

# Grip on data

Green paper on Big Data



VERBOND VAN VERZEKERAARS

## Contents

<b>1</b>	<b>Summary</b>	<b>3</b>
1.1	Big Data is coming our way	3
1.2	Opportunities and concerns	3
1.3	Checks and balances	3
<b>2</b>	<b>Introduction</b>	<b>5</b>
<b>3</b>	<b>Insurers and data: business as usual</b>	<b>5</b>
3.1	Why do insurers collect data?	6
3.2	Which data?	6
<b>4</b>	<b>Big Data</b>	<b>7</b>
<b>5</b>	<b>Big Data: the opportunities</b>	<b>7</b>
<b>6</b>	<b>Big Data: the concerns</b>	<b>9</b>
6.1	Privacy	10
6.1.1	Privacy as a means of payment	13
6.2	Solidarity	14
6.3	Exclusion and inclusion	15
6.3.1	Discrimination	16
6.4	Concerns about the business model	16
<b>7</b>	<b>Checks and balances</b>	<b>18</b>
7.1	Privacy checks	18
7.2	Solidarity Monitoring and Evaluation	19
7.3	Business model	21
7.4	Legislators and monitoring bodies	22
<b>8</b>	<b>Conclusion</b>	<b>22</b>
	Questions or comments?	23

“The black box with Big Data is becoming larger and blacker. The data recorded in the box is becoming more complex and decisions taken on the basis of the black box are becoming more important. Both administrators and consumers need to be able to have confidence in the content of the black box. With this discussion paper, it is great to see the Dutch Association of Insurers taking a proactive stance in this regard.”

Sander Klous, Big Data Ecosystems Professor at the University of Amsterdam and Managing Director of Big Data Analytics at KPMG.

# 1 Summary

## 1.1 Big Data is coming our way

Each day, and indeed for many years now, Insurers<sup>1</sup> have been collecting data on their customers for risk assessment purposes. The insurance business depends on this data. But things are changing fast: the volume of data is increasing multifold and improved techniques are quickly coming on stream to find contextual links. Big Data has arrived. The question is what all of this will mean for the customers of insurers. There is great optimism about Big Data but ‘the revolution’ is also raising questions for society. Are customers better served by Big Data and what about the risks that are attached? And what kind of risks are these? With this green paper, the Dutch Association of insurance companies hopes to make a contribution in responding to these questions and in bringing the discussions on Big Data a step further. The objective: a society and customers who will get the utmost benefit from Big Data and insurers who can work towards a strong sector that is ready for the future.

## 1.2 Opportunities and concerns

Big Data has many opportunities to offer society. For example, products that are better geared to the wishes of the customer, faster claims handling and increased prevention options. But such opportunities can only be capitalized on if the concerns are properly addressed. Because these concerns do exist. For example, Big Data may lead to privacy risks: What will happen with all that personal data? What do insurance companies do when they know that the risk for some people is greater than for the rest? Will people still be able to get insurance? And what will this knowledge about risks mean for the insurance model? If only those who run a big risk take out insurance, the premiums would be sky-high.

Concerns such as these are of course justified but that these risks will become reality. It should

not be seen as a reason for rejecting Big Data. What is clear is that we must use Big Data in a responsible way, and be attendant to the potential risks. Where risks do arise, there are ways of dealing with them directly and effectively.

*“The objective customers who will get the utmost benefit from Big Data and insurers who can work towards a strong sector that is prepared for the future.”*

## 1.3 Checks and balances

To get the maximum benefit from the data for society, for the customer and the insurer alike, we will need to act responsibly. With this in mind, the Dutch Association wants to be in the frontline in the coming years taking a constructive approach to the developments. This green paper is a first step in that process. In it we present a number of *checks and balances* to give customers a better grip on the data. The Association intends to put measures in place aimed at further increasing awareness of privacy in the sector and helping insurers to convert the awareness into deeds, for example by exploring the creation of more effective privacy statements. That done, the Association plans to monitor whether the segmentation between customer groups by Big Data will lead to the exclusion of certain groups. To this end we will introduce an annual Solidarity Monitor. Finally, the Association will take measures to ensure that the sector gains the utmost benefit from Big Data and thus continues to operate within a vigorous business sector into the future. In this context, the Association would suggest that supervisory bodies and legislators exercise caution with regard to new legislation, alongside the EU's General Data Protection Regulation.

---

<sup>1</sup> This paper does not apply to health insurers in view of the very different nature of their business and the manner in which data is made use of by them.

Aside from these concrete guarantees, the Association will also conduct discussions on the use of data. This is never easy: discussions about data usage tend to be quickly viewed in terms of black and white. Our hope is that this paper will contribute to providing more clarity in the discussion. The Association also calls on its members to keep discussing the issue. The Association also has a close involvement in the *High Level Expert Group Big Data* of the Dutch

Ministry of Economic Affairs. Many external parties were consulted during the preparation of this paper, including the Personal Data Authority, TomTom, TNO (Netherlands Organization for Applied Scientific Research), the Rathenau Institute, Bits of Freedom, various universities and ministries. Every effort has been made to present the various perspectives on this issue but that is not to say that the parties stand behind this paper.

### Big Data

There are various definitions of Big Data. One of the most used comes under the three Vs: volume (quantity), variety (various sources) and velocity (speed with which data is generated and analysed). Other definitions combine two facets: internal/external data and structured/unstructured data. The definitions vary and indeed there are many ways of explaining the term Big Data, but on the whole it pertains to the increase in the volume of data and the ever-increasing possibilities of drawing links between all the data coming on stream. A new type of information is thus generated which can be useful to consumers, companies and governments alike. See Chapter 4 for more information on Big Data.

#### The Association will:

- set up a focus group on Big Data to give content to and monitor the framework and operation of these guarantees;
- carry out an annual Solidarity Monitoring and Evaluation;
- assess the precise difference between discrimination and premium differentiation based on risk;
- examine privacy by design;
- pay special attention to more effective privacy statements in the update of the Code of Conduct for the Processing of Personal Data;
- pay due regard to knowledge requirements concerning data in our information and training work;
- further elaborate in future documents on the principle that customers remain in control of their own data;
- watch out for the development of data monopolies when future legislation is drafted.

#### The Association advises its members as follows:

- Use Privacy Impact Analyses where necessary;
- deploy anonymous Big Data applications where possible;
- when differentiating between groups of customers, make sure to avoid any unintended discrimination and try to ensure that people with a greater risk can also obtain insurance at an affordable premium;
- ensure that all data obtained is well protected;
- communicate openly in relation to the use of Big Data and enter into dialogue with customers and stakeholders;
- review the means of communication: is the result of the Big Data analysis something that can be explained to consumers?

## 2 Introduction

Insurers have always relied on data in an effort to properly review the risks attaching to their customers. The opportunities have continued to increase in recent times due to new techniques and technologies which are fast outpacing each other. In this paper, we take a closer look at this phenomenon. How can insurers benefit from these developments? And, more in particular: what will it mean for the customer? We look at the major forces that have been unleashed by the Big Data revolution and outline the attendant opportunities and concerns for both customers and insurers. It should also be borne in mind that Dutch insurers still make limited use of Big Data analyses, particularly when compared to the United States. However, developments in this area are increasing at lightning speed. This paper is therefore a first step and of course depending on developments and responses in the community, viewpoints may change. The Association will keep abreast of new developments and discussions in this area. We will encourage discussion and, where possible, take a position.

## 3 Insurers and data: business as usual

'Data', according to Dikke van Dale, means: 1. *particulars; facts*; 2. *presentations of facts that can be processed in information technology*. Such particulars and facts are very important to insurers in their core task: assessing risks and executing the contract. That is how it is today but 50 years ago it was no different. Two examples:

1. A car insurer in the process of quoting for insurance asks for the vehicle registration number. This enables the insurer to check certain basic details about the car in question with the RDW (National Vehicle and Driving Licence Registration Authority): make, model, type, year of manufacture, type of fuel. This saves the customer on form filling work and provides the insurer with correct details. Next, the insurer asks the customer to estimate the

number of kilometres s/he drives per year. Other important details are the driver's age, postal code and the number of claims-free years. All of these details help the insurer to estimate the appropriate premium and to provide a quotation. To execute the contract, the insurer will need some further information such as the full name of the insured party and his/her contact details. This helps the insurer in ascertaining whether the prospective policyholder has been involved in any fraudulent claim.

*“Data is essential to insurers. That is how it is today but 50 years ago it was no different.”*

2. Term life insurance (ORV) is often obligatory when taking out a mortgage. If the person making the mortgage payments dies, the insurance payout can be used to clear the mortgage debt in full or in part. In calculating the premium for term life insurance, the insurer seeks details on the prospective policyholder's date of birth, the insured amount, length of cover and, possibly, whether the applicant is a smoker or not. These details enable the insurer to determine the premium and to issue a quote. If the consumer decides to proceed with the insurance, the actual application process begins. Essential also to the application process is the health of the prospective policyholder. The health is determined on the basis of a health declaration which the prospective policyholder is obliged to complete truthfully. Based on the replies, the insurer sets the premium. If the state of health gives rise to a different premium than the one in the offer, the consumer will receive a proposal in this regard. For higher amounts, the prospective policyholder may be obliged to undergo a medical test. To execute the contract, certain other details are required, such as the name and the sex of the policyholder (the premium payer), his/her address details and the full name of the party to be insured.

### 3.1 Why do insurers collect data?

An insurance cannot be issued in the absence of data about the customer and/or the item to be insured. An insurer has to be able to estimate the extent of the risk involved and assess the premium due in order to cover the claim if it arises. However there is no directly proportionate link between the volume of data and the extent to which the insurer can assess the risk of a claim. The vehicle insurer can actually assess quite accurately how great the risk of a claim will be based on the data from the car insurance sample above. Additional data on issues such as the maintenance history of the car will not necessarily improve the assessment of the risk. It has more to do with the experience insurers have built over the years and is due to the law of large numbers (see box). Insurers already know on the basis of this limited set of data that of all customers with certain characteristics, an x amount of people will make claims amounting to a y amount of euros.

#### The law of large numbers

The law of large numbers forms an important basis to the phenomenon of insurance. Insurers generate risk-assessments in relation to their customers based on a limited volume of data. They then group a body of customers together who have a sufficiently similar risk profile. If customer x dies earlier (or even later) than estimated, this will be compensated for by another person who lives longer (or shorter as the case may be) and who continues to pay the premium for a longer period. This is how the insurer creates or facilitates solidarity.

When collecting the data they need, insurers are also required to keep a record of data in order to comply with (international) legislation. For example, insurers are obliged to comprehensively identify prospective customers in order to prevent criminals from using life assurance for money laundering

purposes. Following acceptance too, certain policies or customers have to be monitored in order to prevent money laundering and terrorist financing. Transactions that are suspect must be reported to the national supervisory authorities.

### 3.2 Which data?

Can it be said that insurers have access to a lot of data in 2016? Yes and no. The biggest insurers in the Netherlands have millions of customers. For each and every customer, insurers do record and have access to a lot of data, but compared with companies on the web using search engines or social media, it cannot be said that insurers have access to large volumes of data. Insurers have compiled extensive data on the identity of customers (the name, address and place of residence) and information on the insured risk (e.g. residence, car, and perhaps the insured's lifestyle), but that is nothing compared to what the likes of social media companies know about their customers. The difference becomes all the more clear if we look at the source of the data. Data, according to the World Economic Forum, is *volunteered*, *observed* or *inferred*<sup>2</sup>. In other words: data is provided by the data subject (the customer), it is compiled on the basis of observation (although this should always be with the customer's knowledge!) or it is obtained on the basis of an analysis of patterns. Insurers usually have access to *volunteered* or *inferred* data. This means that the customers of insurers often have an idea of what the insurer knows about them.

By relying on the law of large numbers alone, insurers can analyse the probability and extent of the risk attaching to customers. It has to do with patterns; the *inferred* data. Further on in this green paper we will see that with the coming of Big Data, insurers will be in a position to make the switch to the category of *observed* data (which, for them, is new).

Moreover insurers, like many other companies, procure all kinds of anonymous aggregated

<sup>2</sup>[http://www3.weforum.org/docs/WEF\\_IT\\_RethinkingPersonalData\\_Report\\_2012.pdf](http://www3.weforum.org/docs/WEF_IT_RethinkingPersonalData_Report_2012.pdf)

data and they use information from the Central Bureau of Statistics (CBS). For example, criminality information, property value or composition of the family at municipality or district level. Such information helps insurers to match their products more to the customers' wishes.

## 4 Big Data

Data is nothing new to insurers. Why then present a separate paper on the subject? The big data phenomenon, as it is known, has been developing rapidly. Some ninety per cent of all data in the world has come about in the last two years alone, according to IBM<sup>3</sup>. This rapid increase is due to the world wide web and the search engines used on the internet. Computers are becoming smaller and faster. Developments have also led to an enormous increase in the capacity for storing data. In 1990, the storage of one gigabyte of data cost about one thousand dollars, in 2000 ten dollars and now just a few cents<sup>4</sup>. Every minute, more than one hundred hours of video is uploaded onto YouTube. Since 2014, more money is being spent in app stores each year than is earned in Hollywood making films<sup>5</sup>. Remember that there were no such things as app stores prior to 2008.

The rapid decrease in the costs of equipment and storage have made it possible to apply the use of sensors in all kinds of processes and objects. We already have sensors that can be placed on the water mains in houses to measure if there is a sudden drop in pressure, thus possibly signifying a leak. This allows for a signal to be sent to the valve to stop the flow of water to the main cock. This can save a lot of money. This is but an example of the Internet of Things (IoT) and data is set to grow exponentially as a result: lights that switch on automatically at a particular time of the day, or

when the owner gives a signal via an app, garage doors that open on arrival of the owner's car in the street, fridges that detect when food is going off, and so on.

This huge increase all has to do with 'Big Data'<sup>6</sup>. The capacity to find links in this enormous mountain of data is also increasing fast. This creates new opportunities for people and companies in identifying useful patterns in the Big Data. Credit card fraud can now be detected easier than ever before because credit card companies *can view card transactions in real time* and the system is immediately alerted when a card is used in an unusual manner. When this happens, the credit card company contacts the customer to check if anything is wrong.

Examples such as these can be found in all areas. Storms and floods can be predicted more easily. The police can see where there is a likelihood of criminality or a breach of the peace and deploy its forces accordingly. Consumers can get the products that are more suited to their requirements. Peaks in the demand for products and services can be better estimated, which of course is to the benefit of the customer. All of which is good news. For insurers too; Big Data makes it easier for them to assess risks. But there are concerns too. Concerns in relation to major issues such as privacy and solidarity. These concerns deserve to be given careful consideration. Failure to allay these concerns will diminish public backing for Big Data applications.

## 5 Big Data: the opportunities

The combination of large volumes of data and the capacity to link this data in a useful way has

<sup>3</sup> <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>

<sup>4</sup> <http://www.mkomo.com/cost-per-gigabyte-update>

<sup>5</sup> <http://www.asymco.com/2015/01/22/bigger-than-hollywood/>

<sup>6</sup>There are various definitions of Big Data. One of the most used comes under the three Vs: volume (quantity), variety (various sources) and velocity (speed at which data is generated and analysed). Other definitions combine two facets: internal/external data & structured/unstructured data.

opened up a whole new world for customers and insurers alike. And opportunities.

### Big Data and car insurance

Cars are changing slowly but surely into computers on wheels. A modern car can easily be equipped with an extra chip to collect data on driving behaviour (e.g. breaking habits, emergency stops, acceleration and sharp bends) and to forward this to a third party. Already, in the US and in the UK, insurers offer their customers this option. The driving behaviour data is then used to ascertain how safe is the driving behaviour. Each insured person receives a safety score which in turn leads to a discount on the basic premium. The customer can see from the premium how safe he or she is driving compared with other insured persons, and can check the dashboard for additional information regarding driving behaviour. The purpose of this is to encourage all insured persons to drive more carefully in order to lower premiums. The customer is charged a premium that is based also on his or her driving behaviour.

See the summary below for a sample of how Big Data can benefit customers. Just because something is *possible* does not mean that it is *necessary*. Firstly, insurers will need to weigh up carefully the opportunities as well as the concerns. Then it is up to the customer to choose whether or not to make use of products or services in respect of which additional data will be required.

**1. Faster claims handling.** If a car fitted with a chip that records driving behaviour is involved in a collision, a signal could be sent immediately to the insurer who would then be in a position to initiate the claims handling process without delay. The customer would then save the hassle of having to find a claim form and instead receives a phone call from the insurer asking what they can do to help.

**2. Service provision that is even more customer oriented.** Big Data helps insurers

gain an even better understanding of their customers' wishes and to conduct their business in a more customer oriented and constructive way. If a customer happens to take out a travel insurance with world cover and the data would show that s/he never holidays outside Europe, the insurer could advise the customer to opt for a more limited cover. In the short term this might result in some loss of income for the insurer, but in the long term it may well enhance customer satisfaction. Big Data analyses can be of assistance to companies in giving their online customers the information they need: terms and conditions, termination options, a telephone number.

**3. More safety on the street.** Insurers could link and draw conclusions from the data compiled on all insured parties which could be of benefit in resolving social issues. For example, the composite data on all insured parties could be used to analyse the link between safety and street lighting or between flooding and the state of the sewage system. In this way, data usage contributes to a safer district. This benefits society, the customer and the insurer.

**4. Prevention rather than curing.** Customer data, sometimes in combination with external data, can be of help to customers in avoiding risks. Prevention is better than the cure, not only for the customer, but also for the insurer (less damage). In the case of a burglary the insurance will cover the financial loss, but one cannot compensate for the sentimental value of things. Big Data, the pattern of burglaries or the characteristics of a dwelling can be used by insurers to give specific advice on home security matters. The more specific the advice, the more it will be to the customer's benefit.

**5. Premiums tailored to the customer.** Big Data analyses could also be used to set a customised premium for customers. Many customers who drive safely are reluctant to pay for other insured parties who do not. If customers reveal their driving behaviour to the insurer, the insurer could take this into account.



On the other hand, this principle can provide a stimulus for a healthier, safer and more socially conscious way of life. Because: with less safety and an unhealthy lifestyle, you pay more. This would call for clear acceptance guidelines and policy conditions.

**6. Better insurability.** Data provides information. In that way, it offers opportunities. In the past it was not possible for insurers to offer people with HIV a life assurance policy. Not enough was known about the life expectancy of people with HIV to prepare a meaningful risk calculation. Due to the increased availability of data, it became possible to insure the risks involved. Big Data may also be of benefit to more groups in this way too. The lack of data can make underwriting very difficult, whereas more data can lead to more insurability.

**7. Fraud prevention.** Insurance fraud costs millions of euro each year. Fraud is nothing more than theft from the co-insured parties: a portion of the premiums is used not to cover compensation costs but to cover fraudulent claims. Big Data analyses can help identify these claims. Analyses could give a better insight of which claims are likely to be fraudulent. Insurers can then carry out a targeted investigation in order to enable more fraud to be identified. Well-meaning insured persons, who luckily constitute the largest group, naturally stand to benefit from this development.

*“If there is added value for customers in Big Data applications, the inclination to share data will be greater.”*

We repeat: these are opportunities. In the Netherlands, Big Data applications are still used to a rather limited extent. Aside from their own business operations, the only new development we can see for insurers in 2016 are in *pay how you drive* car insurance, and *Usage Based Insurance*. Insurers are being rather cautious with the use of Big Data. Maybe this is because customers are sceptical about having their data used for commercial purposes, unless there is

a clear benefit in it for them. One therefore needs to be aware at all times whether there is a level of support for this. The basis for that support is *shared value*: if there is a value to be had for customers in the Big Data application, they will be more inclined to share the data. It is clear therefore that insurers also will respond to this in different ways.

## 6 Big Data: the concerns

These great opportunities unfortunately may have a negative side: people also have concerns. One important concern has to do with the privacy of customers, and their right to protect their own privacy. Another concern is the potential tension between Big Data and solidarity. One aspect of this is the possibility that Big Data usage will lead to indirect discrimination. Finally there is a concern which affects not so much the customer, but rather the insurers themselves. Big Data also raises concerns in relation to the model of insurance. Will customers still need an insurer in the coming period? And if insurance is only sought by customers who run high risks, won't the premium then be too high? Such concerns are of course conceivable but it is also worth bearing in mind that new technologies do raise concerns at times which later transpire to be groundless or which are much outweighed by the benefits. The Internet itself does raise many concerns in regard to privacy but the benefits clearly outweigh the disadvantages. There are also conceivable ways of the keeping the drawbacks to a minimum: cookie legislation, firewalls, etc. The recent switch by a large messaging service to encrypted messages shows that there is a need for privacy-friendly services.

## 6.1 Privacy

Big Data has to do with the data of customers. Therefore it also has to do with the issue of privacy. Many of the concerns it raises are addressed in legislation. The Dutch Personal Data Protection Act is also a fundamental basis for this. In an article, the chairman of the Dutch DPA outlines the key issue in terms of privacy: *The statutory privacy principles are intended to counter the unbridled collection and use of personal data: be transparent about the data you compile and use (transparency); do not*

### Key concepts of the Dutch Personal Data Protection Act

Under the Act, processing of personal data is defined as all actions involving the processing of personal data. It deals with

the destruction of the personal data, including everything in-between. The main conditions governing lawful processing of personal data are: (i) establishing the objectives of the processing and 'compatible usage'; (ii) establishing a basis for the processing; and (iii) the duty of disclosure that rests on the controller (i.e. the party who determines the purpose of the processing and the manner of processing). An outline of the main points of the Dutch Personal Data Protection Act follows below.

### Point of departure

The point of departure is that personal data must be processed in a proper and meticulous manner that is compatible with the law.

### Purposes

Personal data may only be compiled in relation to specified, clearly defined and legitimate objectives. The purpose for which the personal data is compiled applies as the assessment criteria governing many other conditions, such as compatible usage, retention periods and the condition that no more personal data may be compiled than is necessary for the purpose in hand.

### Legitimate basis

The processing of personal data must be in accordance with one of the bases laid down in the Dutch Personal Data Protection Act (Section 8). The processing of personal data is prohibited if no valid basis exists. Insurers base the processing of personal data on the ground that the processing is necessary for the conclusion and execution of a contract with the relevant party (whose personal data is processed), in order to comply with their statutory obligations or because the processing is necessary in relation to the legitimate interests of the financial institution. Aside from these three bases, financial institutions also process data with the unequivocal consent of the relevant party.

use data for any purpose other than that for which you compiled it (purpose limitation); do

### Compatible usage

Data may be used at a later stage for other purposes subject to strict conditions. These other purposes may not be incompatible with the purposes for which the data was originally compiled. The term compatible usage depends, among other things, on the connection between the two purposes, the nature of the data (sensitive or not), the consequences of the processing for the relevant party and the extent to which suitable guarantees have been provided for the relevant party. All circumstances need to be taken into account and weighed accordingly on a case by case basis in order to ascertain whether any further data processing is justified. If the relevant party has given permission to proceed with further processing, the requirement of compatible usage will at least have been met.

### Quality of the personal data

The quality of personal data comprises of two aspects. Firstly, the personal data that is to be used may not be more or any different than that which is necessary for the purpose of the processing. In addition, the controller must take whatever measures are reasonably necessary to ensure that the personal data are correct and precise.

### Retention period

Section 10 (1) of the Dutch Personal Data Protection Act provides as follows: *Personal data must not be retained for any longer than is necessary for the realization of the objectives in respect of which the data was compiled or subsequently processed.* The controller therefore needs to ascertain whether there are grounds on which the personal data may remain recorded.

### Security

The controller must safeguard the personal data in a manner that is appropriate against loss and unlawful processing. The type of data must also be taken into account as well as the technical possibilities that arise.

### Duty of disclosure

The controller must notify the relevant party of the purposes for which the personal data will be processed (unless the relevant party is already aware of this).

The controller is accountable to the relevant party in this context. As a rule, when commencing a relationship with an insurer, the initial and/or proposal form should set out the purposes for which the personal data is being compiled.

### Code of Conduct

The points of departure referred to above, as well as the other obligations provided for under the Personal Data Protection Act, are set out in more detail for insurers in the Code of Conduct for the Processing of Personal Data by Financial Institutions (GVVFI). This code of conduct also contains an extensive explanation. All members of the Federation of Insurers are obliged to comply with the GVVFI. Compliance is monitored by the independent Foundation for the Evaluation of Insurers (Stichting Toetsing Verzekeraars).

*not use any more data than is necessary for the purpose (data minimization)*<sup>7</sup>. In this paper, the Association wishes to deal with these principal complaints.

In the case of Big Data, compliance with the statutory parameters is not enough. This is because Big Data also raises ethical questions, such as the extent to which insurers are obliged to determine which high-risk behaviour should and should not result in a higher premium. These questions cannot be addressed by reference to the law alone. Additional concerns arise also in relation to the usage of *observed data*. This is because customers are often not in a position to exercise control over such matters. A customer may have once ticked a box to avail of a service but is unaware of the data consequently being compiled and perhaps even shared with third parties. Therefore in the

<sup>7</sup> <https://www.ictmagazine.nl/columns/big-data-en-het-voorkomen-van-digitale-predestinatie/>

case of *observed data* it is important to state all the more specifically how this data will be used. This gives customers a grip on what happens with the data.

Such matters are regulated by cookie legislation where they relate to observation of online surfing behaviour but when it comes to observing offline behaviour, this too should be clear for customers

Details regarding how the existing rules apply to Big Data applications were provided in a letter issued to the Dutch House of Representatives by the Minister of Economic Affairs in November 2014. Despite some overlap with the summary provided above, we have chosen to refer in more express terms to the letter from the Minister because this was the first time that attention was paid in such detail to the right to data protection and the Big Data revolution. In the letter, reference is made to the two most important privacy principles regarding Big Data: *legitimacy* of the processing and *information provision* to the customer proper.

**1. Legitimacy.** Personal data may only be processed in a proper and careful manner that is compatible with the law. Personal data must be compiled for specified legitimate purposes that are clearly outlined in advance and subsequently may only be processed for purposes that are compatible therewith. This is of importance with regard to the use of Big Data. An insurer who compiles data for car insurance purposes is not automatically entitled to use this for other purposes. Let us assume that the car insurance data reveals that a customer often parks near a sports school: the insurer would not be entitled to go ahead and offer the customer another insurance on that basis. Customers must be informed about this in advance and they must provide their consent.

**2. Provision of information.** The person whose personal data is being processed (i.e. the relevant party) must at least be aware of the

identity of the organisation or person processing the personal data (i.e. the controller). Not only must the relevant party be informed but s/he also has the right to see which data is being processed and the right to correct or remove data that is incorrect or irrelevant, not to mention the right to object (particularly against its usage for marketing purposes). This existing legislation is designed to give customers a grip on the data. The proposed European directive will create new and more extensive rights for customers, such as the right not to be subject to automatic decisions.

The Minister also pointed out that this framework development will safeguard against unlawful breaches of privacy, but emphasised also that Big Data is in contradiction with these principles: [...] 'in contradiction with the principle of legitimacy, because it is useful and profitable to use data for another purpose and in another context than for which it was compiled, while also retaining the data for a longer period with an eye to future analyses and to process as much data as is possible to facilitate a better analysis'. However the Minister did say there were possibilities, providing companies carefully weighed up their legitimate interest against the right to privacy. A *Privacy Impact Assessment (PIA)* could be of use in this regard. This is a method that compels organisations to consider pro-actively how the intended project will impact on the privacy of the relevant parties and what the risks will be for relevant parties as well as their own organisation<sup>8</sup>.

The anonymisation of customer data could also be of assistance here. Take for example the data compiled for a particular contract where such data need not be retained for long periods at the end of the contract. Such data must be erased in accordance with the law. This is a pity, because valuable knowledge can be lost. Anonymisation could be the solution.

<sup>8</sup>A frequently used format for a PIA was developed by Norea, the organisation for IT auditors: <http://www.norea.nl/readfile.aspx?ContentID=82987&Obj>

*“Insurers have always been trusted to date. It is important that this trust be maintained.”*

There can also be a conflict with the principle of information provision because notifying people about the many complicated analyses and the volume of data used in the process can be quite cumbersome, as the Minister stated in his letter. It is indeed cumbersome; privacy issues often tend to be raised only when something goes wrong. Recent research shows that even the best designed privacy statements are not read and have no effect on the behaviour of people<sup>9</sup>. There is an important challenge here for companies in keeping customers informed and interested in such information. The Association will also play its part in this.

### 6.1.1 Privacy as a means of payment

Some people are critical about the *pay how you drive* insurance. This scheme offers insured parties a premium discount in return for data on the driving behaviour. Criticasters state that privacy is being used here as a means of payment and claim that if this development continues, privacy will only become affordable for the well-off<sup>10</sup>. This assertion can be refuted. No direct exchange is taking place here between privacy and premium but rather between data and the premium – something insurers have always been doing. There has always been a big difference between people who choose or opt not to share data with an insurer. With no data insurance is not possible, with data it is. The only difference here is that insurers now *have more* data at their disposal; this allows them to make a better risk assessment and to deliver a better product. Where the boundary lies in terms of this trade-off between data and premium and between privacy and premium is a question that can probably only be answered in practise. What kind of exchange is acceptable to society and what kind is not? In life assurance, it was inconceivable 40 years ago to charge higher

premiums for smokers than for non-smokers; now it is entirely acceptable. See also Section 7.2 on the Solidarity Monitoring, which also deals with the issue of providing or not providing behaviour data. There is of course a difference between the type of data and how privacy-sensitive it can be. However sensitive information can also be processed confidentially in a manner that cannot be used by others or for undesirable purposes. The question is what are people comfortable with. The precondition in every situation will be that insurers must treat data confidentially.

### 6.1.2 Proportionality

An important prerequisite under the Dutch Personal Data Protection Act is that no more data should be processed than is necessary. This is termed ‘proportionality’. The consideration as to whether more data may be processed must always be made on an individual basis, but generally in this context the question can also be put whether insurers really do need Big Data. Are they not already in a position to make very good risk assessments based on the law of large numbers?

Indeed, the limited data set suffices for the purposes of calculating a premium for a large group of customers. But we live in an individualised society. And customers expect a more personally tailored service. We can see this for example in the car insurance market, where insured parties with claims-free years are rewarded for their good driving behaviour. Customers also increasingly expect their insurer to provide a more comprehensive service. The insurer who can liaise on these issues with the customer within the statutory frameworks and who can respond at the customer’s behest will win the competition. For example, by providing tips on safer or more efficient driving behaviour. Or about burglary prevention. Prevention is better than the cure. While matters of proportionality have to be considered by each individual insurer, it can also be said that in general, Big Data may be

<sup>9</sup> *Simplification of Privacy Disclosures: An Experimental Test*. Paper published in 2015 by Omri Ben-Shahar & Adam Chilton.

<sup>10</sup> <http://www.nrc.nl/handelsblad/2015/11/28/zo-behouden-alleen-de-rijken-hun-privacy-1561104>

necessary in order for customers to be provided with the best possible service. In that case the added value must clearly be for the customer.

## 6.2 Solidarity

Big Data can also have consequences for the solidarity. Big Data does not exclude solidarity but Big Data could for example lead to more segmentation. This is because more data on the risk and on their customers will place insurers in a better position to offer customised products and premiums. Take the *Usage Based* car insurance, where the insurer adjusts the premium based on a safer or more careless driving behaviour by the customer. In such cases insurance can be offered to a limited group of customers who all have the same score. But what if the group continues to reduce in size? Does the solidarity fall by the wayside?

Another threat to this solidarity could be in the dissipation of the information symmetry. When a customer takes out an insurance, both parties must have the same risk information at their disposal. If a customer knows more than the insurer, this could be problematic. It is a good thing for customers to have a better understanding of the risks they face. But if they subsequently choose not to participate in an insurance that also covers people who run a much higher risk than themselves, this could pose a problem for the insurer. Based on this example: the safer driver may not wish to continue paying for other drivers who happen to be more careless. This first driver then opts out of the insurance. It all sounds good for customers as it will be easier for them to take control in such instances. But the customer who runs a heavier risk will be the victim; if the insurer is only left with unsafe customers, the cost of claims will end up being higher than estimated and the premium will have to be increased. Will this not lead to an end to the solidarity?

The answer to this question depends also on what is meant by the term solidarity. If we regard solidarity simply in terms of the size of the group that shares the risk (a segment), then Big Data

will indeed have consequences for solidarity. But the Dutch dictionary Van Dale defines solidarity as ‘an awareness of unity and a willingness to bear the consequences that follow from this’. It is doubtful whether the individual who takes out a car insurance in 2016 will have a feeling of solidarity towards co-insured parties. This is how modern insurance started at the outset, when people and companies approached each other with the idea of sharing the risk of a fire at their home or business premises. If persons in the community wanted to join such a group of insured parties who themselves had a more carefree attitude regarding safety in the home or at work, they would most certainly be challenged or in fact not even allowed to join the insured group.

*“When you know that you belong to a select group of safe drivers you prefer to stay that way. In that context, Big Data does facilitate solidarity.”*

Today, however, insurance is fully anonymised: people are no longer called to account by co-insured parties for their behaviour. What’s more: who knows if they have the same insurer as the neighbours? When an insurer with access to Big Data has better knowledge regarding which insured parties carry a similar risk, they can be put in a group and be offered a customised product. This can in fact give rise to behaviour connected with more solidarity. When you know that you belong to a select group of safe drivers you prefer to stay that way. In that context, Big Data does facilitate solidarity. But for that solidarity, the size of the group is not so important. We can see this in 2016, with more and more small cooperative societies coming on stream. Bread funds are a fine example of this<sup>11</sup>. These funds comprise of a maximum of 50 independent entrepreneurs. They make a monthly contribution and when one of them falls ill, the fund pays out. The feeling for solidarity and the obvious preparedness to bear the consequences is much greater with such a group than in the case

<sup>11</sup> [http://www.broodfonds.nl/hoer\\_het\\_werkt](http://www.broodfonds.nl/hoer_het_werkt)

of traditional insurance. However the size of the group is very limited.

Big Data can also prevent the misuse of solidarity. The anonymity of the modern insurance system also leads to insurers having to contend with a lot of fraud. On a yearly basis Dutch insurers are dealing with fraud amounting to a hundred million euro. We suspect this is but a part of the problem. Big Data analyses can also prove very cost effective on this front not only by means of detection methods but also because it greatly facilitates the prevention of fraud.

Big Data and solidarity therefore surely do not exclude each other. Big Data can be of help to insurers in creating more groups of people with a more or less similar risk profile, thus offering them a customised insurance product and premium. As a result, Big Data can increase the solidarity within such groups but it can also increase the differences between these groups. This need not necessarily be inconvenient, as long as everyone can get insurance at more or less acceptable conditions. The question is how far will this trend continue.

#### **There is always room for solidarity**

With the arrival of self-driving cars we sometimes hear about the end of the need for car insurance. The likelihood of claims will undoubtedly fall, because such cars will seldom be involved in collisions. However one assumes there will still be risks involved. A heavy shower of hail can result in expensive claims and cars can always be stolen (perhaps even remotely). Not to mention a *software bug* which could cause serious damage. There still remain the unforeseen 'external contingencies' which call for the need for insurance and which generate this solidarity between people. The question is who will bear these risks: a traditional insurer, a car manufacturer or a platform based on the *block chain* (more about this last option below). To a certain extent, there will always be groups of people

with a more or less similar risk profile, which always leaves room for solidarity.

### **6.3 Exclusion and inclusion**

Big Data makes insurers even more aware of the risk attaching to a customer. Insurers therefore might decide on the basis of this knowledge to exclude high risks or to insure them only against a higher premium. Doomsday scenarios are easy enough to envisage but there are always enough forces at play to prevent this occurring in reality.

By way of example, let us look at the standard home contents insurance. With the arrival of all kinds of devices and sensors in the home which are connected to the web (Internet of Things), the home insurer could know even more about the insured and their home (providing the insured has given permission in this regard). It could lead to an insurer being aware that a certain customer failed to replace the smoke alarm battery on time. However, the same customer may also be the type who very conscientiously locks the dwelling when going on holiday. The insurer now has a lot more data at its disposal and may be in a position to provide good advice to the customer. But does this customer represent a big or a small risk? Such questions are not so easy to answer. Human behaviour is not always consistent and in line with the insured risks. So even though insurers would have more data, this may not always lead to exclusion.

But it is not likely that insurers will exclude large groups of potential customers. After all, the Dutch insurance market is a saturated market. There is little growth potential for insurers, other than at the cost of other insurers. If such a group were to be excluded, it is quite probable that another company would come forward with an alternative product. In the past there were groups of customers who were unable to obtain cover from existing insurers and who organised their own insurance amongst themselves. And there are now also people who in the absence of Big Data are not insured. There is poverty in the Netherlands. Some people can't afford basic insurance, with or without big data.

*“It is not likely that insurers will exclude large groups of customers.”*

Nevertheless, we need to keep a close eye on potential exclusion. If new groups of uninsured come about because of the use of Big Data, action will have to be taken. That is the public duty of insurers: the principle of insurance must be open to everyone in the Netherlands. This is also one of the principles of the Dutch Code of Conduct for Insurers: making it possible. As the Dutch insurance industry did on an earlier occasion with the arrival of Rialto. This company was set up in the ‘Sixties to deal with the fact that motor vehicle insurance was obligatory but that some customers (for example with a pretty poor claims history) were excluded by every insurer. If such a group of uninsured parties (for which the market cannot find a solution) should arise once more, then there are always conceivable ways of finding a solution.

### 6.3.1 Discrimination

Another aspect worth looking at is the discrimination phenomenon. Where groups of customers with a more or less similar risk profile are grouped together by insurers, there is a potential danger that they will be indirectly discriminating unintentionally against certain people. We speak of discrimination when a distinction is made on the basis of race (origin or skin colour), religion or personal beliefs, disability or chronic illness, sexual preferences, sex, age, or issues such as political persuasion, nationality or permanent and part-time contracts<sup>12</sup>. A Big Data analysis could show for example that people in a certain street are less creditworthy or come to the attention of the police and the courts more often than others. If many people living in this street are of a certain ethnic origin or religion, is the insurer then discriminating? Some unjustified and incorrect connections of this kind were highlighted in a recent report by the Federal Trade Commission

<sup>12</sup><https://www.rijksoverheid.nl/onderwerpen/discriminatie/inhoud/verbod-op-discriminatie>

<sup>13</sup><https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf>

(FTC) in America. Fortunately, the FTC was also able to cite many good examples of the use of Big Data<sup>13</sup>. It is all the more important therefore, according to the FTC, that the used analyses be studied very carefully with regard to the use of Big Data: do they contain elements that will give rise to discrimination? How accurate is the model? Is the random check representative? We do need to avoid a situation where Big Data leads to ‘digital predestination’. This term was coined by Mr Kohnstamm of the Dutch Data Protection Authority<sup>14</sup>. He contends that people are no longer capable of organising their own lives when all kinds of choices are made by companies on the basis of self-constructed profiles of (potential) customers. Based on a particular profile, people belonging to one group are made an offer while people from another group are ignored, despite the fact that those in the second group stand to benefit from the offer or that they have been wrongly placed in that group. Particularly when this transpires on the basis of *observed data*, in which people have little or no control, it can have negative consequences. This poses a challenge for insurers because insurers have always classified people in groups and then dealt with these groups accordingly depending on their risk profile. Insurers therefore need to be extra vigilant when using big data models and need to make sure that customers are aware of what is happening and be in a position to complain if necessary. Offer the customer ‘grip on the data’.

### 6.4 Concerns about the business model

Big Data analyses carry risks for customers, but also for insurers, which in the end impacts the customer. Big Data can complicate the business model of insurers but on the other hand, Big Data makes things easier, especially for new competitors. More data can give insurers a better insight, but as stated in the section on solidarity, customers too may adjust their behaviour accordingly. If customers

<sup>14</sup><https://www.ictmagazine.nl/columns/big-data-en-het-voorkomen-van-digitale-predestinatie/>



become better aware of the risks they face it could lead to an outcome where people who face lesser risks seek less cover or do not bother with insurance at all. That might be good news for such customers but the average loss per insured customer would then rise, and so too the premium would have to rise. This could give rise to a downward spiral so that in the end of the day all that remains is the group of customers with the highest risks who must pay a very high premium. What we have then is a de facto issue of uninsurability. This is not improbable, because with all this information technology, self-driving cars and sensors, customers are in a better position to be aware of the risks they face and what they need to insure themselves against and what not. This development should therefore be closely monitored. Some uncertainty is quite likely to remain, such as a hail shower or a storm. And all this dependence on IT can also give rise to new risks: risks in relation to identity fraud, hacks or system failures. There will always be the need to share risks or to transfer these to another party such as an insurer.

This point clearly shows that insurers do need to remain conscientious and to keep well abreast of the data revolution. Take for example the data that is compiled in a modern *connected* car: vehicle manufacturers are trying to keep this data exclusively to themselves. In this context, consumer organisations are already arguing that the 'owner' of such data should not be the manufacturer or another company but rather the customers themselves<sup>15</sup>. This is the current phenomenon with cars but it is set to become a broader issue with the Internet of Things when more and more sensors are placed in homes etc. Market power based on exclusive access to data can pose a threat for customers who will no longer be able to change to another service provider, but also for the insurers' own business operations. Ideally it should be the customer proper who decides who will access their data, not an IT company and not the insurer. This raises some difficult

issues, particularly in relation to data that was compiled in the past. Nevertheless, the Dutch Association wants to have this grip on the data by the customer as a point of departure for future developments. Fortunately, provision is made in the General Data Protection Regulation for a 'right to data portability' for customers. This is the right of customers to transfer the data kept by one provider to another.

*“Ideally it should be the customer proper who decides who will access their data, not an IT company and not the insurer.”*

Aside from the fact that Big Data causes concerns in relation to the business model of insurers, the Big Data revolution has also led to new competition. This issue is of no concern to insured parties, but for insurers it is. It is now possible to make calculations with the use of an average computer and data from the *cloud*, where previously this could only be done by large companies. This means that the competition could suddenly come from an unexpected quarter. Even the transactions can be easily arranged. Banks have been very busy studying the *block chain* technology. This is a public register of transactions: participants can have full confidence when using this system that their transaction can be executed as arranged without any intervention by *trusted third parties*. It is a register that records every few minutes who owes what amount to whom. Deutsche Bank expects to have the first commercial *block chain* alternatives in place by the end of 2017<sup>16</sup>.

It is quite conceivable that *block chain* applications will be used for the insurance market. Why wouldn't ten thousand individuals undertake not to sign up to making a payment to an unfortunate nobody who has been struck by fate? It is conceivable that such a platform could make the role of an insurer superfluous. Moreover someone would have to develop and manage that platform: perhaps an insurer. For the time being, the various difficulties attaching

<sup>15</sup> <http://www.anwb.nl/auto/connected-car/my-car-my-data>

<sup>16</sup> <http://www.coindesk.com/deutsche-bank-blockchain-tech-will-go-mainstream-next-decade/>

to this kind of *block chain* insurance make it too problematic. For example, who will assess the extent of the loss if the home of an insured burns down? Can this kind of *block chain* be hacked? What happens if there are suddenly many more claims than originally estimated? Who decides who will be allowed to participate? Difficult questions, but they do not necessarily have to stand in the way of a *block chain* insurance. There will be some bumps on the road, but the prospects are so great that the investment would probably be justified.

## 7 Checks and balances

It is clear from the opportunities and concerns described above that Big Data can be a powerful instrument. At least: if an eye is kept on the concerns that arise and provided that the disadvantages are kept to a minimum. Only then will insurers and customers stand to benefit from Big Data. And it should be noted: these are *possible* disadvantages. These risks have not yet arisen and they may even never arise.

In the first place, we have legislation to prevent such occurrences, particularly the Dutch Personal Data Protection Act and the coming General Data Protection Regulation. The current Act is being worked out in greater detail for insurers in the Code of Conduct for the Processing of Personal Data by Financial Institutions (GVVFI)<sup>17</sup>. At the same time, a certain amount of time is needed with every new development to see how the legislation is being applied to it. Will we need new rules? No. The Association has been arguing that we first should be keeping a close eye on the developments, so that targeted action may be taken wherever it is needed. If it is not known what the precise problems are that need to be tackled, how great they will be and where exactly they will arise, one might find that

because of the rules not only the risks are being prevented but the opportunities as well. In addition, the rules may also prove to be ineffective.

It is the belief of the Association that the insurance sector itself also has a big role to play here – perhaps in cooperation with other parties. In this Chapter we have a number of recommendations to make in that regard. To provide a good framework for this, the Association will first of all set up a focus group on Big Data to monitor developments and to hold regular discussions with stakeholders to see if the checks and balances that we propose here are effective.

### 7.1 Privacy checks

Firstly, the concerns about privacy. The point of departure here has to be that insurers remain in compliance with the regulations governing data protection. Also important in this regard is that insurers incorporate the privacy awareness in their organisations and particularly in the development of Big Data applications. Privacy is not just something to be borne in mind by the Compliance Department or the Data Protection Officer, but also by the product developers, marketers and by management. Aside from the Association's contribution here in terms of the green paper, it also proposes to organise or other ways of supporting its members with privacy compliance. And the Association will of course monitor the situation across other sectors. For example, *privacy by design* is a method designed to build in protection for personal data from the beginning of a product<sup>18</sup>; as from 2018 this will be obligatory for everyone who falls under the General Data Protection Regulation. In anticipation of this, the Association will review in the coming year how the principles of *privacy by design* can be applied to the world of insurance and promote this amongst its members.

The regulations governing the protection of data are becoming ever stricter. For example, the notification requirement in relation to data leaks

<sup>17</sup> <https://www.verzekeraars.nl/overhetverbond/zelfregulering/>

Paginas/Gedragcodes/Gedragcode-Verwerking-Persoonsgegevens.aspx

<sup>18</sup>[https://en.wikipedia.org/wiki/Privacy\\_by\\_design](https://en.wikipedia.org/wiki/Privacy_by_design)

has become obligatory in the Netherlands since the 1st of January 2016. The Association is making its own contribution to data protection: insurers exchange information in relation to potential security problems among themselves and with the National Cyber Security Centre, the National Police and the General Intelligence and Security Service (AIVD). This is done via the *Information Sharing and Analysis Center* (ISAC). During 2016, the Association will also establish a CERT, i.e. a *Computer Emergency Response Team*. This is an operational collaborative project in the area of cyber security that arranges for participating members to be informed and advised on a continuous basis. For example in relation to threats, vulnerabilities and incidents, but also regarding prevention possibilities and response perspectives. Aside from this, the Association prepared a policy for the sector in 2014 in the area of *responsible disclosure*. This policy helps insurers consult with well-meaning hackers in order to identify faults in the security systems of insurers. Also of importance for insurers is to develop their own means of data protection.

Of importance in the protection of privacy is the duty of disclosure: insured parties must know what is happening with their data. Insurers make this as transparent as possible via the privacy statement. Making insurers responsible for ensuring that customers are informed as well as possible concerning what is done with their data is also provided for in the Code of Conduct for the Processing of Personal Data; the insurance sector is one of the few branches of industry that have adopted a code of conduct. This code is to be updated in 2016, with extra attention being given to ways of informing customers about the privacy policy.

*“If it is not possible to explain the use of Big Data to the public, then you need to declare why you are doing it, or you should not do it at all.”*

Insurers are acting prudently in not waiting for this updating to come about and are finding their own ways of creating guarantees in order to make Big Data usage possible. In the first place,

they need to communicate openly regarding the use of Big Data. Why would they not do this, given that Big Data is to be deployed in a customer-focused manner? If it is not possible to explain the use of Big Data to the public, then companies need to declare why they are doing it, or they should not do it at all. Insurers need to inform customers about the usage of *observed data* and afford them the opportunity of correcting any errors so that customers have a ‘grip on their data’. On the web, this is done by means of *cookie* legislation. Offline, insurers make use of the *opt-in* option, in combination with information on what the data will be used for. Customers trust insurers with data. Insurers must do their utmost to keep that trust.

Other options include phenomena such as *data vaults* and other solutions that are even more innovative. The *personal data vault* enables customers to have control over their own personal data; a company may view this once only when preparing an offer. With such a model, the insurer will have only a few contact details on record. This model however will need to be further developed. Innovations such as these are expected to rapidly gain ground in the coming years, given the great tension between the Big Data opportunities on the one hand and the regulatory framework on the other. The Association will closely monitor these developments and where possible encourage insurers to avail of these methods in order to increase the customer’s grip on the data.

## 7.2 Solidarity Monitoring and Evaluation

We have seen how the solidarity within groups can sometimes increase but also how Big Data applications can engender fears that certain categories of people will end up being unable to get insurance cover. It is still too early to predict if this will happen but we do need to remain alert.

To give tangible substance to this, the Association will organise an annual Solidarity Monitoring and Evaluation, starting in 2016. The objective here is to identify any potential negative consequences of Big Data

applications at an early stage. As at 2016, people in the Netherlands are well insured. Where liability insurance is concerned, the penetration rate stands at 95 per cent; the same figure applies to home contents insurance. Of all car drivers, only a small percentage have been assigned to the Rialto safety net. As from 2016, the Association will carry out an annual check to see how this develops and make its findings public. This will also afford the opportunity for interest groups to notify the insurance sector about any groups of people who cannot get insurance. Should this be the case, the Association will endeavour to ensure that access to cover will remain at the same levels as in 2016. This is in keeping with Principle 21 of the Code of Conduct for Insurers, by virtue of which insurers undertake to offer as many customers as possible the possibility of insuring risks financially and commit themselves to preventing uninsurability<sup>19</sup>.

As part of the Solidarity Monitoring and Evaluation, the Association will monitor the range of development between the lowest and the highest premium together with the ancillary cover and conditions. The precise form this will take has yet to be decided, but the initial plan is to conduct a check with a number of reference persons to assess the development of their premiums and coverage levels. Checking, for example, what the highest and lowest premium is for a specific reference person with a certain type of cover and how the range between these reference persons develops will give us an idea of the emerging trends. The issue of privacy will also form part of the review: to this end, a reference person will be included who is unwilling to share behavioral data.

The important question of course is what is to be done with the results of the monitoring: what constitutes a problem and what happens when it is identified? This is something that the Association will be looking at more closely in close cooperation with others, including consumer organisations. The monitoring therefore is not to be regarded as a solution but

rather as a way of finding out if there is a problem, how big it is and where it arises.

In the Chapter on Solidarity we also observed that Big Data can lead unintentionally to indirect discrimination. This is a phenomenon that needs to be taken into account by each individual insurer. In the Netherlands, the Institute for Human Rights has been assigned to investigate the issue of (alleged) discrimination. The Association itself will also promote studies into ways of preventing (indirect) discrimination: exploring the establishment of guidelines in liaison with universities aimed at preventing situations in which data models unintentionally lead to indirect discrimination. Insurers also need to keep examining how customers can be enabled to correct their own data. Special attention will also be given to the anonymous data that is procured by insurers and used by them in preparing profiles of certain generic levels (for example post code, district or location). These do not constitute customer data, but customers will be affected if this data is a contributory factor in the determination of their risk profile. Even though the profile is fair for many of the people in the district, some of the people will feel that they are not being treated fairly in this process. Customers should be offered the option of correcting data in such circumstances. Here again we refer to the slogan: offer the customer a 'grip on their data'. If requested, people should be notified concerning the customer profile that was prepared on them. This goes beyond what is legally required, because the law only provides for notifying customers about their personal data.

*“Insurers need to consider options that will enable customers to correct their data.”*

When customers have questions about categorising, it is important for insurers to be in a position to explain their model to them. However complex the Big Data models might be, they should be fully understood by insurers so that corrections are always possible. The

<sup>19</sup>[https://www.verzekeraars.nl/overhetverbond/zelfregulerings/Documents/Gedragscodes/Gedragscode\\_NL\\_2015.pdf](https://www.verzekeraars.nl/overhetverbond/zelfregulerings/Documents/Gedragscodes/Gedragscode_NL_2015.pdf)

Association advises insurers to carry out a 'communication check' in this context. Such a check could include a provision whereby insurers must only use links in their models for which a logical, risk-related explanation can be given. There is for example a logical explanation for the fact that the weight of a car or the driving habits of the driver are of relevance to the amount of the premium. The payment behaviour of customers is also a logical factor: customers with poor payment records compel insurers to take collection measures. Factors that cannot be explained should not be included when compiling data models. If Big Data analyses cannot be explained by insurers, how can they expect customers to understand them or that there will be continuous public backing for same? This also ties in with the main principles of the Code of conduct Insurers. Principle 5 sets out that insurers must be clear about the operation and cost of products. Principle 7 states that the acceptance and claims handling process must be transparent for the customer. There are of course limits hereto. Insurers will be unable to disclose any calculation models that are a company secret. Moreover not everything can be explained to everyone: the issues are often rather complex. Nevertheless insurers must exercise their endeavours to be able to explain their models.

### 7.3 Business model

In conclusion, it did transpire that Big Data can raise concerns also about the business model of insurers. If insured parties gain a better insight of risks than the insurers, this could complicate matters for the insurers' model and have all kinds of consequences for customers. In addition, the Big Data phenomenon gives rise to new competition, possibly from a totally unexpected quarter. This is a good thing, as it provides a stimulus for the existing providers. At the same time, the rules that apply to insurers must also apply to these new companies. In this context we would make the following recommendations:

1. Insurers should organise their own data. Some insurers have to contend with some

*legacy* issues: systems in use by their legal predecessors which are not yet in line with their own or more modern systems. Before they commence with making use of Big Data, insurers must get their own data in order.

2. In order to make an informed choice based on the pros and cons of using Big Data, insurers need to have access to sufficient information on the matter. The Association will be organising course programmes to this end for the benefit of insurers. Insurers would be well-advised to work closely with universities and in this regard to gain the interest of talented people for their business. In 2016, the Association was involved for the second year in a course run by the Technical University in Delft where students use the existing data of insurers to gain new knowledge.

3. Any data found to be of significance in 2016 is set to be all the more so in years to come. It is therefore important for insurers to have access to data when desired. However with the ever-increasing value of data, manufacturers (of cars, but also of all kinds of domestic appliances) are trying to protect their data. This poses a risk. A customer may possibly benefit initially from closed systems but in the longer term, customers might find themselves locked into a system thus rendering it more difficult to change to another provider. A boiler supplied by manufacturer x will only heat the radiators supplied by manufacturer x; a thermostat from manufacturer x and a fitter from manufacturer x. This phenomenon has already proved to be a problem in recent years in the market for software packages and mobile phones, although recent times have brought some relief with the coming of new legislation. For this reason it is important for manufacturers to be able to form closed systems such as these so it will be possible for customers to switch. This will benefit insurers in their own usage of data, while also providing a healthy counterbalance where the power of manufacturers is concerned. In the General Data Protection Regulation, the proposal on *data portability* may prove to be very important in this regard: it gives consumers the right to ask

the original service provider to send them ‘their’ data in *amachine readable format*, thus facilitating subsequent transfer to the next other service provider. The Association, in liaison with its European sister organisations, will be on the alert to ensure that such *checks and balances* are incorporated in the regulatory framework on data.

4. Big Data applications appear to be infinite: if insurers fail to make use of it, a *fintech start-up* surely will. Insurers would also be well advised to generate their own competition. Some insurers are already investing heavily in *start-ups*. Insurers can also develop *blockchain* insurance applications. The Association has been encouraging this innovation with its own insuranceLAB: this is a centre where various parties from inside and outside the sector work together developing innovations.

5. The Association, together with the Centre for Insurance Statistics (CVS), has access to a lot of (anonymous) data. This gives insurers access to all kinds of confidential reports about their commercial position. The CVS analyses abstracts of this data to examine trends in society and in so doing works closely with academics on a regular basis. This role could be expanded on, however: on the one hand by compiling more data which can be of use in the development of knowledge for insurers, while on the other making data accessible for scientific research of issues that are relevant to society. The Association therefore also calls on academics to come forward with their ideas on the use of this data.

#### 7.4 Legislators and monitoring bodies

All of the above recommendations are meant for insurers or are to be carried out by the Association. The Association recommends that no new regulations be drafted by the supervisory bodies and legislators. Acting too quickly in order to avert potentially adverse consequences can also knock the opportunities on the head.

*“As long as there are no real problems, no new regulatory framework should be developed.”*

Granted, there are concerns (and some of them justified), but as we have already observed in this outlook, Dutch insurers hardly use Big Data at all in their products. And we should be careful not to fall into the ‘risk-regulation reflex’ syndrome. Regulations should not be based on concerns but rather on facts. Laying down rules on the grounds of concerns alone can hamper innovation and fly in the face of good opportunities.

## 8 Conclusion

Publishing a paper on Big Data without any reference to a ‘scientific’ formula would be inconceivable. For this reason we conclude with a formula for the acceptance of Big Data applications. Big Data undoubtedly has many opportunities to offer. But implementing a good Big Data application that passes all tests with flying colours is not an easy matter. Being sure that a Big Data application ‘will be OK’ is therefore not only a matter for legal consideration. This brings us to a Big Data acceptance formula:

#### Acceptance formula

The acceptance readiness for a Big Data application = customer benefit x openness x predictability x compliance x accuracy x comprehensibility x correction options.

Customer benefit forms the basis for the data analysis: acceptance readiness will grow once the application shows that there are clear benefits for the customer and society alike. Openness relates to how much customers are told regarding the end use of the data. Predictability concerns the extent to which customers can assess how their data will be used. Compliance has to do with whether the application falls within the existing statutory parameters. Accuracy has to do with whether customer’s characteristics are correctly

portrayed in the used model. Comprehensibility has to do with whether the insurer can explain the model in use. This is also necessary with regard to the correction options that should be available to customers if they are incorrectly assessed in the model. In short, this formula implies that customers must have control and that customers can keep a grip on their data.

This formula sets out clearly why the acceptance readiness differs between the various types of companies. The rules are the

same for everyone but the perception of customers varies considerably. It is not just a question of compliance, but also of what customers expect of a company. For this reason it is important for insurers to be open and predictable when it comes to Big Data applications. We hope that this paper will contribute to the transparency.

### Questions or comments?

Email us at [info@verzekeraars.nl](mailto:info@verzekeraars.nl)

“I think that the measures recommended by the Federation — Solidarity Monitoring and Evaluation and a communication check, could be very worthwhile. In my view, the impact of these developments will only become really evident once we see how they work in practice. In that context, experimenting is a good step while at the same time assessing the benefits for society in terms of solidarity and comprehensibility.”

Jelte Timmer MA, researcher at the Rathenau Institute