

Q&A on the use of big data in insurance



This Q&A document aims to respond to the most commonly asked questions about the use of big data in insurance. The document concludes with the insurance industry's view on how policymakers and supervisors can support innovation in this area for the benefit of consumers and insurers.

The importance of data in the insurance business model

Data has always been a key factor for European insurers. In fact, even long before the emergence of the big data phenomenon, insurers made use of data mining techniques, in compliance with the relevant regulatory frameworks.

To provide reliable insurance cover, insurers must carry out sophisticated risk assessments and calculations, using various types of information. In particular, insurers analyse past events with statistical methods to estimate the probability of these events occurring. This data analysis is carried out at the product design stage, allowing insurers to learn and manage the risks of offering a new insurance policy. At a later phase, at the sales process, the insurer runs a data analysis process known as underwriting, where the risk that the new customer brings is assessed. With the results of the underwriting process, insurers can offer the customer the policy terms and insurance cover, in line with the risk represented by the new customer.

Furthermore, the underwriting process involves calculating the probability of risks occurring and the consequences for each insured or category of insureds. The premiums charged by insurers are based on these calculations. Inevitably some uncertainty will remain regarding expected losses; for instance, it is common to have variations in claims costs at different times. Therefore, the premium will also include an additional margin to enable the insurer to build up a reserve to draw on in bad years.

Today, new data mining techniques allow European insurers to pass on more benefits to consumers. Specifically, due to the increased availability of data, today's data mining techniques allow insurers to carry out more accurate risk assessments, meaning insurance products can be better tailored to each consumer's risks and needs. Developing new, or more sophisticated, risk models can enable insurers to offer more competitive rates, or to offer insurance for risks that were previously uninsurable, due to information gaps which today are filled in by the increased availability of data. In this regard, digitalisation makes additional data, as well as new analytical methods, available.

What is big data and what role does data mining have in insurance?

Big data is often defined as "high-volume, velocity and variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making"¹.

Technology experts coined the term big data in 2000 in relation to the information and data explosion driven by ever-growing internet access and the invention of digital, or cloud, storage.

¹ Doug Laney, "3-D Data Management: Controlling Data Volume, Velocity, and Variety", 6/2/2001, Meta Group

In 2010, it was said that: “Every two days we create as much information as we did from the dawn of civilization up until 2003”². The pace and volume of data created has increased dramatically since then and it will continue to do so: it is forecast that “by 2025 the global datasphere will grow to 163 zettabytes (a zettabyte is a trillion gigabytes)”³, which is ten times the 16.1 zettabytes of data generated in 2016.

Put simply, the enormous, unstructured sets of data collected from widely diverse sources form what we know as big data. While big data can be viewed as a raw material, which on its own has no practical value, data mining enables analysts to extract understandable and structured data that can be valuable. In this respect, insurers can use the processing of big data in their data analysing processes.

The extraction of patterns from data, or data mining, has taken place for centuries. This was originally done through manual extraction with the application of mathematical and statistical processes. However, data mining techniques are now becoming ever more sophisticated with the emergence of the internet, the increased availability of data and technological advances.

Today, data mining often involves the use of specialised software, such as analytical tools. These tools can be used, for instance, to quickly sort through decades of accounting information to find a specific set of expenses or accounts receivable for a specific operating period⁴.



What are the benefits of using big data in insurance for consumers?

The use of big data analytics will allow consumers to benefit from improved products and services, because insurers will be able to better analyse whether products worked as intended, reached the right customers and whether further development is needed. Also, the use of big data tools will enable insurers to more efficiently detect cases of fraud, by cross-matching data from different databases (eg tax authorities’ data, credit card information, etc.), and to better advise consumers on how to prevent incidents from occurring.

Big data analytics, together with the increased volume of better data, could also help insurers to cover risks that were previously uninsurable. For instance, looking at past advances in risk assessment, more and better data as well as knowledge about risk determining factors have always increased the possibilities to provide insurance cover. In this regard, the ongoing development of big data analytics and related technological developments could help insurers to fill in information gaps in risk assessments, leading to better overall insurability as well as a fairer premium calculation. Evidence of this can be seen in the examples listed below. Furthermore, accuracy in pricing not only means that low-risk policyholders pay less for their premiums. By using big data, insurers can reduce the additional safety margins they need to include in their pricing, which are often larger in high risk profiles.

The use of big data will allow insurers to better advise consumers on the best measures to adopt to prevent risks from occurring and how to mitigate their impact. Crucially, improved prevention is likely to have a positive impact on premium pricing.

Finally, the use of big data will also enhance fraud prevention systems. This will benefit both consumers and insurers alike, as an increased fraud detection rate will decrease unjustified payoffs and decrease individual premiums.

How will consumers benefit from the use of big data in insurance?

- **Insurers will be able to tailor their products and services to match consumers’ priorities**

Insurers are exploring how their products and services can be improved to meet consumers’ needs and preferences. For example, pay as you drive (PAYD) motor insurance policies are one of the first tailored products to be introduced in some European markets using big data analytics. The use of telematics with real-time transmission of data allows insurers to discount policies based on driving behaviour, providing an additional option for customers, while encouraging good driving behaviour.

Regarding property insurance, new technology means policyholders can benefit from more tailored products, as well as increased risk awareness and risk-reduction services. These can also have a positive impact on premiums.

² Eric Schmidt, Google’s former executive chairman, statement at the Techonomy conference in Lake Tahoe, CA, 4/8/2010

³ IDC white paper: “Data Age 2025: The Evolution of Data to Life-Critical. Don’t Focus on Big Data. Focus on the Data That’s Big”, April 2017 ([link](#))

⁴ Techopedia: “What is the difference between Big Data and data mining?”, 3/7/2018 ([link](#))

- **The use of big data will allow insurers to offer products and services based on real-time risk assessment**

An increased volume of data helps insurers to fill in information gaps. For example, before the use of big data, the pricing of the premium would only be reviewed at the time of renewing the contract. With the use of big data, insurers will be able to adapt insurance premiums in real time resulting in immediate savings for the customer (eg, PAYD motor insurance policies).

- **Big data allows insurers to better cover high risk profiles, increasing access to insurance**

Today, internet platforms use big data analytical tools to negotiate premium discounts for groups of high-risk profiles, which would have found it difficult in the past to find the desired insurance cover at affordable prices. Such platforms provide, for example, access to travel insurance for cancer patients or insurance products to cover high risk sports.

The German insurance market has developed a highly sophisticated system of flood risk classification that allows for almost universal property insurability even in high risk areas. The inclusion of additional data in 2017 further increased insurability and affordability of flood insurance in Germany. For example, in 2002, 10% of all houses were presumed uninsurable, whereas today more than 99% of all houses can access insurance cover.

Moreover, the increasing availability of data together with medical progress has made it possible, under certain conditions, to provide life insurance cover to individuals with HIV.

- **Improved customer satisfaction**

Big data allows insurers to have a better and faster understanding of customers and their needs. This results in more efficient and less burdensome processes for customers, who will no longer need to fill out repetitive questionnaires.

- **Prevention policies**

Insurers will be able to advise their customers on how to prepare and protect property and valuables:

- Property insurance: Insurers can use big data analytics to advise consumers on the type of prevention measures needed to make properties insurable. This is particularly true regarding measures against the risk of flooding or other weather-related disasters.
- Health insurance: When an individual has volunteered to be monitored, insurers can use big data analytics to monitor their health and provide them with lifestyle tips and health advice. As a result, consumers become more aware of preventive measures they can take to reduce the risks associated with chronic diseases and control medical costs.
- Motor insurance: Insurers could use big data analytics to monitor the driving activity of customers and provide them with advice on how to improve their driving, and also to provide them with information on fuel consumption and tips on how to reduce risks.



What is the impact of big data on the insurance business model and risk sharing?

Big data's impact on the insurance business model

The use of big data analytics throughout the insurance value chain (ie, product development, distribution, customer service or claims handling) is still at an early stage. For instance, in several national insurance markets, the use of big data analytics is limited to marketing activities. Moreover, to be able to maximise big data's potential, insurers need to further develop their risk models to be able to perform the advanced level of data analytics that a valuable use of big data requires⁵.

Currently, insurers are exploring the possibilities offered by new technologies to improve the insurance business model to benefit consumers. The characteristics of big data — volume, variety, velocity, veracity and value — allow insurers to obtain more detailed information about their customers. This provides insurers with a better insight into their customers' preferences and needs, resulting in improved products and services⁶.

⁵ Idem

⁶ The European Actuary: "Big Data is coming, are you ready? Insurance principle and actuaries in the age of fintech", by Esko Kivisaari, N. 15, October 2017

Big data's impact on the risk-sharing principle

Some believe that the use of big data analytics in insurance will result in an undesired increase of granularity in risk pooling, which could undermine the risk-sharing principle. However, introducing additional risk categories and calculating more risk-oriented premiums does not mean that risk-sharing is limited to smaller groups. The community of insureds is usually divided into several sub-groups with different levels of risk and premiums, but the risks are still pooled across the entire portfolio of policyholders and across all risk categories. This is the nature of insurance: to pool risk so that individuals share the risks, and no-one has to bear the entire economic loss on their own. As long as there is a risk affecting a group of individuals, insurers will have to pool risk, which will be shared through a premium representing the risk that each individual brings to the pool.

Additionally, there is often a misunderstanding about how innovative insurance products are used and designed. For example, a common misconception is to believe that usage-based insurance (UBI) policies, such as PAYD motor insurance policies, are solely based on a dynamic behaviour risk model. In other words, it is often believed that these new products are based entirely on the use of big data analytics. In reality, PAYD products are based on traditional underwriting processes and on the additional data provided through telematic devices, which do not necessarily function on big data analytics.



Is the use of big data in insurance regulated?

While there is no specific regulation on big data, there are already a number of rules at EU level that are relevant and applicable to its use:

- The [EU General Data Protection Regulation \(GDPR\)](#)⁷ allows insurers and consumers to be well prepared for the big data environment. The GDPR has created a well-balanced legal framework for processing data. It provides insurers with the right level of guidance, allowing them to mitigate the potential risks brought by the use of big data. At the same time, consumers can now rely on strengthened and new rights to protect their personal data. Moreover, the GDPR addresses the fundamental issue of transparency in the use of personal data, providing a comprehensive system of information disclosure and effective protection. Additionally, under the GDPR, consumers have the right not to be subject to a decision solely based on automated processing, leaving consumers well prepared for the further development of automated decisions by insurers. Furthermore, while insurers use personal data which falls under the GDPR, a significant amount of data used in the insurance business is anonymised, and as such does not affect individuals' privacy.
- The [Packaged Retail and Insurance-based Investment Products \(PRIIPs\) Regulation](#)⁸ imposes the provision of a standardised disclosure format — the key information documents (KID) before a retail investor purchases a PRIIP — allowing consumers to compare the characteristics of different offers.
- The [Insurance Distribution Directive \(IDD\)](#)⁹ regulates the distribution of all types of insurance products by all types of distributors, preventing any poor selling practices that the use of big data analytics in insurance could facilitate. Moreover, its provisions on product oversight and governance (POG), along with its delegated Regulation¹⁰, regulate the design of new insurance products. These requirements aim to protect customers from an early stage in the insurance process.
- The [Distance Marketing Directive for financial services \(DMD\)](#)¹¹ protects consumers from unsolicited products.
- The [EU Gender Directive](#)¹² prohibits the differentiation of insurance premiums by gender.
- The [proposed e-Privacy Regulation](#)¹³, currently under debate at EU level, will bring an additional layer of protection by guaranteeing the confidentiality of communications and shelter consumers from online tracking and unsolicited commercial communications.
- The [Solvency II Directive](#)¹⁴ means insurers have an effective system of governance that provides for sound and prudent management of their business (article 41). Therefore, in order to comply with prudential regulations for risk management, insurers have to base their pricing on reliable data.

⁷ General Data Protection Regulation 2016/679 of the European Parliament and of the Council, 27/4/2016

⁸ Key information documents for packaged retail and insurance-based investment products. (PRIIPs) Regulation 1286/2014 of the European Parliament and of the Council, 26/11/2014

⁹ Insurance Distribution Directive 2016/97 of the European Parliament and of the Council, 20/1/2016

¹⁰ Commission Delegated Regulation (EU) 2017/2358 of 21 September 2017 supplementing Directive (EU) 2016/97 of the European Parliament and of the Council with regard to product oversight and governance requirements for insurance undertakings and insurance distributors

¹¹ Distance Marketing of Consumer Financial Services Directive 2002/65/EC of the European Parliament and of the Council, 23/9/2002

¹² Council Directive implementing the principle of equal treatment between men and women in the access to and supply of goods and services 2004/113/EC, 13/12/2004

¹³ Proposal for a regulation concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)

¹⁴ Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance

Regulation relating to the use of big data: GDPR, IDD, DMD & PRIIPs case studies

■ Data minimisation and purpose limitation principles

The GDPR establishes a solid system to protect individuals' privacy, personal data and freedom of choice in an internet and big data environment. This includes two of the core principles behind the GDPR: data minimisation and purpose limitation.

- The data minimisation principle requires controllers to limit data processing to what is strictly needed for the specific service. Collecting personal data that is not strictly needed would constitute a breach of the data minimisation principle. The data minimisation principle, therefore, guarantees the collection of data strictly related to the purposes of the processing while providing individuals with control over their personal data.
- Purpose limitation means that an insurer must specify the purpose of the data collection, which must be clearly and specifically identified. Accordingly, if a controller asks the data subject to consent in the processing for multiple purposes, consent must be granular in a way that the data subjects can consent in the processing for one purpose, but reject processing for an other purpose. Moreover, any further processing of the provided data has to be compatible with the original purpose for which the data was collected.

These principles together with the transparency rules (ie, requirements for consumer information disclosure) and the efficient and strong enforcement system provided in the GDPR, prevent insurers from being unduly intrusive in consumers' personal lives or putting their privacy at risk.

■ Consent and the right to withdraw it

Consent under the GDPR means offering real choice to consumers and requires a positive action to opt in to a given service.

To this end, the GDPR imposes strict rules for obtaining consumers' consent in order to process their data. One of the main features of consent is that it should be "informed". That means that insurers have to provide consumers with a list of information to enable them to make informed decisions and understand what they are consenting to. This information includes the purposes of the processing as well as, where relevant, the information about the logic, significance and consequences of decisions that have been taken by solely automated means.

Additionally, insurers must provide this information in clear and understandable language to allow consumers to genuinely understand the implication of their decisions.

Moreover, positive action to opt in to a given service means that pre-ticked boxes are prohibited when obtaining consumers' consent.

Another important aspect that grants consumers real choice is their right to withdraw their consent at any given moment. Withdrawal of the consent means that the insurer will no longer be able to process the data of that consumer. Most importantly, withdrawal of consent should not lead to any negative consequences for the consumer, when it was given for processing data that is not necessary for the provision of the service.

As a result, insurers provide all the necessary information to consumers on the impact of providing consent to the processing of their personal data.

What is the role played by insurers' telematic devices?

Current insurance policies linked to telematics (eg, PAYD motor insurance policies) require the installation of a device in the insured's car or the installation of a mobile application. The insurer not only informs the consumer about the legal requirements in the GDPR, but also about the technical requirements of the telematics device. Consequently, the consumer is aware of the purposes for which their data is being processed, their rights, and of the functioning of the device or the application installed.

■ Targeted marketing

Insurers may be able to process consumers' personal data for direct marketing purposes. However, the GDPR establishes a very stringent regime for processing personal data for marketing purposes: insurers must prove that such processing serves a legitimate interest that does not override the rights and freedoms of consumers. Moreover, consumers have the right to object to the processing of their personal data for direct marketing purposes at any time. Furthermore, this right to object should be explicitly brought to the attention of the data subject and presented clearly and separately from any other information.

Finally, the DMD bans abusive marketing practices seeking to urge consumers to purchase services or products they have not solicited.

The European legislative framework prevents situations where the consumer's data could be shared or sold to third parties, leaving consumers to confront an unexpected amount of advertising of unsolicited products.

The e-Privacy Regulation proposal also seeks to add an additional layer of regulation concerning the collection of consumers' online data using online tracking tools or cookies. The proposal would notably regulate the information gathered online for profiling and targeted advertising.

■ **Algorithm transparency**

Private insurance cover requires an assessment of the covered risks to be carried out in determining a risk-based insurance premium. The insurance industry has always used algorithms for this purpose.

Prior to processing, the insurer has to provide transparent, intelligible and easily accessible information about all the personal data being processed. Moreover, the consumer has to be informed of all their rights in a clear and transparent way.

This right to be informed includes meaningful general information about the logic used by the algorithm to take a decision and explanations on how the algorithm works and on the plausible consequences of the processing affecting the data subject. Accordingly, consumers can gain an insight into the logic followed by the algorithm and therefore better understand the outcome of the automated decision.

Additionally, data subjects unsatisfied with the outcome of an automated decision can challenge the outcome by requesting human intervention. Consequently, consumers are well protected under the GDPR when their data is being processed automatically.

■ **The comparability of insurance products**

Despite increased individualisation, insurance products will remain comparable in a big data environment because insurers have to provide consumers with pre-contractual information documents, including standardised schemes and structures, such as the KID for PRIIPs and the IDD insurance product information document (IPID) for non-life insurance.

Moreover, the emergence and further development of highly sophisticated comparison websites — via, for example, an increased offer of data filters and tools — could enable consumers to compare tailor made products in a big data environment. However, it remains to be seen how these websites will evolve to match the increasingly personalised offer of insurance products.

■ **Fair treatment and advice**

The IDD provides that "insurance distributors always act honestly, fairly and professionally in accordance with the best interests of their customers" when distributing insurance products. Consequently, in a hypothetical case, where the use of big data in insurance would lead to poor selling practices, the IDD would be breached and enforcement authorities would have a case to impose sanctions. In this respect, any insurance distributor who fails to comply with the IDD's conduct of business requirements will face severe economic penalties.

The IDD also requires that any insurance product that is proposed to a customer shall be consistent with their demands and needs. In addition, it introduces rules on ensuring the suitability and appropriateness of insurance-based investment products (IBIPs) for customers. Insurers are, for example, obliged to obtain detailed information on customers' financial situation, their investment objectives and their knowledge and experience in the investment field.

Furthermore the IDD introduces generalised product oversight and governance (POG) requirements into EU insurance distribution law, with the aim of ensuring that all insurance products for sale to customers meet the needs of their specific target market.

These requirements remain valid in a big data environment and address the risks of unsuitable products being sold to customers. Moreover the use of big data opens up the possibility to improve the identification of customers' demands and needs and specific target markets.



What is the way forward?

While the use of big data in insurance is still in its infancy, it has real growth potential. It offers benefits to consumers in terms of better tailored policies and prevention tools, and to insurers in terms of more efficient processes and accurate calculation of premiums. At this stage, insurers are exploring the opportunities that the use of big data analytics can offer throughout the insurance value chain. Where already tested, insurers have experienced positive effects thanks to the increased availability of data. For instance, the possibility of offering insurance cover for high risk customers where previously it did not exist, such as insurance for breast cancer patients or universal property insurance in flood risk areas.

To ensure the full benefits of big data for consumers and insurers, any future regulatory framework should be supportive of innovation. Currently, Insurance Europe does not see the need for further regulatory measures, as there is already a comprehensive framework regulating the use of big data in insurance. In fact, premature regulation could not only hamper innovation and impair the effectiveness of the insurance market, but could quickly become unfit for purpose due to technological advances and market developments.

Regulators and supervisors should ensure that existing rules, such as the GDPR and IDD, are fully implemented and enforced at national level. These rules already provide a framework to guarantee the responsible use of big data analytics in insurance.

However, regulators and supervisors should continue their efforts to closely monitor the impact of the use of big data on markets and consumers, and work together with stakeholders, including the insurance industry, to support innovation that benefits consumers. In this regard, some national insurance associations, in collaboration with consumer organisations, have developed tools to monitor how insurability is developing in the big data environment and to analyse any trends that could lead to undesired effects.

Finally, regulators and supervisors can also encourage the exchange of information and experiences between insurers on new tools and best practices.



Insurance Europe is the European insurance and reinsurance federation. Through its 34 member bodies — the national insurance associations — Insurance Europe represents all types of insurance and reinsurance undertakings, eg pan-European companies, monoliners, mutuals and SMEs. Insurance Europe, which is based in Brussels, represents undertakings that account for around 95% of total European premium income. Insurance makes a major contribution to Europe's economic growth and development. European insurers generate premium income of more than €1 200bn, directly employ over 950 000 people and invest over €10 100bn in the economy.

Insurance Europe aisbl
rue Montoyer 51
B-1000 Brussels
Belgium
Tel: +32 2 894 30 00
E-mail: info@insuranceeurope.eu
Twitter: @InsuranceEurope

www.insuranceeurope.eu