The personalisation of insurance

Mcfall, Liz; Meyers, Gert; Van Hoyweghen, Ine

Published in:
Big Data & Society

DOI:
10.1177/2053951720973707

Publication date:
2020

Editorial

BIG DATA & SOCIETY

Big Data & Society
July–December 2020: 1–11
© The Author(s) 2020
DOI: 10.1177/2053951720973707
journals.sagepub.com/home/bds
SAGE

Editorial: The personalisation of insurance: Data, behaviour and innovation

Liz McFall[1], Gert Meyers[2] and Ine Van Hoyweghen[2]

Abstract

The adoption of Big Data analytics (BDA) in insurance has proved controversial but there has been little analysis specifying how insurance practices are changing. Is insurance passively subject to the forces of disruptive innovation, moving away from the pooling of risk towards its personalisation or individualisation, and what might that mean in practice? This special theme situates disruptive innovations, particularly the experimental practices of behaviour-based personalisation, in the context of the practice and regulation of contemporary insurance. Our contributors argue that behaviour-based personalisation in insurance has different and broader implications than have yet been appreciated. BDAs are changing how insurance governs risk; how it knows, classifies, manages, prices and sells it, in ways that are more opaque and more extensive than the black boxes of in-car telematics.

Keywords

Behavioural data, Big Data Analytics, insurance, personalisation, surveillance, telematics

This article is a part of special theme on Insurance Personalization. To see a full list of all articles in this special theme, please click here: https://journals.sagepub.com/page/bds/collections/personalizationofinsurance

> Telematics are not intended merely to know but also to do (economies of actions). They are hammers; they are muscular; they enforce. Behavioural underwriting promises to reduce risk through machine processes designed to modify behaviour in the direction of maximum profitability. Behavioural surplus is used to trigger punishment, such as real-time rate hikes, financial penalties, curfews and engine lockdowns, or rewards, such as rate discounts, coupons and gold stars to redeem for future benefit. (Zuboff, 2019: 215)

Has individual behaviour come suddenly to matter to private, commercial insurers? There is an inclination among scholars – Shoshanna Zuboff offers a recent and high-profile example – to use insurance as an exemplar of the individual and social harms processes of datafication can do through unfettered access to, 'behavioural' data. An insurance imaginary, tooled with Big Data analytics (BDA) and the internet of things (IoT), presents possibilities for discrimination, exclusion, behavioural surveillance and, by extension, modification, that seem unprecedented. This being the case it is important that researchers concerned about these harms pay closer attention to insurance 'in the wild' than they have hitherto. Critical data, media and internet researchers have subjected the tech industries to enough scrutiny over the last few years to detail how the automation of inequality and bias works (Bucher, 2018; Eubanks, 2018; Mittelstadt et al., 2016; Noble, 2018; Prince and Schwarcz, 2020). The same cannot be said of the insurance industry. Despite its historical significance in processing data and its current prominence in generating expectations about the future of

[1]The University of Edinburgh, Edinburgh, United Kingdom of Great Britain and Northern Ireland
[2]KU Leuven, Leuven, Belgium

Corresponding author:
Liz McFall, The University of Edinburgh, 21 George Square, Edinburgh EH8 9LD, United Kingdom of Great Britain and Northern Ireland.
Email: liz.mcfall@ed.ac.uk

data technologies (Meyers and Van Hoyweghen, 2018a), when researchers refer to insurance it tends to be in a supporting role, part of a broader analysis of the harms associated with other datafied phenomena.

This is an instance of a wider problem that, in the process of editing this special theme issue, we have come to think of as the 'interestingly uninteresting' character of insurance. Quite simply, insurance is hugely under-researched for an industry of its size and governmental importance (cf. Tanninen, 2020). Over the last two decades, insurance undertakings contributed an average of around 8.5% to GDP across OECD countries, reaching 19.8% in the UK immediately prior to the 2007 global financial crisis (OECD, 2020b).[1] Even as the 'financialisation' literature began to proliferate in the early 2000s, insurance attracted little attention despite being among the largest and earliest investors in capital markets.[2] As the first mass market test of the monetisation of probability theory, insurance is an epistemological forebear of contemporary financial instruments specifically, and statistical reasoning more generally (McFall, 2011; Van der Heide, 2019a).[3] Its capacity to quantify and organise risk and responsibility, loss and protection, led to insurance being deeply woven into the overlapping fabrics of public, social solidarity and private, intimate relationships (Baker and Simon, 2002; Baldwin, 1990; Dobbin, 1992; Ewald, 1991; Hacking, 1990; Lehtonen and Liukko, 2012; O'Malley, 1996; Porter, 1996; Lehtonen and Van Hoyweghen, 2014). These characteristics all shape the regulatory and infrastructural frameworks underpinning the various products and practices of global insurance undertakings. A generic category like health insurance, for example, can signify a highly selective mix of consumer choice, employer benefits and minimal social protection in one location, and universal healthcare in another. Insurance is, in sum, an arcane, esoteric, yet quotidian and ubiquitous socio-political-economic technology and this – almost paradoxically – has made it an unattractive object of study in the humanities and social sciences. As datafication and the associated phenomenon of 'insurtech'[4] seem set to 'disrupt' insurance, the gap in critical knowledge about how it works is becoming more apparent and more pressing.

In this special theme issue, we take a step towards correcting this by exploring the particular controversies surrounding the personalisation of insurance through novel combinations of individual-level data, behavioural tracking and innovation. Data, behaviour and innovation are, in one sense, core components of insurance as a technology that compensates risk through the mechanism of mutuality organised according to the 'laws of large numbers' (Hacking, 1990). There is nothing unprecedented about insurers processing large volumes of data; or assessing behaviour, for example, through questions about smoking, alcohol consumption, weight, criminal convictions, or sponsoring technological innovations that might more effectively measure and promote health or safety (French and Kneale, 2009; Van Hoyweghen, 2007; Yates, 2005). This is all insurance business as usual. Yet, in another sense, BDA bring a capacity to drill down to the level of individual, dynamic behaviour that is existentially threatening to insurers who are, epistemologically and organisationally, committed to classifying risk at an aggregate level. BDA introduces the hypothetical prospect of a risk 'pool of one' to an enterprise founded on modelling collectives. Added to this are all the encumbrances of an old and gigantic industry, the legacy infrastructures and entrenched, institutionalised practices that, in the discourses of big tech, make it 'ripe for disruption'.

Data, behaviour and innovation in insurance practice (on selling Lemonade)

To make the stakes clearer, consider how one of the better known insurtech brands positions the differences between its business model and those of traditional insurers. The US-based company Lemonade Inc. offers home, renters and more recently, pet insurance and is one of a handful of insurtech disrupters to have achieved unicorn status. The company was launched on the stock market with an Initial Public Offering (IPO) in July 2020.[5] In the publicity surrounding the IPO, the founders made much of the well-rehearsed claim that the tech industries' facility with data has now far surpassed that of insurers, allowing insurtech companies to take a completely different approach. In a promotional video released on YouTube just before the launch, co-founder Daniel Schreiber characterised this as a microscopic approach to risk:

> At twenty data points much of humanity looks alike so you price and underwrite large numbers of people as if they are a uniform, monolithic group but we found that the data we were gathering, that hundred X digital zoom we have, it showed us that groups our competitors seemed to consider to be monolithic were actually made up of predictable subgroups with over 600 percent variation in their likelihood to file a claim.[6]

At face value, Lemonade emerges in the video as a pioneer in the application of AI and BDA to classifying insurance risk using the 'volume, variety and velocity' of Big Data (Barocas and Selbst, 2016; Chen et al., 2012; OECD, 2020a; Prince and Schwarcz, 2020). Like others in the venture capital funded tech start-up world, Lemonade has a vested interest in this narrative. Getting the narrative right, as Geiger (2019) demonstrates in her discussion of the short and spectacular history of digital health start-up *Theranos*, is key to securing initial funding and ongoing capital investment. Lemonade is adroit at this, their unicorn valuation and the doubling of their share value on the first day of trading was secured even though the company, after 5 years in business, remains loss making and a very small player in its core markets (Ralph, 2020). There is nothing unusual about this. In tech industry funding, the historical experience of disrupter giants like Amazon, Uber and Airbnb has normalised expectations around very long and very slow, accretive 'paths to profitability'. These financial arrangements are important to understanding the technical and commercial challenges of disrupting traditional insurance in the direction of more 'microscopic', personalised risk assessment. They also hint at how intricately linked the technical and sentimental valuations of risk are in insurance (Jeanningros and McFall, 2020; McFall, 2014).

Since the 19th century the profitability of insurance has derived substantially from investment rather than premium income (Alborn, 2002; Ericson et al., 2000; McFall, 2014; Van der Heide, 2019a). The pattern shifts from sector to sector and over time but until relatively recently insurers were effectively cushioned from variations in premium income by fixed asset investments (Lex, 2020). The decline in investment returns since the 1990s has made insurance a more competitive market and increased the importance of premium income and pricing. This has incentivised firms to search for

> more sophisticated BDA-driven pricing models in order to optimize the profits with the help of the new possibilities offered by technological developments and new data sources [and] enabled a more granular segmentation of risks, increasing the effectiveness of risk selection, and allowing more risk-based pricing. (EIOPA, 2019: 29)

Lemonade's pitch to investors follows this logic to argue that it can use its 100× digital zoom to see, classify and price risk more accurately. The decline in fixed income returns means that Lemonade's commercial viability in the longer term appears tied to the superior efficacy of their underwriting. The broad claim that more granular, risk-based pricing will be more efficient and more profitable has increasing currency across the industry. It is made by both incumbents and new entrants, is the stock-in-trade of numerous consultancy reports, and is a central focus of global insurance regulators (EIOPA, 2019; EIOPA and NAIC, 2018; FCA, 2019; MAS, 2020; OECD, 2020a). The trouble is that what this means exactly, and how it relates to the commercial practice of insurance, is opaque. As a recent report by the OECD puts it:

> Insurance is based on the idea of pooling risks, and underwriting is most often based on past loss experiences and/or risk modelling. The prospect of having more data leads to the possibility of greater data analytics and in particular improving predictive analytics, enabling pricing that is better suited to expected risk, and is more granular or adjusted to policyholder behaviour. However, this is too simplistic a way of understanding how data contributes to insurance, as there are a variety of questions and implications that arise when considering potential scenarios that could occur with the arrival of big data. One of the first and principal questions is what big data means in the insurance context. There is also the question of whether pricing would, in fact, be more accurate in terms of risk and whether this would lead to an overall better production of insurance products. (OECD, 2020a: 9)

Controversies around the personalisation of insurance centre on assumptions about how data, behavioural tracking and innovation are being combined to calculate risk at a dynamic, individual level. Zuboff's formulation that insurers are using connected, telematic and self-tracking devices to 'modify behavior in the direction of maximum profitability', is in line with earlier popular and academic critiques of the direction insurance is taking (Lupton, 2016; O'Neil, 2016; Schüll, 2016; *The Economist*, 2015). In general, this formulation appears sound, but it is the details that matter.

First, there is little clarity about what sorts of data, and what sorts of behaviour are in play in discussions of how BDA permits more granular, personalised underwriting. There is a particular tendency to focus, almost exclusively, on telematic self-tracking data, perhaps because it is both very new and very visible. But it is only one of a number of Big Data sources that insurance undertakings use. According to a report by the UK's Financial Conduct Authority (FCA, 2016), these sources also include proprietary data (e.g. data from connected companies such as personal data of products purchased or loyalty cards); third party data (e.g. aggregated search engine data including credit checks, license details, claims discount databases, price comparison website quote); social media data and connected device data. This means that insurers have a far, far greater range of 'behavioural' data sources than connected devices to draw upon and even connected device data is more nuanced in their application than the idea of behavioural pricing suggests (see Jeanningros and McFall, 2020; Meyers, 2018; Meyers and Hoyweghen, 2020). Credit information, for example, has been used to price motor and home insurance for several years despite its potential for proxy discrimination (Kiviat, 2019; Prince and Schwarcz, 2020). The scope for discrimination has increased dramatically over the last decade with the incorporation of 'alternate' or 'external' customer data, to profile price sensitivity and risk attitude and has attracted scant public attention. This includes a vast and novel range of unstructured data, as Prince and Schwarcz explain:

> Traditionally, firms differentiated among customers, employees, and others based on a limited amount of data that they directly collected. In recent years, however, firms have increasingly come to rely on data secured from a broad number of external sources. These data frequently involve online actions, such as "transactions, email, video, images, clickstream, logs, search queries health records, and social networking interactions . . .."[54] But firms rely on data that increasingly also extends to actions in the physical world, which are measured by "sensors deployed in infrastructure such as communications networks, electric grids, global positioning satellites, roads and bridges, as well as in homes, clothing, and mobile phones." (2020: 1273)

Second, the proliferation of data sources and techniques still leaves open the question of whether more granular 100× zoom means more accurate risk pricing and in turn greater profitability. This is a hard question to settle for a number of interconnected reasons. While BDA can fine-grain risk classification and improve the predictability of policyholders making a claim, this is not, in and of itself, a path to profitability. Insurance premiums are both costed and priced. Costing is a technical process that involves the statistical analysis of an insurers' historical 'experience' of loss to forecast the costs of particular insured groups. Pricing, on the other hand, is a commercial process whereby a policy is offered at a certain premium based on a wide range of calculations about different customers, their cost to acquire, their propensity to buy and their propensity to switch provider. BDA are widely used in this context to 'optimise' dynamic prices; for example, new customers who search extensively and read policy information may be offered lower prices than loyal customers (FCA, 2019; Minty, 2016). Technical refinements in costing do not necessarily mean prices will more precisely reflect risk since pricing decisions also always involve commercial and marketing considerations. Even if risk-based pricing does prove to be more accurate, any gains will not be free. As Swedloff has argued:

> it might be extraordinarily expensive to harness big data and generate more refined risk classifications. Each carrier might have to spend significant sums to make marginal improvements to their risk classification scheme. These costs could be exacerbated because carriers may feel a pressure to follow popular trends. Given the press coverage on the wonders of big data, firm leaders may spend exorbitantly even if the new classification scheme costs more than it generates. (Swedloff, 2014: 359)

In a highly regulated competitive industry, there is also caution around the commercial consequences that may come from pioneering forms of risk-based pricing that could be seen as unfair, exclusionary or discriminatory. Insurance regulators have issued specific prohibitions on unfair discrimination. In the United States, for example, this includes refusing to insure or limiting coverage to individuals on the basis of ethnicity, race, gender, religion, national origin and other protected characteristics.[7] In Europe, anti-discrimination and data protection regulation has been issued at the EU and national levels, to prohibit the use of, for example, genetic information (Blasimme et al., 2019), race and ethnicity (Gellert et al., 2013), and gender (Rebert and Van Hoyweghen, 2015).[8] BDA and AI complicate the regulatory environment since they 'discriminate by proxy' in ways that are harder to detect and may derive from badly prioritised algorithms and the representational biases of the datasets they were trained on (Barocas and Selbst, 2016; Prince and Schwarcz, 2020; Swedloff, 2014). Recent work argues that anti-discrimination and data protection laws, such as the recent General Data Protection Regulation (GDPR) in Europe, do not provide adequate protection against discriminatory pricing algorithms in insurance (Drechsler and Benito Sanchez, 2018; Marelli et al., 2020). To make matters still more complicated, insurance falls under general financial regulation that does not address discrimination but does impose conditions on the kind of data on which actuarial calculation can be based. The EU Solvency II Directive,[9] which was enacted in the wake of 2007–2009 financial crisis and came into force in 2016, requires insurers to base their calculations only on data with actuarially demonstrable reliability. This requirement can result in the kind of Catch-22 described by Meyers and Van Hoyweghen (2020) in which the general regulatory environment, while not designed to restrict BDA experiments, nevertheless has that effect.

BDA, IoT connected devices and AI alter the range of what it is technically possible for insurers to do. The emergence of telematics and self-tracking product features in the motor, health and life sectors described in this issue, showcase the role of behavioural data in personalising insurance. They also demonstrate that while technological innovation may stretch, expand

and transform the technical classification of risk, this has to be accommodated within the commercial and regulatory practices of insurance. For insurance to be commercial, risk has to be transformed in an ongoing orchestration of legally and technically calculable processes and sentimentally appealing products (McFall, 2014).

Lemonade Inc. provides a nice worked example of what that involves. Lemonade's choice of brand name is itself revealing. Historically, insurance companies have chosen names to help solve marketing problems. Initially, that meant reassuring customers that however speculative the insurance business seemed, it was financially solid, safe and respectable with names like the Prudential, Equitable, Provident, Perpetual and Amicable. After decades of consolidation, the giant surviving companies chose initials and abstract portmanteaus such as AXA, AVIVA, Manulife or AIG, in an attempt to retain something of their histories in more pronounceable and memorable names. Insurtechs have adopted distinct naming conventions calculated to break into the low brand differentiation of the major incumbent insurers with quirkier, friendlier and more human names, for example Oscar, Brolly, Ladder, Clover and Rhino. 'Lemonade' fits the trend but adds an additional knowing wink to George Akerlof's classic formulation of adverse selection, the dynamic whereby high-risk individuals ('lemons') are motivated to buy more insurance than low-risk individuals ('peaches') when confronted with the same premium (Akerlof, 1970). Adverse selection, or the lemon problem, means that insurance companies cannot easily differentiate between low-risk and high-risk buyers:

> Some insurance buyers are low-risk "peaches" and other insurance buyers are high-risk "lemons." In many cases the insurance buyers have at least some sense of whether they are lemons or peaches. If the insurance company can tell the difference between lemons and peaches, it will charge the peaches a peach price and the lemons a lemon price consistent with actuarial fairness, and the market will work efficiently ... If insurance companies are not able to tell the difference between lemons and peaches, however, or if they are prevented from charging different prices, then they will have to charge all of the buyers the same price. (Baker, 2010: 1609–1610)

Lemonade's name then, is simultaneously a coded reference to their risk classification technique and a recognition of the significance of consumer sentiments. This is also evident in their pre-IPO video. The video opens with word association customer vox pops – insurance is 'boring', 'complicated', 'a rip-off', 'like a trigger word', 'evil'. Daniel Schreiber cuts in to explain that Lemonade's founders wanted

> to create an insurance company with an entirely different word cloud associated with it – a lot of people talk about Lemonade as being delightful, aligned, socially impactful, words like trust and love, those are the kinds of words that consumers have come to associate with Lemonade.[10]

Cultivating warm words is part of the company's strategy of linking machine learning, customer delight, predictive data and fast growth in a recursive, reinforcing loop 'that grows stronger with every rotation'. Whether Lemonade can actually deliver on their IPO promises is not the point, rather their example illustrates how, even in the most datafied insurtech, insurance technique is also a matter of sentiment.

Too many data controversies in all the wrong places

To paraphrase Donna Haraway (2016), it matters what matters we make trouble with. One of the problems with the attention paid to self-tracking is that it makes insurance look more straightforward and more precise than it is. It encourages us to continue to neglect the more confusing and difficult prospects raised by datafication in an industry that was poorly understood anyway. Insurance's unobtrusiveness means that now, when the data controversies surrounding surveillance and extraction have started to touch it, we don't have a specific vocabulary or grammar to assess what's happening. Instead we have formulations made from the perspective of data controversies *in general*. Cathy O'Neil's summoning of insurance as an instance of the socially destructive force of data science fits this mould precisely.

> Now, with the evolution of data science and network computers, insurance is facing fundamental change. With ever more information available – including the data from our genomes, the patterns of our sleep, exercise, and diet, and the proficiency of our driving – insurers will increasingly calculate risk for the individual and free themselves from the generalities of the larger pool. (O'Neil, 2016: 164)

Deborah Lupton's similar claim that insurance companies are incorporating 'self-tracking data into the calculation of risks and resultant premiums offered to customers' (2016: 108) is part of a trend in which insurance innovations feature only as a part of broader analyses of, for example, self-tracking and quantification, algorithmic and computational logics, the platform or surveillance economy. These discussions seldom specify what particular types of insurance behavioural schemes are part of. There is a world of difference between introducing a health-tracking component in voluntary life insurance and making it a mandatory part of group, employer or publicly funded health cover. The reluctance to delve into insurance classifications is understandable given how vast and variable the types are. Even at sector level insurance language is confusing. What is known as general insurance in the UK, Australia and many parts of Africa and Asia, is called property and casualty insurance in the United States and non-life insurance in the rest of Europe (cf. McFall and Moor, 2018). This fluid vocabulary creates plenty of scope for categorical confusions between critical debate, the idiomatic, internal, sectorally and regionally specific jargon of insurance practice and the regulatory contexts in which it operates.

Confusion is exacerbated by the way actors involved in the insurance technology and allied industries talk about themselves and the way critics respond. In the overlapping languages of consultancy, research, trend forecasting, investment and funding, scenarios of intensely surveillant, personalised insurance use cases are presented as the ubiquitous, nearly present, already unavoidable, now. There is an economy to this of course – insurtech actors have a vested interest in disruption narratives. Curiously, insurtech's drummers and its critics often talk in rhythm and draw on the same sources to document how the tech works. Cory Doctorow (2020) recently remarked that the critique of surveillance capitalism presumes that 'everything Big Tech says about itself is probably a lie' but makes an exception for the claims that tech makes about itself in its sales literature. There is something about surveillance tech in particular that seems to command both a reasonable dread and a disavowed fascination with being seen (cf. Ruckenstein and Granroth, 2020). Public concern about BDA is not well served by inflating the extent and value of its uses on the basis of tech industry claims while neglecting the hard empirical graft of figuring out what is happening in what Ball and Webster (2020: 2) have called 'the mid-range of BDA – the mesh of organisations which mediate between the end consumer, the organisational and societal context, and the marketer of products'.

In a different but not entirely unrelated context, Christopher Kelty (2017) describes how new media and internet technologies almost immediately raised questions about 'participation', about whether it could be caused, cured, enhanced or damaged by technological development. These troublesome questions, he argues, are endlessly, fruitlessly deliberated since:

> we do not really know what participation is that it could be caused or cured by technological development. On one day, participation is the solution to our most practical concerns or even an ethical calling; on the next day it is a containment strategy designed to keep us chillingly in place or to extract data and money from us at every turn. (Kelty, 2017: s77)

Kelty attributes the problem to the absence of any 'grammar' of participation. More than a definition, what is wanted is a sense of the place, the syntax or context of participation. There is tension between the *purpose* of participation and how it has come to be *institutionalised*. 'Participation' serves as both a normative ideal and an indication of the state of social activeness. It becomes both end and means, and even a means to the wrong sort of end, resulting in 'too much surveillance, too much unpaid labour, too much devolution of responsibility, too much democracy in all the wrong places' (2017: 88).

We suspect the word insurance suffers from a similar kind of trouble. Insurance is a malleable political rationality, a calculative epistemology, a means of social solidarity, a fantasy of replaceable totality. It is so very mundane and yet so drunkenly various that we find it hard to parse its grammar. At this moment when the proliferation of data and techniques is finally turning public attention towards insurance, we have not got an adequate shared vocabulary or grammar for a publicly framed discussion of what insurance is and how it works. Can systems trained on external data sources classify, cost and price risk more precisely? Does this vary by sector, in motor, property, health, life? If insurance risk is to be granularised is it still insurance or some other kind of financial instrument, a derivative, a future that is being traded? If so, in a pandemic climate and in the face of the millions globally who lack adequate health coverage should we not be talking urgently about alternative mechanisms for healthcare payment, for social, even interspecies, solidarity (Hendrickx and Van Hoyweghen, 2018; Prainsack and Van Hoyweghen, 2020)?

The contributions

There is no single discipline in which scholars of insurance are primarily concentrated. Instead research from across the social sciences, the humanities, law, business and mathematics combine in transdisciplinary insurance studies dedicated to a disparate range of practices. The articles in this special theme reflect that in the varied theoretical perspectives, methods and tools they bring to interrogating the role of data, behaviour and innovation in contemporary insurance. Together, they work towards establishing a 'grammar' for the growing public debate on the future of insurance by placing critical questions in the practical context of insurance innovation, insurtech and regulation.

We begin this task with Maiju Tanninen's account of how contested the technology of behaviour-based insurance is. Tanninen meticulously reviews the way recent developments in behaviour-based insurance have featured in two main streams of literature, Critical Data Studies (CDS) and the social and technical studies (STS) of insurance. She shows how CDS presents personalisation in insurance as an extreme, and yet still exemplary, case of datafication practices anticipating immense and unwanted consequences. STS-inspired studies, on the other hand, focus their attention on how concrete practices of datafication take place and enact practices of personalisation. Tanninen acknowledges the importance of critical voices on insurance, but she questions the shallowness of the portrayal of insurance arguing that CDS relies 'on the same understandings as the dominant Big Data enthusiastic discourses, ultimately taking insurers' and tech companies' visions on the "digital disruption" seriously' (Tanninen, 2020: 10). She makes the case that more nuanced, substantive and empirically grounded research on both how insurers approach personalisation, and how users experience the products, is required to mitigate the hype.

The next two articles explore how insurers assess risk and translate it into the products and prices offered to consumers in the two sectors most closely associated with personalisation: health insurance and motor insurance. In 'The Value of Sharing', Hugo Jeanningros and Liz McFall describe how the South African company, Discovery International, built their Vitality brand explicitly around the idea of supporting healthy behaviour. The Vitality case, they argue, demonstrates the intricate connections between the technical practices and sentimental appeals at work in how insurance companies value risk and how their customers value insurance policies. Behaviour is Vitality's core brand value and its policies provide incentives to customers to meet behavioural targets, share their data with the company and share their progress on social media. 'The value of sharing' for Discovery is in choreographing the perpetual dance between the technical valuation of risk, the commercial performance of the business and sentimental appeal of the brand.

Parallel twists and turns can be seen in the case study of experimental behaviour-based personalisation practices in car insurance presented by Gert Meyers and Ine Van Hoyweghen. They show how the regulatory requirements imposed on insurance companies create a Catch-22 situation. Currently, in order to be allowed to use BDA, insurers have to provide actuarial evidence on the relation between, for example, driving style and the experience of loss. Yet without collecting these data in real-time situations, it is impossible to build this evidence base. This is why insurers currently rely on the use of 'economic experiments' (Tironi and Laurent, 2014). Based on interviews and documentary analyses, Meyers and Van Hoyweghen outline how the in-vivo experiment was set-up, which interventions and manipulations were imposed to make the experiment successful, and how the study was evaluated by the actors. Following Austin (1962), they argue that economic experimentation 'in the wild' should not (only) be evaluated by its capacity to produce truth statements, but also on its (un)intended positive (happy) or negative (unhappy) consequences.

The final two articles of this theme issue reflect, in rather different ways, on how BDA are transforming what Francois Ewald (1991, 2012) called 'insurantial imaginaries'. Laurence Barry and Arthur Charpentier (2020) consider how pricing practices in motor insurance are changing with the use of personalised analytics. They start by raising the question of whether we are indeed in the midst of a 'new epistemological turn in data analysis'. Such a turn, they argue, would shift insurance knowledge away from its basis in statistics to create homogeneity through the classification of people into risk groups, towards a basis in analytics to generate differences (cf. Mau, 2019). This endangers the principle of the mutualisation of risk which is the historical and epistemological core of insurance technique. Barry and Charpentier explain that as most actuarial scientists recommend the addition of new variables to existing (statistical) risk models, the epistemological break has not – at least not yet – transformed risk practices across the industry. For them, although risk practices are becoming more granular over time the underlying mechanisms of insurance will continue to be based on the constitution of statistical risk groups. There is an important distinction to be drawn between imaginaries and sentiments mobilised around personalisation, whether by insurance actors or critics, and the calculative devices, the techniques, currently used to assess risks.

In a related vein, Alberto Cevolino and Elena Esposito (2020) elaborate on the social consequences of algorithmic prediction in insurance that might ensue in a move from statistical risk pools to predictive personalised profiles. They assemble a corpus of actuarial, scientific and commercial literature and use it to distil a nuanced imaginary of what a fully realised BDA-driven insurance might involve. They focus specifically on the changes in conceptions on fairness and discrimination that would be entailed in a reconfiguration of risk sharing from 'pool to profile'. Discrimination based on profiles can seem fairer, or more palatable, than discrimination based on membership of a social group with characteristics that may be predictive but are subject to protections. In a 'new alliance between actuarial techniques and digital technologies', Cevolino and Esposito conclude, discrimination may be packaged as a consequence of individual life-style but it would generate classification situations that could affect individual life-chances in ways that are unpredictable.

The papers all share this concern with the far-reaching consequences on social solidarity represented by the steady epistemological turn away from risk pooling. Together, they consider how difficult the tendency towards a 'risk pool of one' is for insurance, whether privately or publicly organised, to put into practice. BDA has been positioned as a means of reversing the information asymmetries that structure insurance, allowing the insurance provider to know more about their policyholder as individual risks. But having access to more data is not the same as knowing individual risk. Instead, the pattern and character of asymmetries – the 'market for lemons' – is constantly transformed by subtle changes in insurance practice and in the socio-economic, political and regulatory context it takes place within.[11] BDA reshapes the kinds of moral hazards and patterns of adverse selection that are produced in insurance relationships. Tracking the consequences of the market experiments that are happening backstage in the insurance-technology industries requires that researchers pay attention not just to the idea of surveillance but to the practices of insurance.

Notes

1. For more detail see: https://www.iii.org/publications/a-firm-foundation-how-insurance-supports-the-economy/introduction/insurance-industry-at-a-glance; https://www.eiopa.europa.eu/tools-and-data/insurance-statistics_en#EuropeanInsuranceOverview on OECD, US and European markets.
2. See Froud et al. (2000), Martin (2002), Krippner (2005, 2012), Christophers (2015). Van der Heide (2019a) highlights the particular paradox of the life industry, one of the major investors in capital markets, being largely neglected in this literature.
3. The migration of mathematical modelling between insurance and finance is not, as Van der Heide (2019b) points out, one way traffic.
4. The insurtech, sometimes rendered as insuretech, neologism refers, broadly speaking, to the appearance of market entrants and processes originating in the tech industries cf. McFall and Moor (2018).
5. See: https://tracxn.com/d/emerging-startups/top-internet-first-insurance-startups-2020; https://www.ft.com/content/58c42412-3e91-46d4-833e-1e11806c5057
6. Lemonade is 5 Years Old! Here's our strategy and results to date: https://www.youtube.com/watch?v=j7Q8SyuHWc0&t=25s
7. In the United States, all states have unfair trade practices laws/regulations, modelled after the National Association of Insurance Commissioners (NAIC) Unfair Trade Practices Act (#880) which prohibits unfair discrimination on the grounds of ethnicity, race, gender, religion, national origin and other protected classes: https://www.naic.org/store/free/MDL-880.pdf
8. Discrimination by gender has been forbidden in the EU since December 2012 (According to The Gender Goods and Services Directive (2004) and the European Court of Justice's ruling on the Test-Achat case (Case C-236/09) of 1 March 2011).
9. Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II) (Text with EEA relevance): https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32009L0138
10. See note 5: https://www.youtube.com/watch?v=j7Q8SyuHWc0&t=25s
11. The outcome of the next US presidential election and the decisions of its newly re-constituted Supreme Court (SCOTUS) may soon provide a stark reminder of this in re-legislating what sorts of healthcare, notably reproductive healthcare, certain insurers are permitted to 'know' (cf. McFall, 2019). At the time of writing, confirmation hearings for Trump nominee Amy Coney Barrett to join the Supreme Court of the United States (SCOTUS) are underway. Barrett's stance on the Affordable Care Act (ACA), and the federal subsidy it provides towards the healthcare costs of some individuals, has been closely scrutinised. The Trump administration made the repeal of ACA a 2016 election pledge and although its provisions have been weakened it remains in force. Republican opposition espouses the idea that subsidy is morally hazardous in protecting individuals from the financial costs of unhealthy behaviour.

Acknowledgements

This editorial and the contributions it introduces are the outcome of conversations at a series of insurance workshops and events that once happened in lovely places like Lecce, Leuven and Seville. Since the Covid-19 pandemic changed everything, conversations have had to move online. In addition to the contributors, referees and editors of BDS we owe thanks to the exceptionally interesting people who shared notes, ideas and queries in the back channels of reply Twitter, especially, Arjen van der Heide, Duncan Minty, Rick Swedloff, Jim Davey, Taylor Spears, Paula Jarzabkowski, Jathan Sadowski and Shoshana Wodinsky.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Liz McFall https://orcid.org/0000-0001-8681-2576
Gert Meyers https://orcid.org/0000-0002-5255-529X
Ine Van Hoyweghen https://orcid.org/0000-0002-9402-2918

References

Akerlof GA (1970) The market for "Lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3): 488–500.

Alborn TL (2002) The first fund managers: Life insurance bonuses in Victorian Britain. *Victorian Studies 45*(1): 65–92.

Austin JL (1962) *How to Do Things with Words*. Cambridge: Harvard University Press.

Baker T (2010) Health insurance, risk, and responsibility after the Patient Protection and Affordable Care Act. *University of Pennsylvania Law Review 159*: 1577–1621.

Baker T and Simon J (eds) (2002) *Embracing Risk: The Changing Culture of Insurance and Responsibility*. Chicago: University of Chicago Press.

Baldwin P (1990) *The politics of social solidarity*. Cambridge: Cambridge University Press.

Ball K and Webster W (2020) Big Data and surveillance: Hype, commercial logics and new intimate spheres. *Big Data & Society 7*(1): 205395172092585–205395172092585.

Barocas S and Selbst AD (2016) Big Data's disparate impact. *California Law Review 104*(3): 671–732.

Barry L and Charpentier A (2020) Personalization as a promise: Can Big Data change the practice of insurance? *Big Data & Society*. Epub ahead of print 22 June 2020. DOI: 10.1177/2053951720935143; http://doi.org/10.1177/2053951720935143

Blasimme A, Vayena E and Van Hoyweghen I (2019) Big data, precision medicine and private insurance: A delicate balancing act. *Big Data & Society*. Epub ahead of print 6 March 2019. DOI: 10.1177/2053951719830111; https://doi.org/10.1177/2053951719830111

Bucher T (2018) *If . . . Then Algorithmic Power and Politics*. Oxford: Oxford University Press.

Cevolini A and Esposito E (2020) From pool to profile: Social consequences of algorithmic prediction in insurance. *Big Data & Society 7*(2): 205395172093922–205395172093911. http://doi.org/10.1177/2053951720939228

Chen H, Chiang R and Storey V (2012) Business intelligence and analytics: From Big Data to big impact. *MIS Quarterly 36*(4): 1165. http://dx.doi.org/10.2307/41703503.

Christophers B (2015) The limits to financialization. *Dialogues in Human Geography 5*(2): 183–200. http://doi.org/10.1177/2043820615588153

Dobbin F (1992) The origins of private social insurance: Public policy and fringe benefits in America. *American Journal of Sociology 97*(5): 1416–1450.

Doctorow C (2020) How to destroy surveillance capitalism. Available at: https://onezero.medium.com/how-to-destroy-surveillance-capitalism-8135e6744d59 (accessed 16 October 2020).

Drechsler L and Benito Sanchez JC (2018) The price is (not) right: Data protection and discrimination in the age of pricing algorithms. *European Journal of Law and Technology 9*(3): 1–31.

EIOPA (European Insurance and Occupational Pensions Authority) (2019) Big Data analytics in motor and health insurance: A thematic review. Available at: https://register.eiopa.europa.eu/Publications/EIOPA_BigDataAnalytics_ThematicReview_April2019.pdf (accessed 16 October 2020).

EIOPA and NAIC (2018) EU-U.S. insurance dialogue project big data issue paper. Available at: https://www.eiopa.europa.eu/sites/default/files/publications/other_documents/181031_eu_us_big_data_issue_paper_publication.pdf (accessed 04 August 2020).

Ericson R, Barry D and Doyle A (2000) The moral hazards of neo-liberalism: Lessons from the private insurance industry. *Economy and Society 29*(4): 532–558.

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Ewald F (1991) Insurance and risk. In: Burchell G, Colin G and Miller P (eds) *The Foucault Effect. Studies in Governmentality*. Chicago: University of Chicago Press, pp.197–210.

Ewald F (2012) *Assurance, Prévention, Prédiction . . . Dans L'univers du Big Data*. Paris: Institut Montparnasse.

FCA (Financial Conduct Authority) (2016) FS16/5: Call for inputs on Big Data in retail general insurance. Available at: https://www.fca.org.uk/publications/feedback-statements/fs16-5-call-inputs-big-data-retail-general-insurance (accessed 4 August 2020).

FCA (Financial Conduct Authority) (2019) *General insurance pricing practices*. Market Study, MS18/1.2.

French S and Kneale J (2009) Excessive financialisation: Insuring lifestyles, enlivening subjects, and everyday spaces of biosocial excess. *Environment and Planning D: Society and Space 27*(6): 1030–1053. http://doi.org/10.1068/d7607

Froud J, Haslam C, Johal S, et al. (2000) Shareholder value and financialization: Consultancy promises, management moves. *Economy and Society 29*(1): 80–110.

Geiger S (2019) Silicon valley, disruption, and the end of uncertainty. *Journal of Cultural Economy 13*(2): 169–184. http://doi.org/10.1080/17530350.2019.1684337

Gellert R, Vries KD, de Hert P, et al. (2013) A comparative analysis of anti-discrimination and data protection legislations. In: Custers B, Calders T, Zarsky T, et al. (eds) *Discrimination and Privacy in the Information Society*. Berlin: Springer, pp.61–90.

Hacking I (1990) *The Taming of Chance*. Cambridge: Cambridge University Press.

Haraway D (2016) *Staying with the Trouble: Making Kin in the Chthulucene*. Durham and London: Duke University Press.

Hendrickx K and Van Hoyweghen I (2018) Solidarity after nature: From biopolitics to cosmopolitics. *Health*. Epub ahead of print 16 September 2018. DOI: 10.1177/1363459318800149; https://doi.org/10.1177/1363459318800149.

Jeanningros H and McFall L (2020) The value of sharing: Branding and behaviour in a life and health insurance company. *Big Data & Society*. Epub ahead of print 10 September 2020. DOI: 10.1177/2053951720950350; http://doi.org/10.1177/2053951720950350

Kelty C (2017) Too much democracy in all the wrong places: Toward a grammar of participation. *Current Anthropology 58*: S77–S90. http://doi.org/10.1086/688705

Kiviat B (2019) The moral limits of predictive practices: The case of credit-based insurance scores. *American Sociological Review 84*(6): 1134–1158. http://doi.org/10.1177/0003122419884917

Krippner G (2005) The financialization of the American economy. *Socio-Economic Review 3*: 173–208.

Krippner G (2012) *Capitalizing on Crisis: The Political Origins of the Rise of Finance*. Cambridge: Harvard University Press.

Lehtonen T-K and Liukko J (2012) The forms and limits of insurance solidarity. *Journal of Business Ethics 103*(S1): 33–44.

Lehtonen T-K and Van Hoyweghen I (2014) Editorial: Insurance and the economization of uncertainty. *Journal of Cultural Economy 7*(4): 532–540. http://doi.org/10.1080/17530350.2013.875929

LEX (2020) Insurance/Lemonade: Pipping their rivals. *Financial Times*, June 9.

Lupton D (2016) *The Quantified Self: A Sociology of Self-Tracking*. Cambridge: Polity.

Marelli L, Lievevrouw E and Van Hoyweghen I (2020) Fit for purpose? The GDPR and the governance of European digital health. *Policy Studies 41*: 447–467.

Martin R (2002) *Financialization of Daily Life*. Philadelphia: Temple University Press.

MAS (Monetary Authority of Singapore) (2020) "Fairness Metrics" to aid responsible AI adoption in financial services. Available at: https://www.mas.gov.sg/news/media-releases/2020/fairness-metrics-to-aid-responsible-ai-adoption-in-financial-services (accessed 1 August 2020).

Mau S (2019) *The Metric Society*. New York: Polity.

McFall L (2011) A 'good, average man': Calculation and the limits of statistics in enrolling insurance customers. *The Sociological Review 59*(4): 661–684.

McFall L (2014) *Devising Consumption: Cultural Economies of Insurance, Credit and Spending*. London: Routledge.

McFall L (2019) Personalizing solidarity? The role of self-tracking in health insurance pricing. *Economy and Society 0*(0): 1–25. DOI: 10.1080/03085147.2019.1570707.

McFall L and Moor L (2018) Who, or what, is insurtech personalizing?: Persons, prices and the historical classifications of risk. *Distinktion: Journal of Social Theory 19*(2): 193–213.

Meyers G (2018) *Behaviour-Based Personalisation in Health Insurance: A Sociology of a Not-yet Market*. KU Leuven: Proefschrift aangeboden tot het verkrijgen van de graad van Doctor in de Sociale Wetenschappen, CESO.

Meyers G and Van Hoyweghen I (2018) Enacting actuarial fairness in insurance: From fair discrimination to behaviour-based fairness. *Science as Culture 27*(4): 413–438.

Meyers G and Hoyweghen IV (2020) 'Happy failures': Experimentation with behaviour-based personalisation in car insurance. *Big Data & Society 7*(1): 205395172091465–14. DOI: 10.1177/2053951720914650.

Minty D (2016) Price optimisation for insurance optimising price; destroying value? *Thinkpiece Chartered Insurance Institute*. Retrieved from https://www.cii.co.uk/learning/learning-content-hub (accessed 20 July 2020).

Mittelstadt BD, Allo P, Taddeo M, et al. (2016) The ethics of algorithms: Mapping the debate. *Big Data & Society*. Epub ahead of print 1 December 2016. DOI: 10.1177/2053951716679679; https://doi.org/10.1177/2053951716679679.

National Association of Insurance Commissioners (2004) Unfair Trade Practices Act (#880). Available at: https://www.naic.org/store/free/MDL-880.pdf (accessed 20 July 2020).

Noble S (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

OECD (2020a) The impact of big data and artificial intelligence (AI) in the insurance sector. Available at: www.oecd.org/finance/Impact-Big-Data-AI-in-the-Insurance-Sector.htm (accessed 20 July 2020).

OECD (2020b) Insurance spending. Available at: https://data.oecd.org/insurance/insurance-spending.htm#indicator-chart (accessed 20 July 2020).

O'Malley P (1996) Risk and responsibility. In: Barry A, Osborne T and Rose N (eds) *Foucault and Political Reason*. Chicago: University of Chicago Press.

O'Neil C (2016) *Weapons of Math Destruction*. New York: Crown.

Porter T (1996) *Trust in numbers: the pursuit of objectivity in pubic life*. Princeton, Princeton University Press.

Prainsack B and Van Hoyweghen I (2020) Shifting solidarities: Personalisation in insurance and medicine. In: Van Hoyweghen I, Pulignano V and Meyers G (eds) *Shifting Solidarities. Trends and Developments in European Societies*. London: Palgrave Macmillan, pp.127–151.

Prince AER and Schwarcz D (2020) Proxy discrimination in the age of artificial intelligence and Big Data. *Iowa Law Review 105*(1257): 1257–1318.

Ralph O (2020) Insurance experts put premium on start-ups reaching the big time, *Financial Times*, July 5. Available at: https://www.ft.com/content/47faafe0-1252-48a7-a6d5-899833629f50 (accessed 1 August 2020).

Rebert L and Van Hoyweghen I (2015) The right to underwrite gender. The Goods & Services Directive and the politics of insurance pricing. *Tijdschrift Voor Genderstudies 18*(4): 413–431.

Ruckenstein M and Granroth J (2020) Algorithms, advertising and the intimacy of surveillance. *Journal of Cultural Economy 13*(1): 12–24.

Schüll ND (2016) Data for life: Wearable technology and the design of self-care. *BioSocieties 11*(3): 317–333.

Swedloff R (2014) Risk classification's Big Data (r)evolution. *Connecticut Insurance Law Journal 21*(1): 339–374.

Tanninen M (2020) Contested technology: Social scientific perspectives of behaviour-based insurance. *Big Data & Society*. Epub ahead of print 7 September 2020. DOI: 10.1177/2053951720942536; http://doi.org/10.1177/2053951720942536.

The Economist (2015) Risk and reward: Data and technology are starting to up-end the insurance business. *The Economist*, March 12. Available at: https://www.economist.com/finance-and-economics/2015/03/12/risk-and-reward?fsrc=scn%2Fln_ec%2Frisk_and_reward (accessed 16 October 2020).

Tironi M and Laurent B (2014) Experiments for industrial exploration: Testing a car sharing system. *CSI-i3 Working Papers Series 14–11*.

van der Heide A (2019a) Dealing with uncertainty: A historical sociology of evaluation practices in UK Life Insurance, 1971-Present, PhD Dissertation, University of Edinburgh, UK.

van der Heide A (2019b) Model migration and rough edges: British actuaries and the ontologies of modelling. *Social Studies of Science 50*(1): 121–144.

Van Hoyweghen I (2007) *Risks in the Making*. *Travels in Life Insurance and Genetics*. Amsterdam: Amsterdam University Press.

Yates J (2005) *Structuring the Information Age: Life Insurance and Technology in the Twentieth Century*. Baltimore: John Hopkins University Press.

Zuboff S (2019) *The Age of Surveillance Capitalism*. London: Profile Books, Kindle Edition.