

# Big Data and Insurance: Implications for Innovation, Competition and Privacy

March 2018

## The Geneva Association

The Geneva Association is the leading international insurance think tank for strategically important insurance and risk management issues. The Geneva Association identifies fundamental trends and strategic issues where insurance plays a substantial role or which influence the insurance sector. Through the development of research programmes, regular publications and the organisation of international meetings, The Geneva Association serves as a catalyst for progress in the understanding of risk and insurance matters and acts as an information creator and disseminator. It is the leading voice of the largest insurance groups worldwide in the dialogue with international institutions. In parallel, it advances—in economic and cultural terms—the development and application of risk management and the understanding of uncertainty in the modern economy.

The Geneva Association membership comprises a statutory maximum of 90 chief executive officers (CEOs) from the world's top insurance and reinsurance companies. It organises international expert networks and manages discussion platforms for senior insurance executives and specialists as well as policymakers, regulators and multilateral organisations.

Established in 1973, The Geneva Association, officially the 'International Association for the Study of Insurance Economics', is based in Zurich, Switzerland and is a non-profit organisation funded by its Members.

# Big Data and Insurance: Implications for Innovation, Competition and Privacy

Authored by Benno Keller, Special Advisor for Digital and Innovation, The Geneva Association

Prepared in collaboration with:

Martin Eling, Institute of Insurance Economics, University of St. Gallen

Hato Schmeiser, Institute of Insurance Economics, University of St. Gallen

Markus Christen, UZH Digital Society Initiative, University of Zurich

Michele Loi, Institute of Biomedical Ethics and History of Medicine, University of Zurich

---

## The Geneva Association

The Geneva Association—International Association for the Study of Insurance Economics  
Talstrasse 70, CH-8001 Zurich  
Email: [secretariat@genevaassociation.org](mailto:secretariat@genevaassociation.org) | Tel: +41 44 200 49 00 | Fax: +41 44 200 49 99

Photo credits:  
Cover page—Shutterstock.

---

March 2018

Big Data and Insurance: Implications for Innovation, Competition and Privacy

Copyright 2018 © – The Geneva Association

Published by The Geneva Association—International Association for the Study of Insurance Economics

# Contents

<b>Acknowledgements</b>	4
<b>Foreword</b>	5
<b>Executive summary</b>	6
<b>1. Introduction</b>	7
<b>2. The changing role of data in the insurance business model</b>	8
<b>3. Economic, societal and customer benefits</b>	10
<b>4. Ethical and societal concerns with the enhanced use of personal data</b>	11
Concerns about privacy and data protection	11
Concerns with increasing individualisation of insurance	13
Concerns about competition	14
<b>5. Balancing benefits and risks</b>	16
<b>6. Potential future scenarios</b>	18
Scenario 1: The digital society	20
Scenario 2: Insurance at two speeds	22
Scenario 3: Privacy regulation	24
Scenario 4: Digital backlash	26
Scenario 5: A tale of trust	27
Discussion of scenarios and conclusions	29
Box 1: Telematics	30
Box 2: Use of social media data in underwriting	32
Box 3: Wearables	33
<b>Glossary</b>	34
<b>Appendix</b>	36
Types of data used in insurance	36
Overview of privacy trade-offs	37
Overview of potential future scenarios	38
Overview of concerns and potential public policy approaches	40
Framework for scenario development	41
Overview of propositions, benefits and privacy concerns	42

---

## Acknowledgements

This paper has been developed in collaboration with the research project “Between Solidarity and Personalisation – Dealing with Ethical and Legal Big Data Challenges in the Insurance Industry” (Swiss National Research Programme 75 “Big Data”). It has greatly benefitted from numerous insights, comments and input from a variety of experts. We would particularly like to thank the following persons (in alphabetical order): Robert Bauer, AIG; Courtney Bowman, Palantir; Lamrini Gyftokosta, Insurance Europe; Dietmar Kottmann, Oliver Wyman; Luca Marighetti, Swiss Re; Fabian Sommerrock, CSS; Fabian Winter, Munich Re; Katja Würtz, EIOPA; and Tal Zarsky, University of Haifa.

# Foreword



**Anna Maria D'Hulster**

*Secretary General,  
The Geneva Association*

The emergence of big data analytics and artificial intelligence has triggered a deep transformation of the insurance industry. Established insurers invest in the digitisation of their processes and products, while an increasing number of InsurTech companies are entering the market as insurers, distributors of insurance solutions, and at other points along the industry's value chain. Both incumbents and newcomers are developing insurance products that use large amounts of data to assess, select, price, predict and prevent risks that in some cases were previously considered uninsurable. Going forward, access to data and the ability to derive new risk-related insights from it will be a key factor for competitiveness in the insurance industry.

Large societal benefits arise with the potential to reduce risks and increase their insurability through the use of vast quantities of data. New approaches to encourage prudent behaviour can be envisaged through big data, thus new technologies allow the role of insurance to evolve from pure risk protection towards risk prediction and prevention. However, the use of big data in insurance raises complex issues and trade-offs with respect to customer privacy, individualisation of products and competition. Assessing these trade-offs requires complex value judgements, and the way they are addressed leads to different scenarios for the future development of the sector.

The societal and regulatory debate about the appropriate use of personal data and the implications of the ongoing digital transformation in the insurance industry has only just begun. Policymakers and regulators are increasingly becoming aware that we are at the crossroads regarding the future development of the sector. In this context, policy choices can have far-reaching consequences for the future face of the industry, its socio-economic relevance and the value it creates for its customers.

This report aims to contribute to an informed and fact-based debate by identifying and discussing key trade-offs involved with the application of big data in insurance. The paper discusses the implications of a wide range of uses of data and develops potential future scenarios to highlight likely consequences of different policy choices.

Anna Maria D'Hulster  
Secretary General of The Geneva Association

---

# Executive summary

Advances in big data analytics, artificial intelligence and the Internet of Things promise to fundamentally transform the insurance industry and the role data plays in insurance. New sources of digital data, for example in online media and the Internet of Things, reveal information about behaviours, habits and lifestyles that allows us to assess individual risks much better than before.

In many instances, better data makes it possible to better align premiums and risks and to reduce the overall cost of insurance. This has great economic and societal benefits in that it allows premiums to signal risks, reduces the cost of informational asymmetries in insurance markets, and enhances efficiency, thereby boosting insurance protection.

But arguably the greatest societal benefits come from the potential to reduce risks through better data and new digital technologies. The ubiquitous availability of vast amounts of data and the ability to analyse it allow for individual and dynamic risk assessment and a continuous feedback loop to policyholders, with no or limited human interaction. By providing risk insights to policyholders, such 'digital monitoring' encourages behavioural change to reduce risks. Moreover, new data sources allow for the implementation of advanced risk management systems that use predictive analytics as a basis for early intervention and risk prevention. Ultimately, these technologies allow the role of insurance to evolve from pure risk protection towards predicting and preventing risks.

These benefits do not come without a cost, however. In the public debate, many concerns have been raised. These can be grouped into concerns about privacy, concerns about individualisation of insurance, and concerns about competition.

Privacy concerns include concerns about fairness and discrimination, intrusiveness and contextual integrity of personal data. Concerns about individualisation of insurance refer to affordability and exclusion, implications for solidarity and risk pooling as well as premium volatility. Finally, concerns about competition include potential abuse of market power, the level playing field and market transparency.

We identify and discuss the key trade-offs involved with these concerns. Balancing these trade-offs requires intricate value judgements by consumers, firms, policymakers and regulators alike. Yet, finding an appropriate balance between privacy protection and allowing for innovation is of fundamental importance, as insufficient privacy protection will harm consumers and erode trust, while overly strict regulation may hinder society from reaping the benefits of the data.

As a basis for a fact-based regulatory debate on these trade-offs, we develop five different scenarios that highlight the likely consequences of different policy approaches. These scenarios include 'The digital society', 'Insurance at two speeds', 'Privacy regulation', 'Digital backlash' and 'A tale of trust'.

In these scenarios, we evaluate likely implications for consumers, firms and society at large of the different policy choices regarding privacy and access to data. Specifically, the five scenarios differ in the degree to which benefits from the use of data are realised, the level of privacy protection as well as the degree of competition.

# 1. Introduction

Advances in big data analytics, artificial intelligence, and the Internet of Things promise to fundamentally transform the role of data in the insurance business model. These technologies allow for the development of powerful new business models which in turn enable the role of insurance to evolve from 'understand and protect' towards 'predict and prevent'.

While the changing role of insurance promises great economic and societal benefits, at the same time it raises concerns about privacy and data protection, individualisation of insurance, and competition.

Individuals, firms and policymakers and/or regulators are confronted with intricate trade-offs when balancing the benefits of sharing personal data and the risks and concerns surrounding the use of personal data. What makes balancing these trade-offs difficult is that they are often context-specific, subjective, and non-measurable.

Yet, finding an appropriate balance between privacy protection and allowing for innovation is of fundamental importance, as insufficient privacy protection will harm consumers and erode trust, while overly strict regulation may hinder society from reaping the benefits from data.

This report aims to contribute to an informed and fact-based regulatory debate on these trade-offs. To this end we discuss the societal and economic benefits from the use of big data analytics in insurance and the key concerns that have been raised in public and regulatory debate. Based on this discussion, we identify the key trade-offs deriving from the enhanced use of personal data in insurance.

To evaluate the consequences of different policy choices, we develop five different scenarios that highlight the implications of these choices for consumers, insurers and society.

## 2. The changing role of data in the insurance business model

Since the emergence of probability theory and actuarial science as mathematical disciplines in the 17<sup>th</sup> century, data analysis has played a fundamental role in insurance.<sup>1</sup> These scientific advances allowed insurance to evolve from 'intuitive bets' on future states of the world to an industry based on rational calculus and decision making.

The church played a pivotal role in this transition by collecting data necessary to compute actuarial analysis. In the 16<sup>th</sup> century, parish priests in some European countries were ordered to keep records of baptisms and marriages, and later of funerals and burials. In 1693, such data allowed the English astronomer, geophysicist, mathematician, meteorologist and physicist Edmond Halley (1656–1742) to develop the first annuity table based on mortality data drawn from actual experience, resulting in a major leap forward for life assurance.<sup>2</sup>

Aggregated personal data such as mortality tables and accident statistics are important for insurers to estimate risks at the population level or for a large subset of the population. In addition, insurers have mainly relied on personal information<sup>3</sup> collected directly from policyholders at the time of underwriting a policy to group individuals into risk classes. For example, in personal auto insurance, insurers have typically relied on information such as type of car, age and loss history to group individuals into different risk classes that determine the premium rate of an individual belonging to a specific risk class.<sup>4</sup>

Over the past two decades, insurers have increasingly begun to deploy data from third-party data sources. For example, when empirical evidence emerged that people with higher credit scores also tend to be safer drivers, insurers started to incorporate credit scores into their analysis for personal auto insurance.<sup>5</sup> The

role of data, however, basically remained the same, namely to understand risks and protect policyholders by compensating them for incurred losses.

Today, advances in big data analytics, artificial intelligence and the Internet of Things promise to fundamentally transform the role of data in the insurance business model. These technologies are at the core of a new digital and interconnected infrastructure for the Digital Society, which continuously produces very large amounts of real-time data. Systems based on artificial intelligence and self-learning algorithms use large amounts of real-time data and feedback loops to continuously optimise themselves.

This development is fuelled by the emergence of two new sources of data that are relevant in the context of insurance. The first consists of data that is automatically generated and stored with our online behaviour. Such data includes personal information shared via social media platforms, personal online shopping behaviour generated through e-commerce, and data generated by our personal search and browsing activity.

Personal data on online behaviour can reveal information about the habits and lifestyle of individuals, complementing or substituting data that is traditionally used by insurance companies. The collection of such data is highly concentrated with large technology and e-commerce companies such as Alibaba, Alphabet (Google), Amazon, Apple, Baidu, Facebook, Microsoft or Tencent. Google has about a 90 per cent market share in searches, while Facebook has a penetration of about 89 per cent of Internet users.<sup>6</sup> Alphabet, Amazon, Apple, Facebook and Microsoft are among the companies with the highest market capitalisation worldwide, at least partially based on the value of their customer data as an asset.

1 Franklin, J. (2001) "The Science of Conjecture: Evidence and Probability before Pascal". Baltimore, MD.: The John Hopkins University Press.

2 Kopf, E.W. (1927) "The Early History of the Annuity". New York: Lawrence, p. 248ff.

3 In this report, the terms 'data' and 'information' are used interchangeably. See the glossary for a definition of terms.

4 See appendix for an overview of data types traditionally used in insurance. There we also list additional types of data which become available in the context of big data.

5 Clarke, R. and Libarikian, A. (2014): "Unleashing the value of advanced analytics in insurance". McKinsey & Company, <https://www.mckinsey.com/industries/financial-services/our-insights/unleashing-the-value-of-advanced-analytics-in-insurance>.

6 Zingales, L. and Rolink, G. (2017) "A Way to Own Your Social-Media Data". New York Times, June 30, 2017. Market shares probably refer to the U.S. market, although this is not specified.

The second new source of data stems from sensors built into appliances and other consumer goods in the Internet of Things, for example sensors built into cars (telematics) or wearables, data from smart homes or drones. This data is generally fragmented and specific to a particular real-life purpose.

The emergence of big data analytics and artificial intelligence has triggered an arms race in the development of new applications along the entire insurance value chain, both by InsurTech startups and established insurers. Broadly speaking, new applications have typically focused on one of the following areas:

*New distribution models:* New applications revolutionise customer interaction by means of virtual assistants, digital brokers, chatbots and robo-advisers and use big data and artificial intelligence for enhanced customer segmentation, targeted marketing and dynamic pricing.

*Process automation:* Such applications aim to automate or improve efficiency of internal processes with big data and artificial intelligence. Straight-through processing enables the automation of parts of the value chain or even the entire value chain, potentially including underwriting, claims handling, risk management, finance and investment management as well as regulatory reporting and compliance.

*New propositions:* New data applications enable the development of new products and alternative business models, including peer-to-peer insurance, on-demand insurance, usage-based insurance, as well as insurance products covering new types of risk.

The true potential of the new technologies, however, unfolds with the combination of the different elements into a seamless digital infrastructure. The continuous collection and analysis of behavioural data allows for individual and dynamic risk assessment and the establishment of a continuous feedback loop to customers, with no or limited human intervention. Such digital monitoring not only enhances the quality of risk assessments but can also provide real-time insights to policyholders on their risk behaviour and individual incentives for risk reduction.

Moreover, the combination of new data sources paves the way for the implementation of advanced risk management systems that use predictive analytics as a basis for early intervention and risk prevention.<sup>7</sup>

Such powerful new business model recombinations are already being launched or are clearly visible on the horizon.<sup>8</sup> They include genuine peer-to-peer concepts (such as Bought by Many) and fully digital insurers (such as Oscar, InShared, Haven Life or Sherpa, for example). Ultimately, they will enhance the role of insurance from pure risk protection towards “predicting and preventing”.<sup>9</sup>

7 An example for such an integrated risk management system is the “Together for Safer Roads” coalition. This coalition, which includes private sector companies from different industries and insurer AIG, works with three cities (Atlanta, Sao Paolo and Shanghai) to identify and address the cities’ strategic road safety challenges. In this coalition, companies and public stakeholders share data and expertise to determine road safety challenges and potential solutions based on advanced analytics. See <http://www.togetherforsaferroads.org/>.

8 Braun, A. and Schreiber, F. (2017) “The Current InsurTech Landscape: Business Models and Disruptive Potential”. Institute of Insurance Economics of the University of St. Gallen in cooperation with Swiss Re Institute.

9 See appendix for an overview of some new uses of data in insurance.

## 3. Economic, societal and customer benefits

These new business models have the potential to generate great economic and societal benefits:<sup>10</sup>

### Risk reduction and loss prevention

In many instances, better aligning premiums and risk has clear economic and societal benefits. It allows premiums to signal risk and encourages risk reduction. By establishing a feedback loop to policyholders, digital monitoring allows them to reduce risk by adapting their behaviour.<sup>11</sup> Moreover, enhanced data facilitates the establishment of advanced risk management and early-warning systems that allow for timely interventions to reduce losses and lead to additional benefits for policyholders.

### Cost reductions

A key feature of insurance markets is the prevalence of two types of informational asymmetries: moral hazard and adverse selection. They represent a market inefficiency and imply that insurers must invest considerable resources in assessing the risks of their contractual partners and verifying information provided by policyholders. In fact, a considerable fraction of premiums is spent on claims handling, acquisition and administration.<sup>12</sup>

Accordingly, a considerable amount of employee time is spent on processing data. There is therefore a great potential for automation of data processing. McKinsey estimates the automation potential to be 43 per cent of the time spent by finance and insurance employees.<sup>13</sup> In non-life insurance, insurance fraud alone consumes almost 10 per cent of premiums.<sup>14</sup>

Automation therefore has the potential to considerably enhance market efficiency and lower costs by reducing informational asymmetries. In a competitive market environment, this will ultimately be reflected in lower

premiums, boosting affordability and coverage and contributing to narrowing the protection gap.<sup>15</sup> Moreover, better estimation of distribution functions at the portfolio level through big data and artificial intelligence allows insurers to charge lower premiums by reducing the risk load.

### New and enhanced products

Data that is more granular allows insurers to offer products that are tailored to the needs of the insured, including insurance on demand or pay-as-you-use propositions. Such usage-based insurance ensures that consumers pay based on the actual risk, e.g. when they drive as opposed to when the car stays in the garage.

Better understanding of risks also facilitates the development of new types of coverage and enhances the insurability of existing and emerging risks (such as cyber risk, for example). The enhanced use of data may also enable insurers to develop insurance products for high risks that so far could not be insured. For example, patients suffering from previously uninsurable diseases could share data related to their physical condition and benefit from individualised care offers.<sup>16</sup>

To sum up, the societal and economic benefits of the enhanced use of data are highest in business lines in which:

- the cost of moral hazard and adverse selection is high,
- there is great potential for risk reduction through mitigation and prevention, and/or
- there is a high degree of underinsurance.

An overview of important new uses of data in insurance and their key economic and societal benefits can be found in the appendix.

10 Schanz, K-U. and Sommerrock, F. (2016): "Harnessing Technology to Narrow the Insurance Protection Gap", The Geneva Association, Zurich, December 2016.

11 At least as long as consumers know how to adapt their behaviour to reduce risk and their premiums.

12 Schanz, K-U. and Sommerrock, F. (2016): "Harnessing Technology to Narrow the Insurance Protection Gap", The Geneva Association, Zurich, December 2016.

13 McKinsey Global Institute (2017) "What's now and next in analytics, AI, and automation". Executive briefing, May 2017, Exhibit 6. The data refers to the U.S., but it is reasonable to assume that this is true for other regions as well.

14 World Economic Forum (2015) "The Future of Financial Services". [http://www3.weforum.org/docs/WEF\\_The\\_future\\_of\\_financial\\_services.pdf](http://www3.weforum.org/docs/WEF_The_future_of_financial_services.pdf).

15 Schanz, K-U. and Sommerrock, F. (2016): "Harnessing Technology to Narrow the Insurance Protection Gap", The Geneva Association, Zurich, December 2016.

16 CRO Forum (2017): "Big Data & Privacy: unlocking value for consumers—CROs in a changing environment". <https://www.thecroforum.org/wp-content/uploads/2017/12/CROF-Big-Data-and-Privacy-Final-2017-12-14.pdf>.

## 4. Ethical and societal concerns with the enhanced use of personal data

The above-mentioned benefits do not come without costs. There is a general concern that consumers do not share the benefits from enhanced use of data, or that they come at a disproportional cost to consumers, specific consumer groups or society at large.

This section reviews ethical and societal concerns that have been raised in regulatory and public debate, and identifies the key trade-offs involved.

Ethical and societal concerns can be grouped into three broad categories:

- a) Concerns about *privacy* and *data protection*.
- b) Concerns around an increasing *individualisation of insurance*.
- c) Concerns about the implications of big data and artificial intelligence for *competition*.

Such concerns are not new. Indeed, most jurisdictions have public policies or regulatory frameworks in place to deal with many, if not all, of those issues. For example, existing data protection regulations govern the collection and use of personal data by insurers in most Western countries based on Fair Information Principles.<sup>17</sup> In addition to (horizontal) privacy regulation, in many jurisdictions insurance regulation restricts the use of certain information (such as race, gender, genetics etc.) as underwriting factors in order to address concerns about discrimination. Also, to address concerns about competition, competition policy aims to maintain a competitive marketplace, and in many countries insurance regulators have an explicit mandate to ensure a competitive insurance marketplace.

Nevertheless, these worries are likely to become more prominent in the era of big data and artificial intelligence, and some of them may also develop a novel twist.

In the following, we will explore how new business propositions that use enhanced data technologies affect these concerns and also identify the trade-offs involved.

### Concerns about privacy and data protection

Privacy and data protection concerns relate to issues like fairness and discrimination, intrusiveness and the right of (informational) self-determination, as well as the contextual integrity of personal data.<sup>18</sup>

#### *Fairness and discrimination*

It has been argued in many places that the profiling or scoring of customers can undermine fairness and create discriminatory effects.<sup>19</sup> Such concerns seem particularly relevant in insurance as profiling or scoring—i.e. the establishment of an individual risk score or risk profile—forms an inherent feature of the insurance business model.

The term ‘discrimination’ is used differently in the context of different social sciences. In economics, for example, it typically has no moral connotation and is used to describe a differential treatment, whether this is a good or bad thing. In the legal literature, by contrast, ‘discrimination’ is usually regarded as something that is wrongful.<sup>20</sup> Here, we use the term in a normative way implying that “those who should be treated equal are not”. Avoiding discrimination thus implies that certain differences (such as gender, race, sexual orientation, etc.) should be ignored. On the other hand, however, ‘discrimination’ also includes instances in which “those that should be treated differently are treated the same”.<sup>21</sup>

This definition of discrimination reveals a fundamental dilemma in the context of insurance. On the one hand, insurance customers may be treated according to their individual risk, but doing so implies that protected groups may be disadvantaged if their risk is higher than the average.<sup>22</sup> On the other hand, not treating individuals

<sup>17</sup> See section 5, “Balancing benefits and risks”, in particular footnote 56.

<sup>18</sup> Data security, another important concern, is beyond the scope of this report.

<sup>19</sup> See e.g. Zarsky, T.Z. (2014) “Understanding Discrimination in the Scored Society”, *Washington Law Review* 89(4).

<sup>20</sup> “Discrimination”, *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/discrimination/>.

<sup>21</sup> Zarsky, T.Z. (2014): “Understanding Discrimination in the Scored Society”, *Washington Law Review*, 89(4).

<sup>22</sup> For example, if zip codes correlate with race, using zip codes as risk factors may result in redlining. Block, W., Snow, N., and Stringham, E. (2008) “Banks, Insurance Companies, and Discrimination”, MPRA Paper 26035, University Library of Munich, Germany. <https://ideas.repec.org/p/prapa/26035.html>.

according to their individual risk could lead to the risk classification being considered unfair, as it treats individuals the same even if their risk is not.<sup>23</sup>

There is no easy solution to this dilemma. Simply eliminating discriminatory factors such as gender, race, etc. from the data does not do away with potential disparate impact, as an algorithm would easily infer these features from other factors ('blatant proxies'). For example, online media easily allows inference of an individual's gender, ethnicity, nationality, sexual orientation or other personal information.

In computer science, different approaches have been developed to assess and correct for disparate impact of automated decision-making.<sup>24</sup> However, these approaches have in common that they come at the cost of reduced accuracy of risk classification.<sup>25</sup> In insurance, inaccurate risk classification may not only be perceived as unfair but also has broader implications for efficiency and welfare by reducing the role of premiums as a signal of risk.

It is therefore necessary to strike a difficult balance between the accuracy of risk assessments and the potential for disparate impact on different social groups.<sup>26</sup> How to balance this trade-off will depend on the cultural context and the type of risk considered, among other factors. For example, disparate impact may not be considered an issue if the risk is mainly within the control of the individual, or if all groups (including high risks) benefit from absolute premium reductions, even though to different degrees. In any case, insurers should test and assess algorithms for potential disparate impact.<sup>27</sup>

### ***Intrusiveness and interference with the right of (informational) self-determination***

Privacy is often considered a right and a value in itself.<sup>28</sup> This right includes an individual's control over their digital identity and an individual's right of (informational) self-determination.

The use of big data and automated decision-making may be perceived as interfering with the right of self-determination of individuals. Business models based on digital monitoring reward or penalise certain behaviours or lifestyle choices that are deemed 'good' or 'bad' by the insurance company. While such business models have great potential to reduce risks by triggering behavioural change, at the same time they may be considered intrusive or 'paternalistic' and as interfering with an individual's independence in decision-making. Such intrusiveness may be considered particularly problematic if individuals cannot afford to pay insurance for high-risk behaviour and are thus restricted in their lifestyle choices.

One way to strengthen self-determination would be to assign individuals property rights for their personal data.<sup>29</sup> Individuals would be free to sell access to their personal data based on their individual cost-benefit assessment.<sup>30</sup> From the perspective of economic theory, assigning clear property rights certainly has its merits and could lead to the emergence of information markets.<sup>31</sup> In practice, however, such information markets have not yet emerged. Furthermore, assigning property rights to individuals would not do away with the possibility of coercion, since it would

23 For example, in traditional car insurance policies, young drivers typically pay a higher premium, independent of their actual driving behaviour. While young drivers are on average responsible for higher losses, this must not be true for all individuals in a specific age class. Telematics enhances fairness by taking actual driving behaviour into account.

24 See e.g. Pedreshi, D., Ruggieri, S. and Turini, F. (2008) "Discrimination-Aware Data Mining", in Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, pp. 560–568; Calders, T., Kamiran, F., Pechenizkiy, M. (2009) "Building Classifiers with Independence Constraints", IEEE International Conference on Data Mining Workshops; and Feldman, M., Friedler, S.A., Moeller, J., Scheidegger, C., Venkatasubramanian, S. (2015) "Certifying and removing disparate impact", arXiv:1412.3756 [stat.ML], Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

25 Berk, R., Heidari, H., Jabbari, S., Kearns, M. and Roth, A. (2017) "Fairness in Criminal Justice Risk Assessments: The State of the Art", ArXiv:1703.09207 [Stat ML], March. <http://arxiv.org/abs/1703.09207>.

26 A disparate impact could also result from 'tainted datasets', i.e. if the data collection itself was not neutral. Such biased data would also reduce the accuracy of risk assessments.

27 Zarsky, T.Z. (2014) "Understanding Discrimination in the Scored Society", Washington Law Review, 89 (4).

28 See e.g. article 8 of the European Convention on Human Rights. ([http://www.echr.coe.int/Documents/Convention\\_ENG.pdf](http://www.echr.coe.int/Documents/Convention_ENG.pdf)).

29 Varian, H.R. (2002) "Economic Aspects of Personal Privacy". In: Lehr W.H., Pupillo L.M. (eds) Cyber Policy and Economics in an Internet Age. Topics in Regulatory Economics and Policy Series, Vol. 43, Boston, MA., Springer.

30 Privacy regulations based on notice and consent come close to assigning property rights to individuals in that they must consent to the use of their personal data. See the following chapter "Balancing benefits and risks".

31 Laudon, K.C. (1993) "Markets and Privacy", NYU Working Paper No. 2451/14257.

still be possible for certain groups to be economically dependent on selling their personal data.

Insurers must therefore strike a balance between risk reduction and intrusiveness. In any case, to promote trust and acceptance, insurers should ensure that incentives like premium rewards are based on objective and accepted criteria. A failure to do so would carry considerable reputational risks.

### **Contextual integrity**

The concept of contextual integrity postulates that personal information should flow in accordance with expected context-specific informational norms. According to the concept of contextual integrity, existing contexts in which activities are grounded shape expectations that, when unmet, cause anxiety and resistance.<sup>32</sup> Thus, violating contextual integrity will ultimately undermine trust.

For example, when interacting with their insurer, customers typically expect personal information to be treated confidentially, independent of whether the interaction is face-to-face, via telephone, or online. Likewise, when engaging in social interaction on a social media platform, individuals arguably do not expect such personal information to be used to determine insurance premiums.

The need to respect contextual integrity provides a limit for the secondary uses of personal data, as unwarranted secondary uses may be perceived as disturbing or objectionable by individuals. Thus, ensuring contextual integrity is likely to reduce the commercial value of personal data, and a trade-off ensues between the commercial value of personal data and maintaining its contextual integrity.

## **Concerns with increasing individualisation of insurance**

The changing role of data implies that the individual's premiums are no longer determined based on their

grouping in a specific risk class but on their risk profile. Such individual risk profiles allow for a more granular and accurate assessment of an individual's risk.

Three types of concerns are often mentioned regarding this development: high risks may no longer be able to afford risk cover and may be excluded from insurance; the principle of solidarity may be eroded; and consumers may face frequent changes in premiums, i.e. premiums could become more volatile.

### **Affordability and exclusion**

In some cases, greater premium differentiation could imply that insurance cover becomes unaffordable for high risks. This may raise social concerns, in particular if the risk is correlated with low income and low wealth. Such concerns are not new in the insurance industry, but they are likely to become more accentuated in the Digital Society. They may be considered particularly relevant for risks which individuals cannot avoid, where the costs of risk reduction would be unacceptably high, and when the premium represents a large fraction of disposable income.<sup>33</sup>

Academics have long advocated direct premium subsidies for those unable to afford insurance, because this allows the positive effects of premium differentiation to be maintained.<sup>34</sup> In practice, though, such approaches have rarely been implemented. Instead, some jurisdictions restrict the use of certain risk indicators or resort to direct rate regulation.

These approaches have considerable drawbacks as they distort the price mechanism, leading to economic inefficiencies, insufficient insurance coverage, or adverse selection.

Another approach regulators have taken is to create high-risk pools that are based on a distinct financing mechanism. Depending on their design—in particular on how they are funded and who ultimately bears the risk—such high-risk pools may also lead to considerable competitive distortions and distorted incentives.<sup>35</sup>

32 Nissenbaum, H. (2011) "A Contextual Approach to Privacy Online". *Daedalus*, Journal of the American Academy of Arts & Sciences 140(4).

33 An example of a risk which is beyond the control of the individual is genetic predisposition that raises the risk of a certain illness, as opposed to health risks that are the consequence of behavioural choices (e.g. dangerous sports). An example of a risk for which risk reduction is very costly is exposure to flooding, as risk reduction may require the dislocation of tenants or even entire cities. In practice, the distinction between controllable and uncontrollable risks is not always obvious.

34 See e.g. Kousky, C. and Kunreuther, H. (2014) "Addressing Affordability in the National Flood Insurance Program", *Journal of Extreme Events*, 01 (01): 1450.

35 An example of such high-risk pools are Coastal Wind Pools in the U.S., see Hornstein, D.T. (2016) "Lessons From U.S. Coastal Wind Pools About Climate Finance and Politics", 43 *B.C. Envtl. Aff. L. Rev.* 345.

In sum, there is a trade-off between individualisation of insurance and potential consequences for affordability of insurance for individuals considered to be high-risk.

### **Implications for solidarity and risk pooling**

Some commentators have mentioned that increasing individualisation threatens risk pooling as the fundamental premise of insurance<sup>36</sup>, and ultimately undermines solidarity as an inherent social role of insurance. Such claims often mix up the roles of solidarity and risk pooling in insurance.<sup>37</sup> Solidarity is an important motive for social insurance that ensures risks (particularly those that consumers cannot control) are shared by all members of society on equitable terms. In many jurisdictions, such equity considerations are at the root of social health insurance programmes and public pensions systems, for example.<sup>38</sup>

In contrast to social insurance, private insurance is not based on the premise of solidarity, but relies on private risk pooling.<sup>39</sup> Risk-averse individuals have a unilateral interest in engaging in such risk pooling. As long as individuals are charged proportionally to their respective risk, there is no cross-subsidisation involved in private insurance.<sup>40</sup>

The changing role of data does not threaten risk pooling as a fundamental device of insurance. As long as individual risks retain some level of uncertainty and are not predictable with certainty, risk pooling has a role to play, even when big data allows a much better assessment of the risks. It is true, though, that the better individual risks can be predicted, the lower the value of insurance for policyholders and hence the lower an individual's willingness to pay.<sup>41</sup>

Thus, there are two types of trade-offs to be considered. In the realm of social insurance, there is a trade-off between individualisation of insurance and equity. In private insurance, there is a trade-off between individualisation and value of insurance for consumers.

### **Premium volatility**

With dynamic risk assessment and greater individualisation of premiums, the premiums of an individual may vary over time together with changes in risk. While such risk-based pricing enhances actuarial fairness, increased premium volatility reduces the value of insurance for an individual and thus their willingness to pay. Insurers will therefore have to balance the frequency of premium adjustments with consumers' interest in a stable and predictable premium. While premium volatility does not represent a market failure per se, it may nevertheless raise concerns if it significantly reduces market transparency and the comparability of product offerings and their prices (see also section 'Market transparency', p.15).<sup>42</sup>

### **Concerns about competition**

Digitisation has resulted in the disruption of several industries, stirring up existing market structures and marginalising incumbent market players. Take the taxi industry, for example, where Uber gained large market shares within a short period of time.

Such fundamental transformations of industry structure are not unusual with the emergence of new and superior production technologies. They are a feature of the market economy and the process of creative destruction. Such shifts, however, are problematic if they are based on

36 See, e.g. The Economist (2015) "A tricky business", available at <https://www.economist.com/news/leaders/21646203-data-driven-underwriting-contains-great-promise-and-grave-perils-tricky-business>

37 Definitions of solidarity typically emphasise the feeling of belonging to a particular group or community and/or involve voluntary or involuntary transfers of wealth.

38 In some countries, this extends to the protection of assets against natural disasters.

39 Wilkie, D. (1997) "Mutuality and solidarity: assessing risks and sharing losses". In: *Philosophical Transactions of the Royal Society B: Biological Sciences*, London, 352(1357): 1029-1044. See also Reichel, L. and Schmeiser, H. (2017) "Digitales Monitoring in der Assekuranz". In Institut für Versicherungswirtschaft der Universität St. Gallen (ed.) "Assekuranz 2025: Quo vadis?" I-VW-HSG Schriftenreihe, Band 63.

40 Risk pooling does not require premiums for risks belonging to the pool to be equal, see Reichel, L. and Schmeiser, H. (2017) "Digitales Monitoring in der Assekuranz". In Institut für Versicherungswirtschaft der Universität St. Gallen (ed.) "Assekuranz 2025: Quo vadis?" I-VW-HSG Schriftenreihe, Band 63.

41 See Reichel, L. and Schmeiser, H. (2017) "Digitales Monitoring in der Assekuranz". In Institut für Versicherungswirtschaft der Universität St. Gallen (ed.) "Assekuranz 2025: Quo vadis?" I-VW-HSG Schriftenreihe, Band 63.

42 Thouvenin, F. (2016) "Dynamische Preise—Eine Herausforderung für das Datenschutz-, Wettbewerbs- und Vertragsrecht" In Jusletter IT 22. September 2016.

predatory behaviour that aims to drive out competition in order to gain a position of market dominance, or made possible by an unlevelled playing field.

### **Abuse of market power**

Digital technologies are often characterised by strong network externalities that favour the emergence of oligopolistic or even monopolistic market positions. As mentioned above, global technology companies have a very large market share in their specific market segment which provides them with unique access to customers and their data.<sup>43</sup> Such dominant positions could be abused to extend monopoly to insurance markets or to extract value from existing market players through abusive prices.

This is not a mere theoretical possibility. In fact, there has been an increasing number of related antitrust cases in recent years. For example, in June 2017 Google was fined a record EUR 2.42 billion by the EU competition authorities for abusing search engine dominance.

### **Unlevelled playing field**

Disruption may also be considered unfair if it is made possible by the exploitation of regulatory differences (regulatory arbitrage).

In fact, large technology companies that operate at the global level have found ways to circumvent local data protection requirements, e.g. by choosing a location for their headquarters that has less stringent requirements.<sup>44</sup> Insurers, by contrast, are required in many cases to maintain a physical presence in the country and have to abide by local data protection regulations, preventing them from large-scale data collection.<sup>45</sup>

Such technology companies may decide to use data collected in the past to enter the insurance market. Even if such market entry required technology companies

to be licensed as insurers, data collected in the past could provide them with a competitive advantage over traditional insurers. The use of data collected in the past would also conflict with the concept of contextual integrity, as consumers most probably did not expect their personal data to be used for insurance purposes.

It is questionable whether competition policy alone will be sufficient to prevent such practices and ensure a competitive marketplace. As a legislation dealing with abuse, competition policy kicks in only after the fact, i.e. when the abuse has already happened and precedents have been created. Furthermore, cases of abuses of market power typically take several years, and the standards for proof are high. To promote a competitive marketplace, legislators may therefore choose to enforce *ex-ante* data access rules.

An example of such access rules are the data portability requirements introduced by the EU General Data Protection Regulation (GDPR) which give consumers the right to obtain a copy of their data that they can share with their preferred supplier. To what degree data portability will be successful in promoting competition, however, remains to be seen, as related standards for data exchange have not yet been developed.<sup>46</sup>

### **Market transparency**

Regulators have voiced concerns about the potential reduction of comparability of increasingly personalised products, which would make it harder for consumers to compare different offers and thus reduce competition.<sup>47</sup> Certainly, the proliferation of new business and pricing models renders buying decisions more complex. On the other hand, however, digital tools such as comparison platforms or virtual assistants can help to create transparency for customers. Nevertheless, insurers should strive to provide transparency about their products to customers.

43 See chapter 2, "The changing role of data in the insurance business model".

44 For example, in May 2016 the administrative court of Hamburg decided that German data protection requirements are not applicable to Facebook whose European headquarters are located in Ireland. "Social Media und Recht", 23. Mai 2016. The General Data Protection Regulation (GDPR) of the EU, entering into force in May 2018, will require all companies dealing with EU customers to adhere to privacy regulations, irrespective of their location.

45 See Reichel, L. and Schmeiser, H. (2017): "Digitales Monitoring in der Assekuranz". In: Institut für Versicherungswirtschaft (ed.), "Assekuranz 2025: Quo vadis? ". I-VW-HSG-Schriftenreihe, Band 63.

46 CRO Forum (2017) "Big Data & Privacy: unlocking value for consumers—CROs in a changing environment", available at <https://www.thecroforum.org/wp-content/uploads/2017/12/CROF-Big-Data-and-Privacy-Final-2017-12-14.pdf>.

47 See e.g. European Supervisory Authorities (2016) Joint Committee Discussion Paper on the Use of Big Data by Financial Institutions, JC 2016 86.

## 5. Balancing benefits and risks

Balancing the various trade-offs discussed in the previous chapter requires difficult value judgements by consumers, firms, policymakers and regulators alike.<sup>48</sup> What makes balancing these trade-offs difficult is that they are context-specific, often ambiguous and sometimes intangible.<sup>49</sup> Yet, finding an appropriate balance between privacy protection and allowing for innovation are of fundamental importance, as insufficient privacy protection will harm consumers and erode trust, while overly strict regulation may hinder society from reaping the benefits of data.

There is no 'one-size-fits-all' approach, and new uses of data therefore require a case-by-case assessment of their respective benefits and risks. In order to inform this debate, we have summarised the most salient propositions and their most important benefits and privacy risks in the appendix.

For consumers, individual cost/benefit assessments of sharing personal data are becoming increasingly complex in the digital era. It is difficult, if not impossible, for individuals to assess potential consequences of sharing personal data in different contexts. Furthermore, in many cases individuals are confronted with a decision to either "take or leave it".<sup>50</sup> Some commentators have therefore argued that the principle of notice and consent, a key principle of privacy regulation in most Western countries, has failed to ensure privacy in the digital era.<sup>51</sup>

Empirical studies of individuals' privacy choices conclude that privacy choices appear to be inconsistent, highly context dependent, and affected by biases. There are also considerable variations between countries, and many studies have highlighted the dichotomy between self-

professed privacy attitudes and actual self-revelatory behaviour ('privacy paradox').<sup>52</sup> According to a recent survey, more than two-thirds of businesses and consumers said they are willing to share data if they perceive some benefit. There are, however, considerable differences between countries and industries, both in terms of willingness to share data and regarding the benefits of sharing data.<sup>53</sup>

Therefore, the principle of notice and consent—which is attractive because it comes close to assigning property rights on personal information to individuals<sup>54</sup>—may not always ensure an appropriate balance. Insurers therefore have a heightened responsibility in ensuring that uses of personal data are transparent and in line with reasonable expectations of consumers, which differ across regions and countries. A failure to do so would create considerable reputational risks for insurers and erode trust.

Policymakers and regulators are confronted with the question of whether the existing regulatory frameworks are adequate to ensure an appropriate balance. Existing data protection laws date back to the 1970s, reflecting concerns about the emergence of computer and communication technologies, with their ability to remotely process large volumes of data.<sup>55</sup> They stem from a time without Internet, smartphones and the Internet of Things.

Commentators have pointed out that some of the principles embodied in the OECD Fair Information Practices of 1980<sup>56</sup>— which underlie privacy regulation in most Western countries— are in conflict with the nature of big data, and unduly restrict innovation. For example, the principles of data minimisation and purpose specification

48 An overview of the trade-offs is provided in the appendix.

49 Acquisti, A., Taylor, C. and Wagman, L. (2016) "The Economics of Privacy". In *Journal of Economic Literature* 54(2).

50 See e.g. Thouvenin, F. (2017) "Wem gehören meine Daten? Zu Sinn und Nutzen einer Erweiterung des Eigentumbegriffs". In: *SJZ* 113(2), pp. 21–32.

51 Nissenbaum, H. (2011) "A Contextual Approach to Privacy Online". In: *Daedalus, Journal of the American Academy of Arts & Sciences* 140 (4).

52 Acquisti, A. (2009) "Nudging Privacy: The Behavioral Economics of Personal Information", *IEEE Security & Privacy*, November/December 2009.

53 AIG (2017) "The Data Sharing Economy: Quantifying Tradeoffs that Power New Business Models". *RISK + INNOVATION | PART 3 IN A SERIES*, <https://www.aig.com/content/dam/aig/america-canada/us/documents/insights/aig-the-data-sharing-economy.pdf>.

54 See Thouvenin, F. (2017) "Wem gehören meine Daten? Zu Sinn und Nutzen einer Erweiterung des Eigentumbegriffs". In: *SJZ* 113(2), pp. 21–32. The European General Data Protection Regulation (GDPR) strengthens the rights of individuals over their data by requiring 'unambiguous consent' for processing personal data and by introducing the right to be forgotten and data portability.

55 UNCTAD "Data protection regulations and international data flows: Implications for trade and development", New York and Geneva, 2016.

56 OECD 1980 "Guidelines on the Protection of Privacy and Transborder Flows of Personal Data". Core principles include collection limitation ('notice and consent'), purpose limitation, data minimisation and transparency (or 'openness'). In Europe, these principles—among others— are being strengthened by the General Data Protection Regulation (GDPR) which comes into effect starting in May 2018.

are difficult to reconcile with the prospect of big data analyses. At the time of data collection it might not be clear which data is useful for which purposes, making it hard to strike a balance between minimising data collection and providing room for innovation.<sup>57</sup>

To comply with the purpose specification rule, entities striving to engage in big data analysis will need to inform their data subjects of the future forms of processing they will engage in (which must still be legitimate by nature) and closely monitor their practices to assure they do not exceed the permitted realm of analyses. Carrying out any one of these tasks might prove costly, difficult, or even impossible.<sup>58</sup>

In practice, much depends on how these principles are applied. In fact, despite the Fair Information Principles, there exist substantial differences in data privacy or data protection legislation between different regions and countries, and “there is no single agreed model for data protection law at this stage.”<sup>59</sup>

---

57 CRO Forum (2017) “Big Data & Privacy: unlocking value for consumers—CROs in a changing environment”. <https://www.thecroforum.org/wp-content/uploads/2017/12/CROF-Big-Data-and-Privacy-Final-2017-12-14.pdf>

58 Zarsky, T.Z. (2017) “Incompatible: The GDPR in the Age of Big Data”. In: *Seton Hall Law Review*, Vol. 47, No. 4(2), 2017.

59 See UNCTAD: “Data protection regulations and international data flows: Implications for trade and development”, New York and Geneva, 2016.

## 6. Potential future scenarios

In this chapter, we develop five stylised future scenarios for the evolution of the insurance industry.<sup>60</sup> The scenarios are not meant to be predictions of the future development of the industry. Rather, they serve to identify the likely consequences of different policy choices for the industry's future development. In particular, we will discuss the implications for competition, for economic and societal benefits and for privacy for each scenario.

In all scenarios, regulation, technology and consumer behaviour are treated as external factors that determine—in dynamic market interaction—the market outcome, including market size, industry structure, degree of competition and the level of innovation, as well as societal welfare, including consumer and firm surplus and the level of privacy. An overview of the scenarios and their implications is provided in the appendix.

The scenarios differ along three key dimensions: access to data, privacy regulation and the distribution of information between consumers and companies. To assess the economic benefits and consequences for the privacy of each scenario, we have developed an overview of the most salient propositions and their most important benefits and privacy risks (see appendix). Depending on which propositions are feasible in a given regulatory setting, one can aggregate the benefits and privacy risks of these provisions to assess the consequence of policy choices.

### **Access to data**

It is assumed that either all market players have access to new sources of data, or that only new market players (technology companies, Internet of Things data collectors) have access, while traditional insurers do not.<sup>61</sup>

Access to new sources of online and Internet of Things data will have an important impact on the insurance industry. Such an impact can come from two directions:

- New forms of risk classification can lead to advantages in risk selection and cost-adequate pricing. Given a competitive market, companies with higher abilities in this field are in the lead. If the market is not fully competitive, a better knowledge of the underlying cost structure allows companies to focus more on profitable products and customer segments. Moreover, insurance companies face better information in order to derive profit-maximising price and/or quantity combinations.
- Since risk classification is already a very well-established discipline in the insurance sector, an even more important point could be that the future needs of customers and their willingness to pay for differentiated types of insurance coverage and product features become much more transparent to the provider.

### **Privacy regulation**

A second determining factor for the scenarios is the regulatory framework. We will analyse three cases: no restrictions on the use of data, strict regulation that impedes the use of data, and intermediate cases in which regulation addresses some specific privacy concerns.

60 These scenarios have been developed in a joint workshop with the Institute of Insurance Economics of the University of St. Gallen and the University of Zurich.

61 For the purpose of these scenarios, technology companies include companies that collect large amounts of online media data and possess the ability to analyse and derive insights from this data (e.g. Amazon, Apple, Facebook, Google, and Microsoft). Internet of Things data collectors include car manufacturers, smart home providers, and providers of wearables, for example.

### **Distribution of information between individuals and firms**

For any scenario, the informational distribution between providers (insurance company or technology companies) on the one hand and policyholders on the other is central. If policyholders have access to tools that allow proper risk classification, an absorption of the policyholders' willingness to pay is hard to obtain.<sup>62</sup> In such a scenario and given a competitive market, not many incentives are given to join the market. This holds true for insurers, but also for technology companies.

In the following, we discuss five scenarios:



#### **Scenario 1: The digital society**

This scenario is characterised by free flow and open access to data. Insurance and technology companies have equal access to a broad range of data and can use it without restriction.



#### **Scenario 2: Insurance at two speeds**

In this scenario, insurance companies are prevented from access to or use of enhanced data to which technology companies have access. Those who have access to data can use it without restrictions. In a sub-scenario, only a few insurers have access to data through exclusive cooperation agreements with technology companies, but other insurers do not.



#### **Scenario 3: Privacy regulation**

In this scenario, regulators intervene to protect certain privacy values. In one sub-scenario, regulators try to prevent discrimination at all cost. In another sub-scenario, regulators aim to avoid intrusiveness. In a third scenario, regulators apply a zero-tolerance approach to the risk of abuse in using personal data.



#### **Scenario 4: Digital backlash**

In this scenario, we assume that increasingly restrictive regulation prevents established insurers and new market players from the use of enhanced data in insurance.



#### **Scenario 5: A tale of trust**

In this scenario, people are no longer willing to share their private data such as health-related information with technology companies in general or social networks in particular. Insurance companies can act as a 'safe harbour', but face similar conditions to the rest of the industry with regard to accessing data.

<sup>62</sup> For example, dedicated service providers may provide individual risk insights to customers. Or genetic tests may provide consumers with detailed risk information.



## Scenario 1: The digital society



### Scenario description

This scenario is characterised by free flow of data. Consumers are in general willing to share their data for a respective benefit or for free (for instance via social media). They feel confident that personal data is kept safe—or they may not really care about the security of personal data. Insurance companies have access to the same data as technology and Internet of Things companies at very low cost, either through their own applications (such as telematics devices, wearables, smart home devices etc.), through cooperation agreements with data collectors or by means of data portability.

There are no notable regulatory restrictions on the use of data. Insurers thus use the available data for purposes that promise the highest risk-adjusted return, including risk selection and pricing, first degree price discrimination, and to identify uncovered needs. Only insurance companies that use all cost-relevant information can survive a fully competitive environment. BigTech or InsurTech companies may enter the insurance market as risk carriers, in which case they will be subject to insurance regulatory requirements such as risk-based capital standards, investment regulation and other requirements.

Two cases must be distinguished in this scenario, depending on whether only firms have access to insights generated from data, or whether consumers also have access to these insights.



### Implications for competition

The availability of new data sources reduces market entry barriers in that new players do not need proprietary data, e.g. loss histories, to enter the market. Insurance regulatory requirements, however, still work as market entry barriers. While economies of scale and scope in data handling and analysis may trigger industry consolidation, there may be increasing competition from customer-driven business models such as peer-to-peer insurance and captive insurance. The threat of market entry is likely to reduce margins and prevent market players from raising rates above the competitive level, as is in fact already happening.

Lower margins will reduce the expected risk-adjusted return on investment from entering the insurance business and thus reduce the attractiveness for BigTech companies to enter the market as fully integrated risk carriers.



### Economic and societal benefits

This scenario allows for maximum risk reduction. In motor insurance, for example, the number of accidents, injuries and casualties is reduced by incentivising prudent driving behaviour and the implementation of integrated prevention and emergency response systems.

Likewise, in life and health insurance, integrated systems enable improved health management e.g. in the management of chronic diseases. The use of wearables data allows the monitoring of compliance with health treatments as well as measurement of their outcome, and thus leads to continuous quality improvement in health care.

The ability of insurers to automate the provision of insurance through the use of personal data drives significant cost reductions. Low cost of insurance and attractive insurance products that match individual needs enhance coverage, thereby narrowing the protection gap.

If consumers have access to risk insights from their own data, i.e. they have the ability to understand their risk profile, adverse selection and moral hazard can be significantly reduced, if not eliminated. This has considerable economic and societal benefits and represents a 'first-best' solution from an economic perspective.

If, on the other hand, consumers do not have access to their own data or are unable to derive risk insights from it, insurers would have an informational advantage over consumers. Insurers may be able to use their advantage to appropriate consumer surplus by charging premiums according to consumers' willingness to pay (first degree price discrimination). The ability to do so, however, depends on the degree of competition, as charging premiums above cost would provide opportunities for market entrants to lure customers away.



### *Implications for privacy*

This scenario raises several privacy concerns.

While the increased use of personal data may enhance the fairness of risk classifications, there is a risk that not all societal groups will equally benefit. The fact that some uses of data will have a disparate impact on different social groups raises concerns about discrimination.

Furthermore, as insurers engage in differentiated pricing, individuals may pay a different premium, even if the underlying risk is the same. Such differentiated pricing, even though very common in other industries, may be regarded as unfair discrimination, particularly if the purchase of insurance is compulsory or if the premium is a considerable fraction of the individual's income. However, to the extent that competition increases under this scenario, the ability of insurers to exert first-degree price discrimination actually decreases.

A system of rewards and penalties for different lifestyle decisions implemented by the insurance industry may be considered as intrusive, conflicting with self-determination, and lacking legitimacy. This may be particularly relevant if such rewards and penalties are based on non-transparent and not generally accepted criteria.

Violation of contextual integrity and (accidental) dissemination of personal data may lead to instances of stigmatisation and may imply that individuals refrain from personally or socially beneficial activities.

Finally, greater premium differentiation implies that high risks may be impacted negatively through higher insurance premiums. In an extreme case, insurance cover may become unaffordable for high risks.



## Scenario 2: Insurance at two speeds



### Scenario description

In this scenario, insurers do not have access to online or Internet of Things data collected by technology companies. Technology companies use their data assets to enter specific insurance lines or market segments as risk carriers. As gatekeepers of personal data, they have a competitive advantage over traditional insurers who would find it difficult to compete.

Such a scenario may also arise if consumers, who willingly accept data collection by technology providers, are unwilling to share the same data with established insurers.

In a sub-scenario, called 'Champions league', it is assumed that some insurers, most likely large global carriers, have access to data through cooperation agreements with technology providers, while smaller and local insurers would be excluded from such opportunities.<sup>63</sup>

It is assumed that there are no notable regulatory restrictions on the use of data. As in scenario 3, two cases need to be distinguished, depending on whether consumers do have access to insights about their risk or not.



### Implications for competition

Technology companies are likely to use their informational advantage over traditional insurers to enter market segments that promise the highest margins by specifically targeting low risks and offering them an attractive price ('cream-skimming'). As low risks migrate to market entrants, incumbent insurers see their portfolio deteriorate and their loss rates increase.

This could lead to a market dynamic by the end of which all customers migrate to the market entrant. Eventually, technology companies will dominate the market, and traditional insurers will see their role reduced to pure risk carriers without contact with end customers, or they will be driven out of the market entirely. This will reduce competition, and the market will be characterised by an oligopoly of large technology players or even by a monopolistic market structure.

However, this is not a necessary—or even likely—outcome of this scenario. Depending on consumers' willingness to share their data and the distribution of high and low risk types, there may exist a market equilibrium in which low risks willing to disclose their data insure with the market entrant, while high risks and low risks not willing to share their data insure with the traditional insurer.<sup>64</sup> Low risks willing to share their data may benefit from such an outcome, while low risks pay a price if they are not willing to share their data.

<sup>63</sup> Global technology players may seek cooperation with large insurers that allow the scaling of respective propositions at a global level, whereas such cooperation agreements could prove prohibitively costly for smaller insurers.

<sup>64</sup> Gemmo, Browne and Gründl (2017): "Transparency Aversion and Insurance Market Equilibria". ICIR Working Paper Series No. 25/2017. Preliminary version, February 2017.

Furthermore, the competitive advantage of technology companies may only be temporary, as traditional insurers learn the risk type of policyholders based on past claims experience and other information. Traditional insurers may also invest in technology to collect their own data. In this dynamic view, the distribution of market share between traditional insurers and technology companies will depend on the traditional insurers' speed of learning as well as the cost of digital technologies.<sup>65</sup>

In the 'Champions league' sub-scenario of scenario 2, the dividing line is not between insurers and technology companies but between insurers that have access to data and insurers that do not. Insurers that do not have the possibility to enter into a partnership with a technology player may be driven out of the market. This could lead to a reduction in competition if no other players enter the market.



### ***Economic and societal benefits***

Innovative propositions by technology players will allow for risk reduction. However, if traditional insurers coexist with technology players, such risk reduction will only be partial and lower than in scenario 1.

If the market is dominated by an oligopoly of large technology players, informational asymmetries may be significantly reduced, as in scenario 1.

In the case of the market outcome being characterised by the coexistence of traditional insurers and technology companies, informational asymmetries will persist in the traditional market segment. However, the cost of adverse selection and moral hazard may even increase if consumers have enhanced insights into their risk profile, while traditional insurers do not.



### ***Implications for privacy***

If technology companies dominate the market, similar privacy concerns arise as in scenario 1.

If traditional insurers coexist in the market, individuals have the choice between technology players with low levels of privacy protection and traditional insurers with higher levels of privacy protection. This, however, will come at a cost for individuals seeking privacy in terms of higher premiums.

<sup>65</sup> Eling, M. and Jia, R. (2017) "It's all about speed and costs: The impact of digital technology on the insurance market structure", Preliminary version, July 2017.



## Scenario 3: Privacy regulation



### Scenario description

In this scenario, it is assumed that the regulator intervenes to mitigate some of the negative implications of scenarios 1 and 2. In the following, we analyse three sub-scenarios in which regulation aims to avoid discrimination and intrusiveness or the risk of abuse of personal data, respectively. In each of these sub-scenarios, some propositions are still feasible while other are not (see appendix).

Different cases can be distinguished in this scenario, depending on whether traditional insurers do or do not have access to new sources of data and on whether consumers do or do not have access to insights about their risk.



### Sub-scenario 3a: Avoid discrimination

Regulators disallow uses of data that have a high risk of discrimination through having a considerable disparate impact on protected social groups. Only propositions with low likelihood of discrimination can be implemented. These are typically propositions that are based on the observation of actual risk-relevant behaviour, as opposed to relying on proxies.

In this sub-scenario, it will therefore be possible to implement propositions which are based on digital monitoring (telematics, wearables, smart home devices). However, propositions that rely on the use of proxy indicators from online media data for risk classification would be banned, as these proxies carry a risk of being discriminatory.



### Implications for competition

In this sub-scenario, new competition would mainly come from those who have access to risk-relevant Internet of Things sensor data. These include car manufacturers, health service or fitness providers and providers of smart home devices, for example, but may also include smartphone providers. These may bundle insurance together with their other product offerings, which would reduce competition if insurers did not have the ability to access such data.



### Economic and societal benefits

Propositions based on digital monitoring would allow for risk reduction benefits as well as cost reductions.



### Implications for privacy

Propositions based on digital monitoring, while reducing the risk of discrimination, may be considered particularly intrusive.



### Sub-scenario 3b: Avoid intrusiveness

Regulators disallow uses of data which are intrusive in the sense that they are based on some form of surveillance and on influencing the behaviour of individuals. In this sub-scenario, it would therefore not be possible to implement propositions that are based on digital monitoring (telematics, wearables, smart home devices). Conversely, it would be possible to mine online media data. This sub-scenario is thus the opposite case of the sub-scenario 'Avoid discrimination'.



### **Implications for competition**

Competition would in this sub-scenario come from big technology companies that collect large amounts of online media data. An oligopolistic or monopolistic market structure may emerge if traditional insurers do not have access to such data.



### **Economic and societal benefits**

While the use of online media data would allow an improvement of risk assessment and selection, risk reductions through influencing behaviour and through integrated risk management systems could not be realised. Improved risk assessments may lead to greater premium differentiation. In the extreme case, insurance may become unaffordable for individuals classified as high risks.



### **Implications for privacy**

While individuals would not be exposed to surveillance and monitoring of their risk behaviour, there is a high likelihood that new propositions would have a discriminatory impact. Furthermore, the pervasive mining of online media data for insurance purposes may violate the contextual integrity of personal data.



### **Sub-scenario 3c: Avoid risk of abuse of personal data**

In this sub-scenario, it is assumed that regulators aim to avoid the risk of abuses of personal data that may result from criminal activities or through accidental dissemination. As all uses of personal data carry the risk of hacking or other criminal abuse, zero tolerance for the risk of abuse implies that no innovative propositions can be realised. The implications of this scenario are like the ones described in scenario 4.



## Scenario 4: Digital backlash



### *Scenario description*

In this scenario, it is assumed that restrictive regulations preventing the use of enhanced data in insurance<sup>66</sup> and/or widespread resistance by consumers to share their data<sup>67</sup> makes it impossible to use enhanced data to offer insurance products. Such regulation may include restrictions on the use of indicators for underwriting, rate regulation, or other requirements that render the use of data impossible.

While it would not be possible to use enhanced data to offer insurance products, consumers may very well benefit from enhanced risk insights based on data analytics and artificial intelligence. As a result, the costs arising from asymmetric information may considerably increase in this scenario.

In sum, this scenario is likely to offer the lowest welfare compared to the scenarios discussed above.



### *Implications for competition*

The impossibility of using advanced data technologies would eliminate the threat of market entry by technology companies and new players. Such a scenario would therefore reduce the level of competition in insurance.



### *Implications for privacy*

The privacy of individuals is protected to a maximum degree. There are no instances of stigmatisation due to (accidental) dissemination of data on personal habits or behaviour or unwarranted secondary use of data (e.g. for marketing purposes). Individuals are in full control of their digital identity.



### *Economic and societal benefits*

In this scenario, the potential for risk reduction could not be realised in the insurance sector. The inability to automate and the absence of outside competition result in the persistence of elevated cost for insurance.

<sup>66</sup> For example, as a regulatory response to the concerns mentioned in chapter "Ethical and societal concerns with the enhanced use of personal data".

<sup>67</sup> For example, massive data leaks and widespread criminal abuse of personal data destroy consumers' trust in the ability of insurers to safeguard their data.



## Scenario 5: A tale of trust



### Scenario description

In this scenario, people are no longer willing to share their private data regarding health, lifestyle, travel, social status, personal relationships, etc. within social networks. Similarly, they refrain from posting pictures of their children and do not provide information on their current job situation, etc. There are many plausible reasons for such a development. One possible scenario, for instance, is that one of the large providers of social networking (such as Facebook, Twitter, etc.) is subject to a cyberattack. Consequently, sensitive user data is made publicly available. In such an extreme event, people might also lose trust in governments since they are unable to protect them appropriately. Insurance companies could therefore make use of such an event and establish themselves as a third party of trust. That is, they could act as regulated and trusted data managers that take over responsibility for their clients with respect to all third-party service providers.

One means of doing so could be to establish personal connections with customers through, for example, new and innovative products, providing incentives for healthy lifestyles, or issuing text alerts that inform customers on their premium payment, claims status, etc. As a recent study by Mintel shows, millennials in particular consider insurers more trustworthy than they were five years ago.<sup>68</sup> While 53 per cent of the millennial participants perceive them as trustworthy institutions, only 31 per cent of the overall population do so as well.

Another possible approach is to serve customers beyond their basic coverage. That is, insurers could position themselves as platform providers that offer ancillary services for customers' homes, cars, health and lifestyle. For instance, insurers could be integrated into auto sales and leasing, fitness club memberships, and other third-party partnerships such as flood monitoring in home basements via smart home devices.

Generally, to be successful, insurers are required to provide their customers with some central interaction and controlling tools. First, customers must be the ones that decide which data they want to share with their insurer or with any other party (e.g. through platform services). At the same time, they must ensure that every step in the communication process is made transparent via customer portals, apps, etc. so that customers can understand and check what data has been shared, for what purposes it is used, etc. Second, regarding communication, insurers are advised to ensure that all information is regularly updated and that customers are informed as soon as possible. Particularly in cases where the policy has changed or a data breach has occurred, a sound and proper communication strategy is needed to ensure that customers are kept up-to-date. Third, although most customers care about their (data) security and potential (cyber) attacks, their behaviour does not always match. Therefore insurers also need to invest in educating their customers on how to keep their data secure. In summary, insurers could make good use of their data access and data usage through providing additional services to their customers. Scenario 5 could also be supported by specific knowledge of the insurer when it comes to claims handling. For technology companies it is generally not easy to develop this expertise.

Besides these aspects, we would generally expect that a defined and regulated insurance market continues to exist. But instead of only offering insurance products and providing risk transfer and payments, insurance companies could indeed expand their business model and act as a general 'problem solver' for the policyholder based on additional data knowledge.

68 Mintel (2016) "Innovations in the Insurance Market – US—April 2016".



---

## Scenario 5: A tale of trust (continued)

---



### *Implications for competition*

The main economic implications are similar to the ones described in scenario 2; the main difference would be that insurers, rather than technology companies, would be in the lead. Competition might decrease if only large insurers are able to establish themselves as trusted data managers.



### *Economic and societal benefits*

This scenario offers large economic benefits in terms of risk reduction and decreased informational asymmetries.



### *Implications for privacy*

Analogous privacy issues emerge in this scenario as in scenarios 1 and 2.

## Discussion of scenarios

These scenarios are of course sketchy, and in reality, many nuances and 'shades of grey' are possible. The reality will most likely be reflected in a combination of all scenarios. Nevertheless, these scenarios allow the identification of different possible trajectories and the drivers for those trajectories.

The scenarios demonstrate the importance of building trust to enhance consumers' willingness to share their personal data. According to a recent survey, increased customer trust will likely translate into customers being more willing to share more data with their insurer.<sup>69</sup> Without consumers' willingness to share data with insurers, we will inevitably end up in 'Insurance at two speeds' or 'Digital backlash' scenarios, which offer reduced levels of competition, welfare and innovation. The need to earn consumers' trust may offer an alternative scenario for insurers to establish themselves as trusted data managers ('A tale of trust').

Furthermore, the scenarios highlight the importance of regulation, in particular policies regarding access to and use of personal data. These policies have to strike a difficult balance between ensuring privacy and promoting competition, innovation and welfare.

## Conclusions

New business models based on the enhanced use of data have great potential benefits in insurance by enhancing the efficiency of insurance markets, promoting risk reduction and mitigation, and by enhancing consumer choice and insurance coverage.

These benefits do not come without a cost, however, and there are complex trade-offs involved in the enhanced use of personal information. The discussion has shown that there is no easy, 'one-size-fits-all' approach to address the trade-offs involved with new business models. Rather, new business models need to be analysed on a case-by-case basis. This report discussed several scenarios that could arise as a consequence of different policy choices.

---

<sup>69</sup> IBM Institute for Business Value (2017) "Data: gold or kryptonite? An insurer's guide to the resource of the future". In association with the Institute of Insurance Economics of the University of St. Gallen, [https://www.ivw.unisg.ch/wp-content/uploads/2017/11/data\\_gold\\_or\\_kryptonite2017.pdf](https://www.ivw.unisg.ch/wp-content/uploads/2017/11/data_gold_or_kryptonite2017.pdf).

## Box 1: Telematics

### *What is it and how does it work?*

One of the earliest applications of enhanced use of data within the Internet of Things is telematics in car insurance. Telematics combines telecommunication and information technology into one solution that sends, receives, or stores information in a vehicle.<sup>70</sup> Several incumbent insurers and InsurTech startups have developed and launched products that are based on monitoring certain parameters of driving behaviour by means of sensors. This can be done either by accessing the built-in vehicle information system or by installing an on-board diagnostics (OBD) device ('black box') that is equipped with a SIM card to transmit data over the mobile network.

Typically, data that provides information on driving behaviour is collected, including geographic position, speed, acceleration and braking severity, vibration and impact events. This data—or part of the data—allows insurers to calculate a risk score. This risk score, combined with other data such as age of the driver, is used as the basis for pay-as-you-drive insurance propositions.

Telematics data may either be directly transmitted to the insurer or to a third-party telematics service provider which transmits a summary of the collected data to the insurer on a regular basis.

### *How is the data used?*

Telematics insurance propositions often provide an upfront premium discount and a cash-back discount at the end of the contractual period, depending on the risk score. Such propositions may also include value-added services such as emergency roadside assistance, automatic emergency crash response, stolen vehicle tracking, and interactive platforms that allow

consumers to review their driving behaviour online. Fleet management propositions provide additional services such as risk management, vehicle maintenance cost management and fuel consumption management.

### *What are the benefits?*

#### **Consumers**

Consumers can benefit by a considerable reduction of their premium. Consumers also benefit from information that helps them to improve their driving behaviour and become a better risk. In fact, telematics propositions are gaining increasing acceptance from consumers. Between 2015 and 2016, enrolment in a usage-based insurance program increased by 20 per cent in the U.S.<sup>71</sup>

#### **Society**

By providing information on driving behaviour, such propositions have the potential to lead to significant risk reduction. Typically, less than three per cent of telematics customers do not respond to feedback on their driving behaviour.<sup>72</sup>

### *What concerns does it raise?*

#### **Fairness and discrimination**

Telematics enhances the fairness of risk classification as it is based on actual driving behaviour. However, it may raise concerns about discrimination if, for example, driving in neighbourhoods considered unsafe correlates with belonging to a specific societal group.

#### **Interference with the right of self-determination**

Consumers may consider the insurer's monitoring of their driving behaviour intrusive, especially if it includes information on when and where they drive.

70 Technavio (2013) "What is telematics and why should we want it?", available at <https://www.technavio.com/blog/what-telematics-and-why-should-we-want-it>.

71 LexisNexis(2016) "2016 Usage-based insurance (UBI) research results for the U.S. consumer market", White Paper, August 2916. <https://www.lexisnexis.com/risk/downloads/whitepaper/2016-ubi-study-white-paper.pdf>.

72 Cavanagh, S., managing director of Wunelli and vice-president of LexisNexis, in Raconteur, "Future of Insurance", June 14, 2017.

### **Contextual integrity**

There may be a concern that insurers could use the data collected for other purposes. For example, an insurer may find that driving behaviour not only correlates with the risk of a car accident but with other types of risks as well. If this is the case, consumers may be concerned that their driving behaviour has an influence on the premium for other types of insurance cover.

### **Concerns about competition**

Car manufacturers are increasingly building telematics systems into standard car models. Ernst & Young expects the penetration of global integrated telematics for new cars to touch 88 per cent by 2025.<sup>73</sup> Car manufacturers could use the data recorded by these systems to offer car insurance policies as a package with the car itself. In fact, Tesla has already announced such plans.<sup>74</sup> This bundling could raise competitive issues if combined with market power.

---

73 EY (2013): "The quest for Telematics 4.0: Creating sustainable value propositions supporting car-web integration", [http://www.ey.com/Publication/vwLUAssets/The\\_quest\\_for\\_Telematics\\_4.0/\\$File/The\\_quest\\_for\\_Telematics\\_4\\_0.pdf](http://www.ey.com/Publication/vwLUAssets/The_quest_for_Telematics_4.0/$File/The_quest_for_Telematics_4_0.pdf).

74 Niklowitz, M. (2017) "Weniger zahlen dank Autopilot". In: Schweizer Versicherung, September 2017.

## Box 2: Use of social media data in underwriting

### *What is it and how does it work?*

Data shared over social networks such as posts, likes, pictures, videos, friendship connections etc. could be used to establish personal risk profiles for underwriting purposes.

While insurers generally have not systematically used social media data for underwriting purposes<sup>75</sup>, in November 2016 Admiral Insurance planned to launch a car insurance product that was based on analysing the Facebook accounts of first-time car drivers to look for personality traits linked to safe driving. Under the scheme, Admiral planned to identify personality traits through examining posts and likes by Facebook (although excluding photos) and an inspection of certain habits.<sup>76</sup> Facebook stopped the product due to privacy concerns shortly before its launch.<sup>77</sup>

### *How is the data used?*

Admiral said its algorithm looked for correlations between social media data and actual claims data. For example, Facebook users who write in short, concise sentences, use lists, and arrange to meet friends at a set time and place, rather than just 'tonight', would be identified as conscientious. In contrast, those who frequently use exclamation marks and phrases such as 'always' or 'never' rather than 'maybe' could be overconfident.<sup>78</sup>

As the number of customers increases and more evidence is gathered about correlations, the algorithm will evolve further, changing the importance of items identified on social media. The company would only have access to information gathered during the quote process and would have no ongoing access. Admiral would not have access to information about what its customers look at on Facebook or what their friends do.

### *What are the benefits?*

#### **Consumers**

Purchasers of the product would be offered discounts of up to £350 a year. Consumers may also benefit from a better customer experience and reduced information requests by the insurer.

#### **Society**

Such products may offer considerable cost reduction potential through automation. However, the use of social media for underwriting is very unlikely to provide additional societal benefits through risk reduction or mitigation, as it is not transparent to consumers how they need to adapt their behaviour to reduce their premium (if it was, consumers could 'game the system').

### *What concerns does it raise?*

#### **Fairness and discrimination**

The use of social media data in underwriting is unlikely to enhance fairness of risk classification, but it does raise considerable concerns about discrimination, which is why the launch has been stopped. For example, the use of certain words or expressions may be correlated with belonging to a specific social or ethnic group. In this case, analysing social media posts, likes or friendship connections would lead to 'blatant proxies' and thus reintroduce banned scoring factors. Furthermore, there is no direct causality between social media data and driving behaviour that could justify a disparate impact.

#### **Interference with the right of self-determination**

Individuals may find it intrusive if their insurer uses social media data for underwriting. They would have to worry about the consequences of their social network activities for their insurance premium. At the same time, consumers are unlikely to possess the information necessary to appreciate these consequences (otherwise they could use this information to 'game the system'). In extreme cases, individuals may renounce social media activities altogether.

These concerns are somewhat eased if participation is voluntary and there is no continuous monitoring of social media activity (as in the case of Admiral Insurance).

75 Like companies in other industries, some insurers use social media data for marketing, customer segmentation and customer engagement. In addition, some insurers use social media data for forensic purposes.

76 Ruddick, G. (2016) "Admiral to price car insurance based on Facebook posts", The Guardian, November 2, 2016, and Ruddick, G. "Facebook forces Admiral to pull plan to price car insurance based on posts". The Guardian, November 2, 2016.

77 According to article 3.15. of Facebook's Platform Policy, the site's data must not be used to "make decisions about eligibility, including whether to approve or reject an application or how much interest to charge on a loan", see <https://developers.facebook.com/policy/>.

78 Ruddick, G. (2016) "Admiral to price car insurance based on Facebook posts". The Guardian, November 2, 2016, and Graham Ruddick: "Facebook forces Admiral to pull plan to price car insurance based on posts". The Guardian, November 2, 2016.

## Box 3: Wearables

### *What is it and how does it work?*

The market for wearable devices is growing rapidly. In 2015, 78 million wearable devices were sold.<sup>79</sup> By the end of 2016, Fitbit alone had sold 60 million devices.<sup>80</sup> A first generation of wearables focused on fitness and healthy lifestyle by tracking physical activity (e.g. number of steps) or sleep patterns, for example, but were limited in what they measured. Today, types of wearables are proliferating. Increasingly, medical-grade wearables are being developed that measure and track a broad range of vital signs such as heart rate, blood oxygenation, skin temperature, skin/blood perfusion and blood glucose levels.<sup>81</sup>

Several insurers have developed propositions in life and health insurance that use data from wearables.

### *How is the data used?*

Wearables data can be used in different parts of the insurance value chain. It can be used to identify customer needs, define new customer segments and develop new propositions to address specific needs. In underwriting, wearables provide data that allows a better assessment of health risks and the implementation of pricing models that incentivise healthy lifestyles. The wealth of data enables enhanced understanding of how different variables affect morbidity and mortality as well as improved measurement of the outcomes of medical treatments as a basis for quality improvements. Wearables data can further be used in claims to initiate and speed up the claims process, reduce the need for burdensome documentation by policyholders, and identify fraudulent claims. Importantly, wearables data allows the implementation of proactive risk management and early intervention mechanisms.

### *What are the benefits?*

#### **Consumers**

Consumers typically benefit from premium rebates and reward programmes for sharing their wearables data with the insurer. These benefits are often linked to meeting certain targets connected to a healthy lifestyle. Furthermore, consumers receive information and advice on how to moderate their behaviour to influence their overall health outlook.

Medical-grade devices offer a range of new possibilities for consumers with specific conditions. For example, monitoring the health conditions of patients with a chronic disease enables an enhancement of their quality of life by reducing the need for hospitalisation as well as affording the possibility of instant intervention and treatment in the case of abnormal conditions, thereby positively influencing the course of disease.

#### **Society**

In times of rapidly rising health costs, wearables can be an important ingredient in managing these costs by incentivising healthy lifestyles. At the same time, wearables technology offers the prospect of improving quality of healthcare by measuring outcomes of medical treatments and supporting the management of chronic conditions.

### *What concerns does it raise?*

#### **Fairness and discrimination**

The use of wearables data is likely to enhance the accuracy of risk assessments. However, it may raise concerns of discrimination as health-related lifestyles often correlate with income and education level.

#### **Interference with the right of self-determination**

Continuous surveillance may imply that individuals feel constrained in their lifestyle choices, particularly if unhealthy choices are associated with high premium costs. This may raise a concern of interfering with an individual's right of self-determination.

#### **Contextual integrity**

Inadvertent or abusive disclosure of sensitive health data could harm affected individuals and lead to stigmatisation. Secondary use of health-related personal data, e.g. for marketing purposes, could intimidate individuals who may feel limited in their lifestyle choices. In an extreme scenario, secondary use of data may limit opportunities of individuals in the labour market or in their career choices.

A violation of contextual integrity of personal health data could deter individuals from using wearable devices.

#### **Affordability and exclusion**

Individuals with high health-related risks may face high and potentially unaffordable premiums. This may limit their access to basic medical provisions, leading to a further deterioration of their condition.

79 Wearables.com.

80 "How many Fitbit devices have been sold in total?", available at <https://www.quora.com/How-many-Fitbit-devices-have-been-sold-in-total>

81 Swiss Re (2017), "The Integration of Wearables and Insurance" available at [http://institute.swissre.com/research/library/Medical\\_Wearables\\_Kelvyn\\_Young.html](http://institute.swissre.com/research/library/Medical_Wearables_Kelvyn_Young.html).

# Glossary

**Artificial Intelligence** is a branch of computer science dealing with the simulation of intelligent behaviour in computers. More commonly, the term is used to refer to the capability of a machine to imitate intelligent (human) behaviour.<sup>82</sup>

**Big data** is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision-making. Big data may be assessed through 5 'V' parameters: volume, velocity, variety, veracity, and variability.<sup>83</sup> Some commentators have added visualisation and value to those parameters.<sup>84</sup> Other definitions emphasise the complexity of big data. The National Institute of Standards and Things (NIST) defines big data as data that exceeds the capacity and capability of current methods and systems.

**Data** are facts and statistics collected for reference or analysis.<sup>85</sup>

**Data mining** is a procedure that uses algorithms to analyse large databases for patterns and correlations between data. These correlations indicate a relation between data without establishing causes or reasons.<sup>86</sup>

**Data protection**, technically speaking, is the process of safeguarding important information from corruption and/or loss.<sup>87</sup> European jurisprudence tends to treat data protection as an expression of the right to privacy. While there are overlaps in the concepts of data protection and privacy, there are also differences in their scope, the scope of data protection being broader than the scope of privacy.<sup>88</sup>

**Data science** is the extraction of knowledge from data. Data science includes big data and is conceived as a broader discipline that employs techniques and theories from mathematics, statistics, computing, and information technology, for example machine learning, to uncover patterns in data from which predictive models can be developed.<sup>89</sup>

**Decisional privacy** refers to the freedom to make one's own decisions without interference by others in regard to matters seen as intimate or personal.<sup>90</sup>

**Digital monitoring** refers to the continuous collection of significant amounts of behavioural data and the use of this data for dynamic and individual risk assessment and pricing.

**Information** is facts provided by or learned about something or someone. Both data and information may be used as a basis for reasoning or calculation. While there used to be more of a distinction between data as underlying facts and statistics, and information as knowledge gleaned from these facts and statistics, the definitions have now become quite close and may be used synonymously.<sup>91</sup>

**Informational privacy** is concerned with the interest of individuals in exercising control over access to information about themselves.<sup>92</sup>

**Internet of Things (IoT)** has been defined by the International Telecommunication Union (ITU) as "a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual)

82 <https://www.merriam-webster.com/dictionary/artificial%20intelligence>.

83 Swan, M. (2015) "Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services" in 2015 IEEE First International Conference on Big Data Computing Service and Applications.

84 See e.g. Devan, A. (2016) "The 7 V's of Big Data", available at <https://www.impactradius.com/blog/7-vs-big-data/>.

85 Swan, M. (2015) "Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services" in 2015 IEEE First International Conference on Big Data Computing Service and Applications.

86 Hildebrandt, M. (2008) Defining Profiling: "A New Type of Knowledge?" in Hildebrandt, M. and Gutwirth, S. (eds) "Profiling the European Citizen: Cross-Disciplinary Perspectives", Dordrecht, Netherlands: Springer pp. 17-45.

87 <http://searchstorage.techtarget.com/definition/data-protection>.

88 Kokott, J. and Sobotta, C. (2013) "The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR" International Data Privacy Law, 3(4), available at <https://academic.oup.com/idpl/article/3/4/222/727206/The-distinction-between-privacy-and-data>.

89 Swan, M. (2015) "Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services" in 2015 IEEE First International Conference on Big Data Computing Service and Applications.

90 "Privacy and Information Technology", Stanford Encyclopedia of Philosophy, first published Nov. 20, 2014.

91 Swan, M. (2015) "Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services" in 2015 IEEE First International Conference on Big Data Computing Service and Applications.

92 "Privacy and Information Technology", Stanford Encyclopedia of Philosophy, first published Nov. 20, 2014.

things based on existing and evolving interoperable information and communication technologies.”<sup>93</sup>

**Machine learning** is an application of artificial intelligence that provides systems with the ability to automatically learn from experience and improve without being explicitly programmed. Machine learning focuses on the development of computer programmes that can access data and use it to learn for themselves.<sup>94</sup> A distinction is made between supervised and unsupervised machine learning.

**Metadata** is “data that provides information about other data.” Descriptive metadata describes a resource for purposes such as discovery and identification. Structural metadata helps in understanding the format and definition of the information.<sup>95</sup> It is possible to derive personal information from metadata: authorities tracked down computer security pioneer and fugitive John McAfee through the metadata associated with a photo of him posted on a Central American online magazine blog. The digital photo included metadata—date, time, and geo-location—that pinpointed McAfee’s whereabouts.<sup>96</sup>

**Personal information or data** is information or data that are linked or can be linked to individual persons.<sup>97</sup> In the European General Data Protection Regulation (GDPR), personal data is defined as “any information relating to a person who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person.” So in many cases online identifiers including IP address, cookies and so forth will now be regarded as personal data if they can be (or are capable of being) linked back to the data subject without undue effort.

**Privacy:** there is no universally accepted definition of the concept of privacy. For the sake of this report, we define privacy as the ‘appropriate use of personal data’. See also ‘informational privacy’ and ‘decisional privacy’.

**Profiling** is the process of ‘discovering’ correlations between data in databases that can be used to identify and represent a human or non-human subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category.<sup>98</sup>

93 Internet of Things Global Standards Initiative, available at <http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx>.

94 <http://www.expertsystem.com/machine-learning-definition/>.

95 <https://www.merriam-webster.com/dictionary/metadata>.

96 Marotta, P. (2013) “The Whodunnit of Big Data”, Risk and Insurance, available at <http://riskandinsurance.com/the-whodunit-of-big-data/>.

97 “Privacy and Information Technology”, Stanford Encyclopedia of Philosophy, first published Nov. 20, 2014.

98 Hildebrandt, M. (2008) Defining Profiling: “A New Type of Knowledge?” in Hildebrandt, M. and Gutwirth, S. (eds) “Profiling the European Citizen: Cross-Disciplinary Perspectives”, Dordrecht, Netherlands: Springer.

# Appendix

## Types of data used in insurance

Type of data	Examples	Personal / non-personal	Use <sup>99</sup>	Data source
<b>Panel A: Traditional data</b>				
Demographic data	Age, gender, civil and family status, profession, address	Personal	Risk selection	Policyholders
Medical data	Medical history, medical condition, condition of family members, genetic testing	Personal	Risk selection	Policyholders
Exposure data	Type of car, value of building contents, type and features of dwellings	Personal/ non-personal	Risk selection	Policyholders
Behavioural data	Smoking, drinking behaviour, distance driven in a year, deductible choice, life insurance lapse rates	Personal/ non-personal	Risk selection, marketing	Policyholders, industry statistics
Loss data	Claim reports from car accidents, liability cases	Personal/ non-personal	Claims management	Policyholders, information exchange within industry
Population data	Mortality rates, morbidity rates, car accidents	Anonymised and aggregated personal data	Risk selection	Government, industry statistics, academia
Hazard data	Frequency and severity of natural hazards	Non-personal	Risk selection	Government, industry statistics, academia
Other traditional data	Credit reference, claim adjustment reports, information from the auto repair shops	Personal/ non-personal	Risk selection, marketing, claims management	Policyholders, credit agents, partner adjusters or agencies involved in the claim
<b>Panel B: New data in the era of digitisation</b>				
IoT data	Driving behaviour (telematics), physical activity and medical condition (wearables), surveillance (smart home)	Personal	Risk selection, claims management	Data collection devices
Online media data	Web searches, online buying behaviour, social media activities	Personal	Risk selection, marketing	Technology companies (internet providers, search engine providers, e-commerce providers, social media platforms)
Insurers' own digital data	Interaction with insurers (call centre data, users' digital account information, digital claim reports, online behaviour while logging in to insurers' websites or using insurers' app)	Personal	Marketing, claims management	Insurers' own customer service or call centre, insurers' websites and apps
Other digital data	Selfie (to estimate biological age for life insurance), flight information for flight delay insurance	Personal and non-personal	Risk selection, marketing, claims management	Policyholders, all other possible data related

99 Here risk selection includes pricing and underwriting; marketing includes distribution and sales activities; claims management includes fraud detection.

*Overview of privacy trade-offs*

Issue	Benefit	Cost
Discrimination	Accuracy of risk classification	Equal treatment
Intrusiveness	Risk reduction	Intrusiveness
Secondary use	Value of data	Contextual integrity
Individualisation	Individual pricing	Affordability
Solidarity (social insurance)	Individualisation	Equity
Risk pooling (private insurance)	Individualisation	Value of insurance

### Overview of potential future scenarios

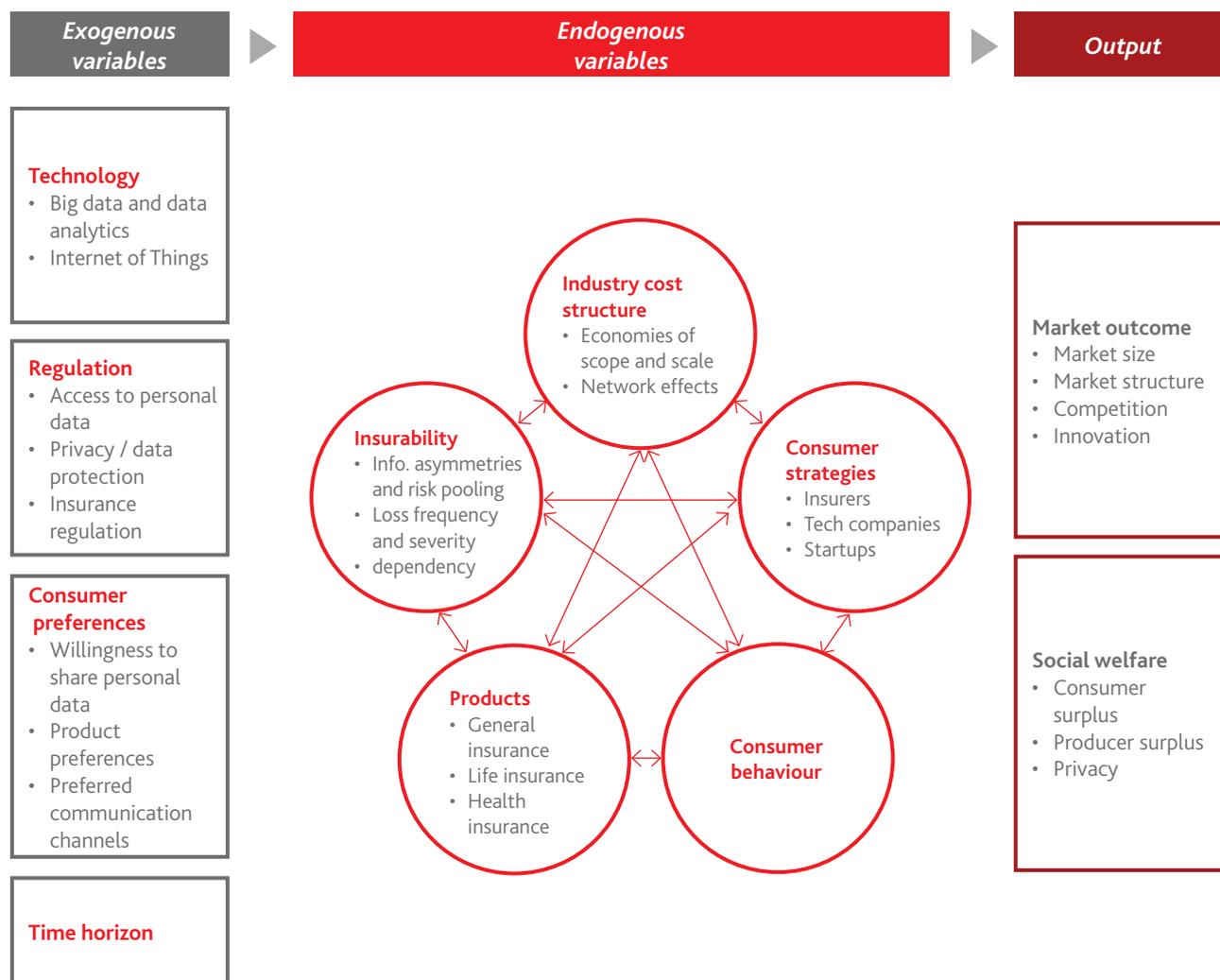
Scenario	Role of insurers	Impact on competition	Impact on informational asymmetries	
<b>Scenario 1:</b> <b>The digital society</b> 	Digital insurers	Enhanced competition	Significantly reduced if not eliminated	
<b>Scenario 2a:</b> <b>Insurance at two speeds</b> 	<ul style="list-style-type: none"> <li>Insurers exist as pure risk carriers or are driven out of the market, or exist as traditional insurers for customers not willing to share data</li> <li>Insurers may coexist with technology companies, depending on their speed of learning</li> </ul>	Reduced competition if traditional insurers are driven out of the market	Enhanced informational asymmetries (at least temporarily) if traditional insurers coexist with technology companies	
<b>Scenario 2b:</b> <b>Champions league</b> 	<ul style="list-style-type: none"> <li>Large insurers exist as digital insurers</li> <li>Small insurers are driven out of the market or exist as traditional insurers for customers not willing to share data</li> </ul>	Reduced competition if small insurers are driven out of the market	See scenario 2a	
<b>Scenario 3a:</b> <b>Avoid discrimination</b> 	Insurers engage in digital monitoring for prediction and prevention of risks	Competition mainly from collectors of Internet of Things data (car manufacturers, providers of smart home devices, providers of wearables)	Significant reduction or elimination of informational asymmetries in business lines that lend themselves to digital monitoring	
<b>Scenario 3b:</b> <b>Avoid intrusiveness</b> 	Insurers use broad range of online media data for risk assessment and selection	Competition mainly from collectors of online media data ('BigTech')	Reduction of informational asymmetries	
<b>Scenario 3c:</b> <b>Avoid risk of abuse</b> 	See scenario 4	See scenario 4	See scenario 4	
<b>Scenario 4:</b> <b>Digital backlash</b> 	Traditional insurers	Reduced competition by eliminating threat of market entry	Increased informational asymmetries if consumers have access to enhanced insights	
<b>Scenario 5:</b> <b>A tale of trust</b> 	Digital insurers and trusted data managers	Potential reduction of competition if only large insurers can establish themselves as trusted data managers	Significantly reduced if not eliminated	

	Impact on consumers	Impact on welfare	Public policy issues / concerns
	<ul style="list-style-type: none"> <li>• Lower prices on average</li> <li>• Enhanced and more tailored products</li> <li>• Increased premium difference between low and high risks</li> <li>• Shift of consumer surplus to firms if consumers do not have access to risk insights, and competition is weak</li> <li>• Consumers not willing to share data may face unfavourable conditions</li> </ul>	Large welfare gains through risk and cost reductions	<ul style="list-style-type: none"> <li>• Discrimination</li> <li>• Intrusiveness</li> <li>• Violation of contextual integrity</li> <li>• Affordability for high risks</li> </ul>
	<ul style="list-style-type: none"> <li>• Enhanced and more tailored products for individuals willing to share data</li> <li>• Increased premium difference between low and high risks</li> <li>• Shift of consumer surplus to firms if consumers do not have access to risk insights, and competition is weak</li> <li>• Consumers not willing to share data may face unfavourable conditions</li> </ul>	Ambiguous welfare effects	<ul style="list-style-type: none"> <li>• Affordability for high risks</li> </ul> <p>If traditional insurers are driven out of the market:</p> <ul style="list-style-type: none"> <li>• Discrimination</li> <li>• Intrusiveness</li> <li>• Violation of contextual integrity</li> </ul>
	See scenario 2a	See scenario 2a	See scenario 2a
	<ul style="list-style-type: none"> <li>• Enhanced and tailored products</li> <li>• Increased premium difference between low and high risks</li> <li>• Consumers not willing to share data may face unfavourable conditions</li> </ul>	Considerable risk reduction for risks that lend themselves for digital monitoring	<ul style="list-style-type: none"> <li>• Intrusiveness</li> <li>• Affordability for high risks</li> </ul>
	<ul style="list-style-type: none"> <li>• Enhanced and tailored products</li> <li>• Increased premium difference between low and high risks</li> <li>• Consumers not willing to share data may face unfavourable conditions</li> </ul>	Welfare gains from increased accuracy of risk assessments	<ul style="list-style-type: none"> <li>• Discrimination</li> <li>• Violation of contextual integrity</li> <li>• Affordability for high risks</li> </ul>
	See scenario 4	See scenario 4	See scenario 4
	<ul style="list-style-type: none"> <li>• Persistence of high risks</li> <li>• High cost of insurance</li> <li>• No innovation</li> </ul>	No significant welfare gains, potential welfare loss due to increased informational asymmetries	Maximum protection of privacy of individuals
	<ul style="list-style-type: none"> <li>• Lower prices on average</li> <li>• Enhanced and more tailored products</li> <li>• Increased premium difference between low and high risks</li> <li>• Shift of consumer surplus to firms if consumers do not have access to risk insights, and competition is weak</li> <li>• Consumers not willing to share data may face unfavourable conditions</li> </ul>	Large welfare gains through risk and cost reductions	<ul style="list-style-type: none"> <li>• Discrimination</li> <li>• Intrusiveness</li> <li>• Affordability for high risks</li> </ul>

### Overview of concerns and potential public policy approaches

Concern	Issues	Public policy / regulatory approach	Comments
Individualisation of insurance	Affordability / exclusion	Restriction in the use of risk indicators	Leads to economic inefficiencies and adverse selection
		Rate regulation	Leads to economic inefficiencies and adverse selection
		High-risk pools	May lead to competitive distortions depending on their design
		Separating funding for high risks from insurance premiums	No distortion of the price mechanism
	Undermining the solidarity principle	Clear regulatory distinction between private and social insurance	Private insurance does not rely on solidarity
	Premium volatility		No justification for regulation unless affordability is an issue
Fairness and discrimination	Explicit discrimination	Prohibition of the use of discriminatory risk indicators (e.g. race, gender, etc.)	Also applies to algorithmic decision-making
	Implicit discrimination	Enhance customer choice	
		Audit and test runs of algorithms	
		Restricting the use of blatant proxies if not based on causation	
		Requirement to install process for consumers to appeal against decisions	
	Treating similar individuals differently	Enhance customer choice	
Interference with right of self-determination		Analogous to affordability plus assign property rights to personal data	
Competition	Monopolisation	Competition policy	Intervention only after facts have been established; unlikely to be effective
		Data portability	Question whether data portability as in GDPR is effective
		Assign property rights to personal data	

## Framework for scenario development



This framework is not intended for mathematical modelling or quantitative estimation, but to illustrate the basic logic for scenario development.

### Overview of propositions, benefits and privacy concerns

Product line	Proposition / operational improvement	Data is used in (element of value chain)	Data source / use case	Benefits to consumers	Benefits to insurers
<b>Motor insurance</b>	Telematics	Underwriting	Data on driving behaviour, including geographic position, speed, acceleration, braking severity, vibration and impact events (collected through vehicle information system or on-board diagnostics device)	<ul style="list-style-type: none"> <li>• Upfront premium discount or cash-back discount</li> <li>• Information on driving behaviour</li> <li>• Value-added services such as automated emergency calls, remote assistance, stolen vehicle tracking etc.</li> </ul>	<ul style="list-style-type: none"> <li>• Reduction in claims payments and operating costs</li> <li>• Improved risk assessment</li> <li>• Additional income through value-added services</li> </ul>
	Motor insurance based on the use of social media data for risk profiling	Underwriting	Risk profiling based on text analysis of social media data (posts, likes, etc.)	<ul style="list-style-type: none"> <li>• Reduced premiums for individuals classified as low risks</li> <li>• Enhanced customer experience through digitisation and reduced information requests</li> </ul>	<ul style="list-style-type: none"> <li>• Reduction in operating costs through automation</li> <li>• Improved risk assessment</li> </ul>
	Pay-as-you-drive: use of geolocation data to determine driving distance	Underwriting	Use of geolocation data for risk assessment	<ul style="list-style-type: none"> <li>• Reduced premiums for individuals who drive less</li> <li>• Enhanced customer experience</li> </ul>	<ul style="list-style-type: none"> <li>• Reduction in operating costs through automation</li> <li>• Improved risk assessment</li> </ul>
<b>Property insurance</b>	Smart home: homeowner and renter insurance based on the use of smart home sensor data	Underwriting	Use of smart home sensor data to monitor risk-relevant features	<ul style="list-style-type: none"> <li>• Premium reduction</li> <li>• Enhanced customer experience</li> <li>• Value-added services such as optimisation of energy use, security systems etc.</li> </ul>	<ul style="list-style-type: none"> <li>• Reduction in claims payments and operating costs</li> <li>• Improved risk assessment</li> <li>• Additional income through value-added services</li> </ul>
	Digital insurance: fully digitised homeowner and renter insurance	Distribution/underwriting	Use of data about a particular home or neighbourhood from a variety of sources for risk classification	<ul style="list-style-type: none"> <li>• Reduced premiums for policyholders in low-hazard zones or low-risk neighbourhoods</li> <li>• Enhanced risk insights</li> <li>• Enhanced customer experience</li> </ul>	<ul style="list-style-type: none"> <li>• Reduction in claims payments and operating costs</li> <li>• Improved risk assessment</li> </ul>
	Drones: use of drones for risk and loss assessment in industry and agriculture	Claims	Use of drones to collect data for risk assessment (e.g. industrial sites, agriculture) and loss assessment (e.g. in case of natural disasters)	<ul style="list-style-type: none"> <li>• Enhanced risk insights and risk management services in industry and agriculture</li> <li>• Value-added services such as optimisation of use of pesticides and harvest in agriculture</li> <li>• Faster disaster assistance, disaster response and claims settlement</li> </ul>	<ul style="list-style-type: none"> <li>• Reduction in claims payments and operating costs</li> <li>• Improved risk assessment</li> </ul>
	On-demand insurance: digital insurance to cover assets on short-term basis	Distribution/underwriting	Use of various data sources (e.g. geolocation data, point-of-sale data and behavioural data etc.) to identify customer needs and for distribution and underwriting	<ul style="list-style-type: none"> <li>• Premium reduction</li> <li>• Enhanced customer experience</li> <li>• Enhanced coverage and customer choice</li> </ul>	<ul style="list-style-type: none"> <li>• Additional income through new coverages</li> </ul>

	Benefits to society	Privacy concerns	Consequence of inhibiting proposition	Consequence of no privacy restrictions
	<ul style="list-style-type: none"> <li>Risk reduction: reduced number of accidents and injuries / deaths through promotion of behavioural change and more efficient emergency response</li> <li>Reduced cost of insurance</li> </ul>	<ul style="list-style-type: none"> <li>Intrusiveness: use of data to educate individuals on their driving behaviour</li> <li>Risk of unwarranted secondary use of data (use of driving score outside of motor insurance) and accidental dissemination that could lead to abuse and stigmatisation</li> </ul>	<ul style="list-style-type: none"> <li>Elevated number of accidents, injuries and deaths</li> <li>Elevated cost of insurance</li> <li>Foregone efficiency gains through the integration of value-added services (e.g. integration of insurance with emergency response)</li> </ul>	<ul style="list-style-type: none"> <li>Unrestricted sharing and use of data may lead to stigmatisation based on driving score</li> <li>Surveillance and violation of contextual integrity may lead to individuals choosing not to drive</li> </ul>
	<ul style="list-style-type: none"> <li>Cost reduction through automation</li> <li>Potential efficiency gains through increased accuracy of risk assessment</li> </ul>	<ul style="list-style-type: none"> <li>Potential for discrimination ('blatant proxies')</li> <li>Violation of contextual integrity</li> </ul>	<ul style="list-style-type: none"> <li>Elevated cost of insurance</li> </ul>	<ul style="list-style-type: none"> <li>Violation of contextual integrity: individuals may choose not to be active on social media</li> <li>Informed individuals may 'game the system'</li> </ul>
	<ul style="list-style-type: none"> <li>Enhanced fairness of risk classification</li> <li>Potential reduction of traffic congestion if individuals renounce unnecessary driving</li> </ul>	<ul style="list-style-type: none"> <li>Premium increases for individuals who drive more</li> <li>Potential discrimination if belonging to a specific social group requires having to drive longer distances</li> <li>Risk of unwarranted secondary use and accidental dissemination of data</li> </ul>	<ul style="list-style-type: none"> <li>Inefficiency / unfairness of inaccurate risk classification</li> <li>Elevated cost of insurance</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of data on driving, risk of stigmatisation</li> <li>Unwarranted secondary use of data</li> <li>Individuals may choose not to drive</li> </ul>
	<ul style="list-style-type: none"> <li>Risk reduction through early intervention</li> <li>Cost reduction</li> <li>Enhanced fairness of risk classification</li> </ul>	<ul style="list-style-type: none"> <li>Intrusiveness: use of data to influence behaviour</li> <li>Risk of unwarranted secondary use of data on household contents, personal habits, etc. and accidental dissemination that could lead to abuse and stigmatisation</li> </ul>	<ul style="list-style-type: none"> <li>Inefficiency / unfairness of inaccurate risk classification</li> <li>Elevated cost of insurance</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of data on household contents and personal habits, risk of stigmatisation</li> <li>Unwarranted secondary use of data</li> </ul>
	<ul style="list-style-type: none"> <li>Risk reduction through targeted mitigation measures</li> <li>Enhanced fairness of risk classification</li> </ul>	<ul style="list-style-type: none"> <li>Potential for discrimination if living in a high-risk neighbourhood is associated with belonging to a specific social group</li> <li>Higher premiums and potential unaffordability for policyholders in high-hazard zones or high-risk neighbourhoods</li> <li>Risk of unwarranted secondary use of data on household contents, personal habits, etc. and accidental dissemination that could lead to abuse and stigmatisation</li> </ul>	<ul style="list-style-type: none"> <li>Inefficiency / unfairness of inaccurate risk classification</li> <li>Elevated cost of insurance</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of data and stigmatisation</li> <li>Unwarranted secondary use of data</li> <li>High risks may be priced out of the market</li> </ul>
	<ul style="list-style-type: none"> <li>Risk reduction through targeted risk management</li> <li>Loss reduction through early intervention and faster disaster response</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of accidentally collected personal data (analogous to Google Maps)</li> </ul>	<ul style="list-style-type: none"> <li>Elevated levels of risk</li> <li>Enhanced level of casualties in case of disasters</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination or abuse of commercial secrets</li> <li>Risk of dissemination of accidentally collected personal data and stigmatisation</li> <li>Unwarranted secondary use of data</li> </ul>
	<ul style="list-style-type: none"> <li>Cost reduction</li> <li>Enhanced coverage</li> <li>Enhanced fairness of risk classification</li> </ul>	<ul style="list-style-type: none"> <li>Risk of unwarranted secondary use of data</li> </ul>	<ul style="list-style-type: none"> <li>Elevated cost of insurance</li> <li>Inefficiency / unfairness of inaccurate risk classification</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of data on personal assets and habits, risk of stigmatisation</li> <li>Unwarranted secondary use of data</li> </ul>

**Overview of propositions, benefits and privacy concerns (continued)**

Product line	Proposition / operational improvement	Data is used in (element of value chain)	Data source / use case	Benefits to consumers	Benefits to insurers
<b>Life and health insurance</b>	Wearables	Underwriting	Use of data from health tracking devices for assessment of health status and risks	<ul style="list-style-type: none"> <li>Premium reductions for people with healthy lifestyles</li> <li>Potential new coverages, e.g. for pre-existing conditions</li> <li>Risk insights and value-added services</li> </ul>	<ul style="list-style-type: none"> <li>Reduction in claims payments and operating costs</li> <li>Improved risk assessment</li> <li>Additional income through value-added services</li> </ul>
	Life and health insurance based on the use of social media data for risk profiling	Underwriting	Risk profiling based on text analysis of social media data (posts, likes, etc.).	<ul style="list-style-type: none"> <li>Premium reductions for people with healthy lifestyles</li> <li>Enhanced customer experience through digitisation and reduced information requests</li> </ul>	<ul style="list-style-type: none"> <li>Reduction in operating costs through automation</li> <li>Improved risk assessment</li> </ul>
	Life or health insurance based on image recognition for risk profiling	Underwriting	Use of facial recognition and artificial intelligence for risk profiling (generating information such as gender, how quickly a person is ageing, body mass index and whether the person smokes) to predict life expectancy	<ul style="list-style-type: none"> <li>Enhanced customer experience</li> </ul>	<ul style="list-style-type: none"> <li>Reduction in operating costs through automation</li> <li>Improved risk assessment</li> </ul>
	Use of genetic information for life and health insurance	Underwriting	Use of genetic test results to identify health risks and for health management	<ul style="list-style-type: none"> <li>Reduced premiums for individuals with favourable test results</li> <li>Information on preventative measures to reduce risk of illness where possible</li> </ul>	<ul style="list-style-type: none"> <li>Improved risk assessment and selection</li> </ul>
<b>Across product lines</b>	Fraud detection	Claims	Use of social media data, pattern analysis, voice recognition etc. to identify potential fraud instances and verify information provided by policyholder	<ul style="list-style-type: none"> <li>Reduced premiums through reduction of fraud incidence</li> </ul>	<ul style="list-style-type: none"> <li>Reduced claims payments</li> </ul>
	Automated claims handling	Claims	Use of digital information (e.g. pictures) and artificial intelligence for claims adjustment	<ul style="list-style-type: none"> <li>Enhanced customer experience</li> <li>Faster claims payments</li> <li>Potentially reduced litigation</li> </ul>	<ul style="list-style-type: none"> <li>Reduced operating costs</li> </ul>
	Digital concierge, digital brokers, virtual assistants	Distribution	Use of various data sources and artificial intelligence to assess customer needs and optimise coverage	<ul style="list-style-type: none"> <li>Enhanced customer experience</li> <li>Enhanced customer choice through tailored products</li> </ul>	<ul style="list-style-type: none"> <li>Reduced distribution costs</li> </ul>

Benefits to society	Privacy concerns	Consequence of inhibiting proposition	Consequence of no privacy restrictions
<ul style="list-style-type: none"> <li>Promotion and incentivisation of healthy lifestyle</li> <li>Enhanced risk insights and enhanced quality and efficiency of health care (through integrated system)</li> </ul>	<ul style="list-style-type: none"> <li>Intrusiveness: use of data to influence behaviour of individuals by penalisation of unhealthy and risky activities</li> <li>Higher premiums for individuals with unhealthy habits and lifestyle choices</li> <li>Risk of unwarranted secondary use of data on personal habit and lifestyle choices and accidental dissemination that could lead to abuse and stigmatisation</li> </ul>	<ul style="list-style-type: none"> <li>Elevated cost of insurance</li> <li>Foregone quality and efficiency improvements in health care</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of data and stigmatisation</li> <li>Unwarranted secondary use of data (e.g. pictures may be linked to other data for additional, potentially harmful, uses)</li> </ul>
<ul style="list-style-type: none"> <li>Cost reduction through automation</li> <li>Potential efficiency gains through increased accuracy of risk assessment</li> </ul>	<ul style="list-style-type: none"> <li>Potential for discrimination ('blatant proxies')</li> <li>Violation of contextual integrity</li> <li>High risks may find insurance unaffordable</li> </ul>	<ul style="list-style-type: none"> <li>Elevated cost of insurance</li> <li>Inefficiency / unfairness of inaccurate risk classification</li> </ul>	<ul style="list-style-type: none"> <li>Violation of contextual integrity: individuals may renounce activity on social media</li> <li>Informed individuals may 'game the system'</li> </ul>
<ul style="list-style-type: none"> <li>Reduced informational asymmetries</li> </ul>	<ul style="list-style-type: none"> <li>Potential discrimination based on gender or ethnicity</li> </ul>	<ul style="list-style-type: none"> <li>Elevated cost of insurance</li> <li>Inefficiency / unfairness of inaccurate risk classification</li> </ul>	<ul style="list-style-type: none"> <li>Risk of dissemination of data and stigmatisation</li> <li>Unwarranted secondary use of data</li> </ul>
<ul style="list-style-type: none"> <li>Pro-active disease management and prevention</li> <li>Enhanced accuracy of risk classification</li> </ul>	<ul style="list-style-type: none"> <li>High or unaffordable premium or denied coverage for individuals with unfavourable test results</li> <li>Discrimination based on genetic predisposition</li> <li>Risk of unwarranted secondary use of genetic data and accidental dissemination that could lead to abuse and stigmatisation</li> </ul>	<ul style="list-style-type: none"> <li>Inaccurate risk classification</li> <li>Potential for proactive disease management and prevention not realised</li> <li>Potential for new coverages for pre-existing conditions not realised</li> </ul>	<ul style="list-style-type: none"> <li>High risks may be priced out of market</li> <li>Risk of dissemination of data and stigmatisation</li> <li>Unwarranted secondary use of data</li> </ul>
<ul style="list-style-type: none"> <li>Enhanced efficiency</li> </ul>	<ul style="list-style-type: none"> <li>Violation of contextual integrity</li> </ul>	<ul style="list-style-type: none"> <li>Efficiency gains cannot be realised</li> </ul>	<ul style="list-style-type: none"> <li>Potential for discrimination if data is freely shared among insurers and individuals are falsely identified as fraudsters</li> </ul>
<ul style="list-style-type: none"> <li>Cost reduction through automation</li> </ul>	<ul style="list-style-type: none"> <li>Risk of unwarranted secondary use of data</li> </ul>	<ul style="list-style-type: none"> <li>Foregone efficiency gains</li> </ul>	<ul style="list-style-type: none"> <li>Digital claims history may be shared among insurers</li> </ul>
<ul style="list-style-type: none"> <li>Enhanced coverage</li> </ul>	<ul style="list-style-type: none"> <li>Potential monopolisation of customer relations</li> </ul>	<ul style="list-style-type: none"> <li>Foregone efficiency gains</li> </ul>	<ul style="list-style-type: none"> <li>Customer relations may be monopolised by a few large players</li> </ul>



The use of big data in insurance raises concerns around privacy, innovation and competition, which require intricate value judgements. The paper discusses the societal and economic benefits from the use of big data analytics in insurance and the key concerns that have been raised in public and regulatory debate. It also identifies the key trade-offs deriving from the enhanced use of personal data in insurance.



**The Geneva Association—International Association for the Study of Insurance Economics**  
Talstrasse 70, CH-8001 Zurich | Tel: +41 44 200 49 00 | Fax: +41 44 200 49 99

[secretariat@genevaassociation.org](mailto:secretariat@genevaassociation.org)