Science to a number of insurance developments, including telematics devices in motor insurance; wearable fitness devices in health and care insurance; and advanced risk management (more detailed analysis of experience analysis) in life insurance. *Please see the Case Study for more background on telematics devices.*

As well as benefiting insurance companies, Data Science offers benefits to consumers at each stage of the insurance policy lifecycle, in terms of better:

- consumer targeting and product design
- risk assessment, underwriting and pricing
- consumer engagement during cover and
- claims management, including identification and avoidance of fraud.

We consider each of these in turn:

Better consumer targeting/ product design

Insurers can take advantage of new sources of data to better target intended customers to specific, and potentially more suitable, products. Analysis of internet search histories or of social media content helps predict consumer preferences and behaviours. In turn, these analyses can be used to match consumers to specific products. Similar analysis of trends in preferences and behaviours also improves an insurer's understanding of changes in consumer needs, and this insight can be useful in the development of innovative new products and design of associated features.

The increasing use of smart Apps by society at large and consequent analysis of their use by insurers could also make it easier for some consumers groups to access insurance, including meeting untapped areas of demand for insurance.

A more rounded view of consumers and their needs could also potentially mean that they are not missing out on necessary insurance cover, or indeed doubling-up on cover elsewhere.

More accurate risk assessment, underwriting and pricing

Much of Data Science's potential in the insurance world relates to the greater insight possible in the risk assessment process. It has long been the established practice for insurers to gather data on applicants' (or their property's) characteristics and use this to assess the likely chance and cost of claims. From that, an underwriting decision would be made: either to accept the risk on standard terms, to accept on modified terms, or to decline.

An appropriate premium based on the risk (and underwriting decision) would then be set. However, Data Science allows insurers to see their applicants' risk profiles in much greater detail. This more granular risk assessment should mean better-informed underwriting decisions, and allow insurers to set premiums more accurately, in greater alignment with the corresponding level of risk.

More accurate and detailed risk assessment could increase insurance coverage or make it cheaper in some cases, such as where a risk would be declined under a broad assessment, but deemed acceptable in the light of greater information. For example, greater clarity on the underlying level of risk could see younger drivers individually allocated to higher and lower risk segments, with the latter potentially able to access cheaper motor insurance. Some consumers with disabilities could also potentially benefit from more detailed risk assessment.

Data Science can also reduce the effort involved in obtaining an insurance quote, if much of the information required for underwriting were sourced directly, thus reducing the number of direct questions the consumer would need to answer. More accurate underwriting can also mean less fraudulent disclosure, with the potential for insurers to reduce the risk that applicants take advantage of information the insurer does not have.

Stronger engagement with consumers

Consumer engagement could also improve over the course of an insurance policy, benefiting policyholder and insurer alike. There is potential to reward policyholders with lower premiums if their risk profile improves, such as via a telematics device in the case of motor insurance, or a wearable fitness device in a health context. Furthermore, the prize of lower premiums could encourage policyholders to reduce the 'riskiness' of their lifestyle. By incentivising insurance policyholders in this way, Data Science offers clear benefits to society. Consumer engagement is also being taken to the extent of 'on demand' insurance: at least one motor insurance provider is offering the facility to switch cover on and off via a smartphone.

Better claims management

Data Science has a role to play in making insurance claim management and associated complaints processes more efficient, benefiting insurer and policyholder alike. Data analytics can be used to prioritise claims: with straightforward claims fast-tracked for quick settlement, and more complex cases flagged for further analysis by claims assessors. Analysis of social media activity and connections can also be used effectively to spot fraudulent claims activity by groups of people working together to make a series of false or exaggerated claims.

As consumers' expectations of all things digital, speed of response and ways of doing business increase, their expectations of the insurance industry are also likely to increase, from both a claims management perspective, and more generally.

Finally, emerging high technology 'Insure tech' firms are potential disruptors to the insurance market. Such Insure tech firms can be quite agile in their business model, and often have a very data-driven mindset and digital approach.

3. Public interest concerns

As Data Science delves ever deeper into the insurance world, this raises questions over ethics and the general public interest. Insurers and consumers could be impacted in new and unexpected ways, with potential unintended consequences and possible distortions in the insurance market.

Insurance unavailable for some?

Certain subsets of the general public could find that Data Science has an adverse impact on the cost or availability of insurance. If insurers have a clearer understanding of an individual's risk characteristics, then individuals in certain market segments may then find that insurance is harder or more expensive to obtain. For example:

young drivers with little driving experience accessing motor insurance

people in declining health, finding health insurance becomes too expensive

older travellers requiring travel insurance, where current factors used to set insurance premiums are generally broad-brush rather than precise in nature.

This then gives rise to what is essentially a question of public policy, which will need to be addressed. Charging a premium aligned to the individual's risk might be considered consistent with treating customers fairly. However, individuals or their property could be 'higher risk' due to inherent factors, or changes in circumstance, over which they have no influence. If this lack of availability of cover at a reasonable (meaning subsidised in this case) price is deemed to be a market failure then avoiding this would require some degree of intervention from Government, regulators or the insurance industry itself.

Less pooling of risk

Data Science could also have an impact on the related concept of the degree of pooling of risk. This risk pooling - where risks are shared between policyholders with broadly similar risk characteristics - is a long established feature of insurance. It is likely that the rise of Data Science will steadily reduce the size of each pool, and hence lower existing levels of cross-subsidy between different policyholders. There is already a trend in the UK towards more detailed risk assessment within insurance, such as the emergence in the mid-1990s of impaired life and enhanced annuities, which offer a higher income to smokers or others with chronic/ terminal health conditions. Data Science could see this trend continue, with risk segmentation becoming increasingly fine and less pooling of risk possible within ever smaller segments.

However, society may desire a degree of cross-subsidy by retaining pooling of risk where this is deemed to be 'fair'. Across the EU, gender-specific pricing for insurance was banned in 2012 on the grounds of sex discrimination. In the UK, the emergence of Flood Re* recognises that as flood risk rises, both those at greatest risk and the industry need time to adjust; in this case, a temporary subsidy from policyholders with lower risk of their homes flooding to those with a higher risk. To the extent that Data Science further reduces the pooling of risk in the insurance market, further debate may be necessary to determine the degree of pooling society deems appropriate and what actions are required by Government, regulators or the insurance industry as a consequence.

Flood Re is a flood reinsurance scheme to help flood insurance cover be affordable to those households at highest risk of flooding.

Price discrimination

As the amount of data collected on individuals increases, the extent of price discrimination by insurers could also potentially increase. Price discrimination refers to basing insurance premiums not only on the individual's level of risk and associated costs, but also on wider factors including their sensitivity to price and brand loyalty. It is by no means new, but greater scope to measure such sensitivities increases the risk that vulnerable consumer groups could be impacted adversely by such practices. Insurers would need to consider whether such price discrimination gives rise to any conduct risk concerns.

Data: property of insurer or individual?

A further question relates to who owns the data being gathered on a policyholder. Does the relevant insurer own it, or the policyholder? If insight gathered via a telematics or wearables device were the property of the insurer, this could potentially restrict the consumer's ability to access a better deal elsewhere. Conversely, if it were the property of the consumer, could they then be obliged to share pre-existing data with an insurer if they changed provider?

Transparency and judgement

There is a potential lack of transparency around Data Science and associated analytics, and understanding what is behind the analytical 'black box' can be challenging. Information asymmetries between insurers and consumers could widen with greater use of Data Science. There is also a risk of over-reliance on analytical model outputs: no model can replace human knowledge and judgement entirely.

Human judgement also brings with it further public interest concerns. With Data Science becoming increasingly intrusive, there a risk of prejudice (or bias) creeping into any risk profiling applied. Some advanced pricing models may rely on algorithms rather than explicit rating factors, making it difficult to establish whether price discrimination is being applied. Conversely, it could be argued that greater clarity on the risk could reduce the reliance on analysts' prejudices by making decisions more data-led.

As insurers gather increasing volumes of personal data, it is possible that they could be perceived as being overly-intrusive and acting in a 'Big Brother' fashion. It is therefore important that insurers are transparent in their use of data, to avoid diminishing the level of trust policyholders place in them, and in turn, reduce the scope of data policyholders are happy for the insurer to access. However, if insurers explain the tangible benefits to the policyholder of this data collection - such as a potential reduction in insurance premiums through the use of a telematics device - they may then decide that this outweighs any associated privacy concerns.

Cyber risk

Finally, cyber risk is an important emerging area of risk, and Data Science contributes to insurers' exposure to it as more and more data is accumulated. In particular, the risks of data being lost, corrupted or stolen are important issues for users of Data Science applications to consider. Processes will need to be developed with appropriate controls against hacking and other unauthorised users accessing this data.

4. Regulation

Data protection

Compliance with existing and impending data protection legislation is highly relevant to Data Science. The General Data Protection Regulation (GDPR) will apply in the UK from 25 May 2018; previous data protection legislation in the UK dates from the 1990s, and much has changed in the intervening years. We also note that the UK Government has confirmed that the UK's decision to leave the EU will not affect the commencement of the GDPR in the UK. To the extent that insurers operate on a global basis however, they will have the additional challenge of complying with multiple data protection regimes.

The principle of consent to the use of data is key, but users will need to be clear on the boundaries of data that have permission to be used. For example, where data is gained from social media activity and applied to insurance underwriting or marketing purposes, have policyholders and social media platforms necessarily given their consent to this? Data captured should also be relevant to the purpose for which it is being used. As mentioned above, there is a further question over ownership and access.

Insurers will therefore need to have robust data governance processes together with suitable controls to ensure that relevant processing meets current and evolving data privacy regulation.

Data Science also raises potential conduct risk concerns for insurers. They will need to consider whether product design and pricing, and wider applications of Data Science, are putting consumers' needs first, and are in the general public interest. Failure to do so could risk regulatory intervention to address conflicts of interest, or to improve consumer outcomes.

Regulating actuaries in the public interest

Given the potential ethical and wider public interest issues arising from the increasing use of Data Science, it is important to consider the regulation of professionals working in this field, be they actuaries, data scientists, risk managers or otherwise.

The IFoA regulates our members to ensure the public interest, whilst supporting business and innovation. Ensuring the public interest relates to protection of the wider public and society as a whole, but also extends to ensuring public confidence in the actuarial profession by maintaining its good reputation. The public interest is served by actuaries adhering to high-quality standards of work and also by the ethics and professionalism of the actuaries who deliver it.

As Data Science grows in importance, the IFoA will assess its regulatory framework to ensure that as public interest issues evolve, the regulation (and education) of members remains fit for purpose, without being overly burdensome.

The principle of consent to the use of data is key, but users will need to be clear on the boundaries of data that have permission to be used.

5. Conclusion

As set out above (and in the case study which follows), Data Science is already starting to transform the insurance industry, offering clear benefits to insurers, policyholders and the wider public. However, increasing use of Data Science gives rise to a range of potential public interest issues for the insurance industry, regulators, Government and other relevant stakeholders to consider:

- certain subsets of the general public could find that Data Science has an adverse impact on the cost or availability of insurance
  - when should interventions be made to keep broader pools of risk and hence maintain or increase levels of cross-subsidy between different policyholders?
  - does differentiating by factors other than inherent risk give rise to conduct risk concerns?

- what are the rights but also obligations of policyholders with respect to data being gathered on them?

- how should insurers ensure transparency to maintain policyholders' trust in the use of their data?

- how do participants in the insurance market ensure compliance with current and evolving data protection regulation, whilst gathering and using greater amounts of data?

- more generally, are applications of Data Science in insurance putting consumers' needs first, or are there wider conduct risk concerns?
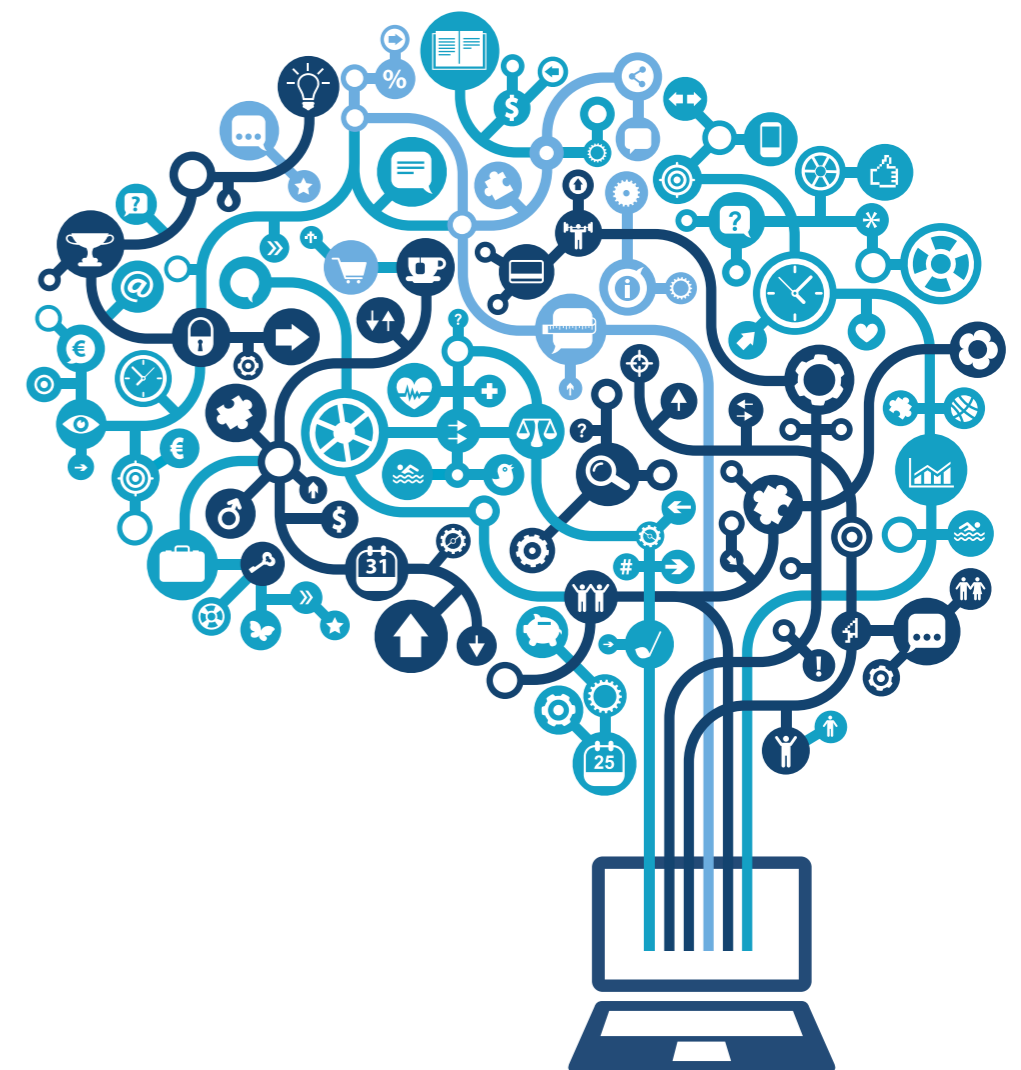
Case study: telematics devices

Insurer   Driver   Benefits
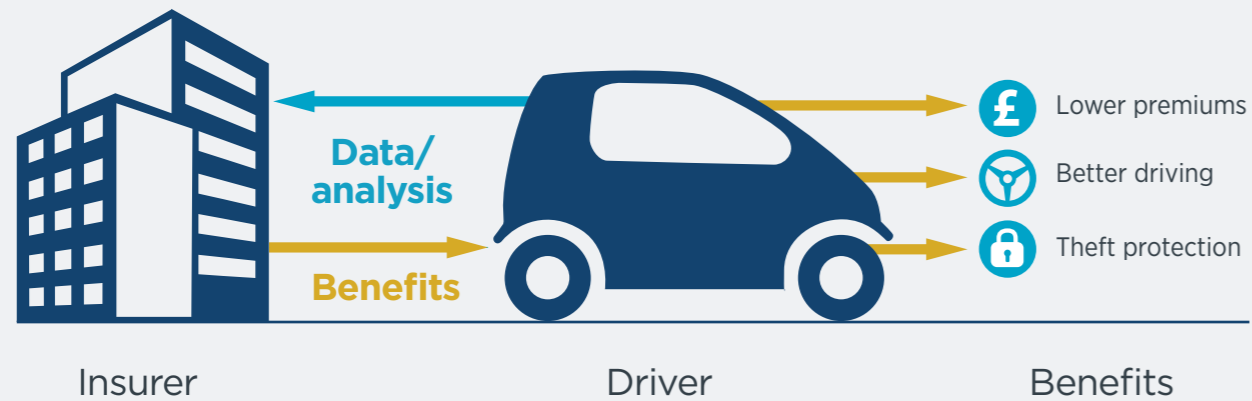
Motor insurers and their policyholders alike are benefiting from the use of Data Science in the form of telematics. A telematics device – often referred to as a 'black box' – is fitted to the insured vehicle, and monitors how, where and when it is driven. For example, devices may record the vehicle's speed, acceleration, rate of braking and other forces such as the impact of an accident. Driving habits can also be inferred through data recorded on when the vehicle is driven, length of single journeys and when the driver takes a break. All of this data measured is relayed back to the insurer on a real-time basis.

Telematics devices are already well established in the North American market. They are also now gaining ground in the UK, where there are now 450,000 devices used in motor insurance, with 40% growth between 2015/16 [3].

Insight from the black box: refined premium rating

By receiving data on how the vehicle is driven and the corresponding mileage undertaken, the insurer is able to see the insured driver in fine detail. The insurer can therefore assess the risk based on the driver's actual driving behaviour, rather than estimating the likelihood of claim by allocating the driver to a broad risk category. By making a more accurate assessment of the risk posed by a given driver, underwriting decisions and premiums charged will correspond more closely to the underlying risk than would be the case using traditional rating methods. Younger drivers with good driving habits have benefited

from this: either by obtaining cheaper insurance (by not being allocated to a broad pool of inexperienced and high risk drivers) or simply being able to obtain insurance, where they previously struggled to get it.

Insurers and drivers also benefit from the real-time relay of data from the telematics device back to the insurer. Continuous updating on driving patterns and the mileage driven means that underwriting and premium setting can be dynamic in nature. This gives drivers the incentive of premium discounts as a reward for improving their driving behaviour, or at least maintaining good driving techniques to avoid an increase in premium. The process can be transparent, where drivers are able to see the updated analysis of their own driving online and how this compares to other groups of drivers. Telematics therefore also promotes safer driving as well as the monetary benefits.

Wider benefits from location tracking

Telematics devices offer a range of wider benefits to insurers and drivers alike. Continuous data on the vehicle location and associated timing is clearly useful in the case of theft, and such information can also assist with vehicle recovery in the event of an accident, and streamline the claims process. As well as being able to pinpoint location, data measuring the impact on the vehicle in a collision can alert the insurer to the potential occurrence of an accident allowing the emergency services to be called, and provide diagnostics on the cause - including insight on where the blame may lie. In such circumstances the

driver would benefit from support at the time of accident/ during the claims process, but there are also advantages to the insurer in more efficient claims assessment and processing. Such opportunities take consumer engagement and claims management to a new level.

Combating fraud

Insurers are also using telematics to combat fraudulent claims. Comparison of claim details submitted by the insured against the detail recorded by the telematics device has flagged discrepancies, and ultimately led to fraudulent activity being uncovered. This can lead to the identification of systematic fraud across groups of what otherwise appear to be unrelated claims, and the prosecution of the people involved.

Switching providers: rights and obligations

Although telematics offer consumers a wide range of benefits, there are potential public interest concerns. The question over data ownership and obligations is relevant here. As telematics grow in popularity, there is a risk that data gathered from such devices may not be transferable easily, particularly if different insurers use incompatible devices, or regard the data as their own property. Lack of

transferability could see policyholders reluctant to switch provider, or face higher premiums if their data cannot be transferred easily. Conversely, where it can be transferred, and in the context of compulsory motor insurance, would it be in the public interest for consumers to be obliged to share pre-existing data on their driving with new insurers, when they changed provider?

Intrusion to privacy: a price worth paying?

Many drivers might be put off by the continual surveillance nature of a 'black box' in their car, seeing it as intrusion into their privacy. However, insurers marketing telematics devices explain the range of benefits to the driver, including promotion of safer driving, the prospect of discounted premiums now and in the future, and the potential enhanced insurer support in the event of theft or accident. Where the driver can appreciate the tangible benefits of a telematics device, they may then decide that this outweighs any associated data privacy concerns.

References

[1] http://www.vcloudnews.com/every-day-big-data-statistics-2-5-quintillion-bytes-of-data-created-daily/

[2] https://techcrunch.com/2015/06/02/6-1b-smartphone-users-globally-by-2020

[3] https://www.biba.org.uk/press-releases/40-increase-in-telematics-motor-policies-in-a-year/

Institute
and Faculty
of Actuaries

Beijing

14F China World Office 1 · 1 Jianwai Avenue · Beijing · China 100004
Tel: +86 (10) 6535 0248

Edinburgh

Level 2 · Exchange Crescent · 7 Conference Square · Edinburgh · EH3 8RA
Tel: +44 (0) 131 240 1300 · **Fax:** +44 (0) 131 240 1313

Hong Kong

1803 Tower One · Lippo Centre · 89 Queensway · Hong Kong
Tel: +852 2147 9418

London (registered office)

7th Floor · Holborn Gate · 326-330 High Holborn · London · WC1V 7PP
Tel: +44 (0) 20 7632 2100 · **Fax:** +44 (0) 20 7632 2111

Oxford

1st Floor · Park Central · 40/41 Park End Street · Oxford · OX1 1JD
Tel: +44 (0) 1865 268 200 · **Fax:** +44 (0) 1865 268 211

Singapore

163 Tras Street · #07-05 Lian Huat Building · Singapore 079024
Tel: +65 6906 0889

www.actuaries.org.uk