

Working paper



17

# Personalization as a Promise: Can Big Data Change the Practice of Insurance?

---

Laurence Barry & Arthur Charpentier

Décembre 2019

PARI

PROGRAMME DE RECHERCHE  
SUR L'APPRÉHENSION DES RISQUES  
ET DES INCERTITUDES

# Personalization as a Promise: Can Big Data Change the Practice of Insurance?

---

Laurence Barry<sup>1</sup> & Arthur Charpentier<sup>2</sup>

## Abstract

Ce papier a pour objet une mesure de l'impact des technologies issues du big data sur la tarification des produits d'assurance automobile. La première partie décrit comment le point de vue agrégé construit par les statistiques permet de mettre en évidence des régularités invisibles au niveau individuel. Malgré une segmentation très granulaire en assurance automobile, l'approche est restée classificatoire, posant comme hypothèse l'identité de risque des individus d'une même classe. La deuxième partie met en avant le retournement de perspective induit par le big data dans l'analyse des données ; avec leur volume et les nouveaux algorithmes, le point de vue agrégé se trouve remis en cause. L'hypothèse d'homogénéité des classes devient de plus en plus difficile à maintenir, d'autant que l'analyse prédictive se vante de sa capacité de prévoir le résultat au niveau individuel. La troisième partie étudie l'influence des boîtiers télématiques à même d'importer le nouveau paradigme en assurance automobile. Pourtant, une lecture des articles de recherches les plus récents sur une tarification automobile incluant ces nouvelles données montre que le saut épistémologique, du moins pour l'instant, n'a pas eu lieu.

## Introduction

Prior to the emergence of statistical knowledge, accidents were perceived and accepted as the decision of God (Bernstein, 1998). In the Middle Ages, the leper was excluded from the community in a religious ceremony that acknowledged the divine decision; his exclusion was also the oblivion of the existence of the disease. In the seventeenth century by contrast, the treatment of the plague is drastically different; Foucault describes how a district would be put into quarantine, and each individual would be attributed a window,

---

<sup>1</sup> Chaire PARI, Laurence.Barry@datastorm.fr

<sup>2</sup> Université du Québec à Montréal (UQAM).

Nous souhaitons remercier Pierre François et Thierry Cohignac pour leurs précieux commentaires sur une première version de ce papier. Toutes les erreurs et omissions restent de notre fait.

where from he had to give a daily indication of his health: the disease was not ignored but precisely recorded in ledgers where the dead and the living were constantly kept apart (Foucault 1975: 228–30). In both cases, a human hand intended to separate the sick from the healthy in a dichotomous and exact way (Foucault 2009: 9-10).

With inoculation in the eighteenth century, a new treatment of diseases is taking shape that can serve as a case in point to illustrate the changing perception on aleatory events. As compared to earlier attempts to categorize each individual as sick or healthy, the vaccine works with the understanding that some will not survive – that cannot be identified in advance- but yet the practice is desirable on the population as a whole. Bernoulli thus publishes a study that shows that, if the mortality rate caused by the vaccination is below 11%, the practice will improve the life expectancy of the population as a whole, by three years on average. He therefore concludes that the vaccination is scientifically desirable. For d’Alembert, Bernoulli’s so called demonstration does not stand due to the high risk *for the individual* of dying in the short term due to smallpox (Colombo and Diamanti, 2015); d’Alembert still reasons in the individualistic approach of the previous centuries. Bernoulli by contrast inaugurates the statistical thinking that accompanies the emergence of modern states (Desrosières, 2008a; Hacking, 1990) and, with them, insurance mechanisms that manage aleatory events at the collective level (Ewald, 1986).

Current technological development combined to the collection of huge amounts of personal data lead some scientists today to claim that we have reached a new epistemological turn as concerns data analysis. Pentland thus emphatically contends that: “we are discovering that we can begin to explain many things— crashes, revolutions, bubbles— that previously appeared to be random ‘acts of God’” (Pentland, 2014, 9). This new apprehension of risk seems indeed to be currently taking shape in the treatment of sickness; the statistical and collective approach that supported the vaccination of populations is now doubled by predictive medicine, focused on the individual. Instead of considering the efficiency of a medicine on the whole population, the aim is to adjust the diagnosis and the treatment to each patient specifically (Herrero et al., 2016; Samerski, 2018).

But where personalization seems to bring true benefits in many domains such as health, the individualization of risks in insurance is more problematic. Insurance mechanisms are indeed intrinsically collective, as they are built on the pooling of risks (Baker 2002, 6; Lehtonen and Liukko 2015, 158). What could then be the meaning of an individualized risk measurement? Furthermore, the new technologies are often defined as *predictive* analytics (Siegel, 2016). Yet insurance has always been about prediction: risk measurement traditionally consisted in the transformation of individual uncertainty concerning the future into something stable, measurable and thus predictable, on the collective (Ericson and Doyle, 2004). The new techniques are therefore particularly challenging for traditional conceptions of insurance.

Yet the paradigm shift seems in many ways already in action: in health insurance, some insurers have started calculating individual risk scores (McFall, 2019). In many countries telematics devices are being installed in cars to collect behavioural and continuous data, and insurers have started implementing Usage Based Insurance products (UBI) (Meyers and Hoyweghen, 2018; Ptolemus, 2012). The aim of this paper is to measure the extent of the shift of paradigm, on these motor products. While Meyers and Hoyweghen focused on the implied changes in conceptions of fairness, our aim here is rather to understand if and how the UBI products actually serve the calculation of an *individualized* risk premium.

The first part demonstrates how insurance was built by the actual *creation* of homogeneity that was artificially obtained thanks to risk classification. Insurance accompanies welfare state mechanisms in the construction of statistical tools for the management of aleatory events. The second part shows how current big data technologies claim to imply a reversal of perspective that is deeply at odds with the core of the insurance practices. The last part is a reading of existing research papers on risk measurement with telematics. It shows their limited impact on pricing techniques. Whether such a position can be maintained in the long run remains an open question, the more so as all researchers have proved the relevance of telematics parameters for crash prediction and risk measurement.

## **The Emergence of Insurance Mechanisms: Building a View on the Aggregate**

The development of statistics during the nineteenth century shows the existence of a regularity at the collective level of events that cannot be explained at the individual one (Foucault 2009: 65-66).<sup>3</sup> For Foucault, a new object of knowledge is taking shape, the population. The statistical approach to diseases marks the emergence of a new rationality that changes the perspective on the individual level by focusing on the collective one:

When the different possibilities of death or contamination are calculated, the result is that the disease no longer appears in this solid relationship of the prevailing disease to its place or milieu, but *as a distribution of cases in a population circumscribed in time or space*. Consequently, the notion of case

---

<sup>3</sup> Interestingly, Foucault shows how the new knowledge was in fact built upon the same ledgers first created for the sake of the management of the plague: by compiling those ledgers for the city of London, Graunt could indeed show that the proportion of death per cause was stable over time and place, including the proportion of suicides (Foucault 2009: 74; see also Mazur 2016).

appears, *which is not the individual case, but a way of individualizing the collective phenomenon of the disease*, or of collectivizing the phenomena, integrating individual phenomena within a collective field, but in the form of quantification and of the rational and identifiable (Foucault 2009: 60, emphasis added).

While statistics helped develop a new management of collective phenomena at large, they also gave new tools to cope with uncertainty. By the end of the nineteenth century, in most western societies,<sup>4</sup> accidents were therefore perceived along Durkheim's terminology as "social facts:" they cannot be predicted at the individual level, but get some predictability at the collective one. Knowledge of aleatory events can be obtained on the aggregate, once the micro level is abandoned. A vertical perspective might be a good metaphor of the statistical gaze (Desrosières 2014, 169). Although statistics builds upon data collected at the individual level, there is a kind of orthogonality of viewpoints; either you stick to the (horizontal) description of the individual, but then the collective level remains out of reach, or you give up the precise knowledge of individuals in order to access the larger picture. This dual understanding of knowledge seems to characterize the period. In his 1896 introductory lesson on probability, using particles and the newly established laws of the kinetic theory of gases as metaphor, Poincaré states the following:

You ask me to predict events that will occur in the future; would I unfortunately know the laws of these phenomena, I couldn't manage without inextricable calculations and I would have to give up answering. Yet since I am lucky enough to ignore them, I will answer immediately. And, most extraordinarily, my answer will be correct (Poincaré, 1912, p. 3).

Insurance mechanisms took shape in this context: they indeed build upon the adoption of a collective and statistical perspective on aleatory events. As Kolmogorov puts it "the epistemological value of probability theory is based on the fact that chance phenomena, considered collectively and on a grand scale, create non-random regularity" (Gnedenko and Kolmogorov, 1954). Following Knight's terminology, insurance can be defined as the transformation of unknown individual uncertainty, or chance, into a measurable aggregate risk (Knight, 1985). Technically, it consists in the pooling of uncertainty, and the application of the law of large numbers.

The first mechanisms of welfare also belong to this epistemological paradigm: "the only means that we have to solve the difficulty is to pool the risks and the advantages, which means to accept in advance that *without knowing who will bear the risk and who will benefit from the advantage*, risks will be bore collectively

---

<sup>4</sup> The United States were late in this matter and established their first workers compensation acts in 1908 (Haller, 1988).

and access to social advantages will be open to all” (Bourgeois in Ewald 1986: 370, emphasis added, personal translation). Work accidents could therefore be taken care of by society as a whole, and only by society as a whole.

Hence the emergence of a level of reality that treats the individual as *a case for the understanding of the group* at the heart of insurance mechanisms (Foucault 2009, 60; see also Ewald, 1986). But this level is actually *produced*, we would like to argue, by statistics as a practice, and the necessary quantification implied by the new science. Statistical knowledge was indeed built by the collection of data via questionnaires (e.g the censuses) and the quantification of the world (Hacking, 1990), that also imposed a vision of homogeneity among people. Actually, the homogenization occurs twice: once in the choice of *what is not asked*, therefore not quantified at all; and once in the averaging of what is measured and collected. The Belgian mathematician Quételet is known for being among the first to have applied probabilistic techniques -formerly used in astronomy- to human phenomena, and to have therefore universalized the use of probability calculus (Ewald, 1986: 147; Stigler, 1986: 161). Measuring the size of the torso of soldiers, Quételet noticed that it is distributed along a bell curve; until then this curve, formalized by Gauss, was used in astrophysics to model the error in measurement of the position of stars. For Gauss, the true position of the star is the one where he has a pick of observations, hence at the mean. By analogy, Quételet thus concludes that the deviation from the mean is also a form of error: the actual torso should be compared to the ideal torso of an ideal man, “the average man” (Desrosières, 2008b). By so doing, he reduces the individual measure to its contribution to the average or, in Foucault’s terms, as a case within a collective phenomenon.

Quantification, confirms Porter, creates standardization and « averages away » the noise of individuals:

Inevitably, meanings are lost. Quantification is a powerful agency of standardization because it imposes order on hazy thinking, but this depends on the license it provides *to ignore or reconfigure much of what is difficult or obscure*. As nineteenth-century statisticians liked to boast, their science *averaged away everything contingent, accidental, inexplicable, or personal*, and left only large-scale regularities (Porter 1996: 85, emphasis added).

In insurance, the statistical treatment of human phenomena therefore assumes an equality of all the members of the group in their exposure to the accident. Desrosières thus explains how quantification creates “classes of equivalence:” within a given class, all the members of the group are considered to have the exact same behavior (Desrosières, 2008c).<sup>5</sup> Statistics indeed build upon a

---

<sup>5</sup> In the same strand of thought, Le Bras contends that mortality tables could not be computed before a notion of universal equality had emerged (Le Bras 2000: 124-128). Curtis further

collective perspective that ignores individual peculiarities and transform them into variables: by interpreting the deviation from the mean as error, Quételet actually *constructs* the homogeneity of the group around the single measure of the means when he *decides* that it represents the whole. This homogeneity is thus both an assumption and a construct.

Insurance products were built and still function along these same lines: they consist in the a priori definition of classes that are supposed to reflect identical risks. Describing current practices, Paefgen et al., describe the process as follows:

In order to differentiate the risk of insurance policies, actuaries use a set of rate factors to separate policies into groups (i.e., tariff classes). The construction of tariff classes is ultimately a clustering task. Each tariff class corresponds to a certain combination of rate factor categories or intervals in the case of continuous rate factors. For each tariff class, actuaries analyze historical claims data to arrive at a reliable estimate of the corresponding pure premium, that is, the minimum required payment per policy to cover the expected losses from its class (Paefgen et al., 2013, 193).

Each class thus functions as a group deemed homogenous as concerns the risk (Lemaire et al., 2016: 42). Reflecting on motor insurance today, Lemaire et al., further mention that the classes are known to artificially group people that are not exactly the same from an insurance viewpoint. This understanding triggered the introduction of the bonus-malus system, also called “merit rating plan.” The system differentiates *within* the class according to *individual* claims record:

In most developed countries, insurers have implemented bonus-malus systems (BMS), which modify the premium according to past claims history. *One of the main goals of BMS is to reduce adverse selection by including indirectly information that could not be taken into account explicitly such as respect of the driving code, alcohol use, mileage driven, etc* (Lemaire, Park, and Wang 2016, 40, emphasis added).

The bonus-malus system, we would like to argue, constitutes a further ramification of classes; in the Taiwanese example studied by Lemaire, it introduces 10 different levels of rates within the rating system (Lemaire et al., 2016, 47). However refined (and the segmentation in automobile insurance is known to be very granular – see for instance Weidner et al., 215), the classification is limited by two kinds of technical constraints. The first is that

---

suggests that the population as an object of knowledge is rendered possible by the French Revolution, that puts all members of society on an equal standing (Curtis, 2002, p. 530).

each class should have sufficient volume so as to remain statistically valid and to show the needed collective regularity granted by the law of large numbers (Pfaegen et al., 2013, 193). The second is that, since the data is collected via questionnaires, there is a practical limit to the amount of information that can be gathered.<sup>6</sup>

Until very recently at least, insurance practice thus belonged to the scientific paradigm inherited from nineteenth century: it consisted in human quantification, i.e. the collection of a limited amount of individuals' information via questionnaires, the constitution of classes based on these variables, the assumption of homogeneity within the classes that allowed for an apprehension of risk on the aggregate; very little could be said on the individual, but a lot could be deduced from the groups manually constructed.

## **Big Data and the Reversal of Perspective**

In an uncritical manner, something radically new seems to happen with the emergence of big data: a “revolution” is supposed to be taking place (Mayer-Schönberger and Cukier, 2013) that will radically transform the data we collect (both *what* and *how* we collect it), and the manner in which it is treated and used. This part tries to sort out from the big claims on big data the real conceptual changes implied for insurance: the latter is a good entry point, we argue, to show what is really changing when compared to the classical statistics paradigm. This part is focused on this conceptual shift, whereas the next part gives a detailed analysis of the current status in the particular field of motor insurance.

The novelty of the data is often characterized along the “three Vs:” volume, velocity and variety (Billot et al., 2017; Kitchin, 2014). In relation to the previous part and insurance, we'd rather insist here on the process of collection itself that has the three Vs as a consequence. The datafication indeed entails that data is generated *directly from activity*, without the need to humanly quantify described in the previous part. Indeed, online navigation, captors and/or bodily sensors continuously collect indicators that transform human behavior into something *natively numerical*. Big data consists of online traces, “bread crumbs” as Pentland (2014, 8) defines them, supposedly delivering exhaustive information on the individuals at stake (Kitchin, 2014, 1).

---

<sup>6</sup> Lemaire et al., (2016, 42) add that the segmentation “ends when the cost of including more risk factors exceeds the profit that the additional classification would create.” This is part of the technical limits mentioned above.

Where insurance is concerned, there are today a couple of domains where big data technologies should, or could, change practices: health, homeowners and automobile insurance. All of them imply the collection of data via sensors: bracelets or wearables that capture bodily information (Gilmore, 2016; Lupton, 2016); home automation sensors to prevent fire, leak and moisture (Kulesa, 2016); and telematics captors that collect location, speed and acceleration in vehicles. The first major change brought by the captors is that they unearth a vast amount of data concerning the individual; tracking the movements of the car or the body, the sensors provide *behavioral* and *continuous* data, two characteristics at odds with traditional insurance data.

In automobile insurance for instance, underwriting information traditionally consisted of drivers' demographic and cars' technical details asked for upfront, at the issuance of the policy; the recent possibility to collect real time information is a real challenge for products where prices are most of the time updated once a year (Denuit et al., 2019). Besides, the data was static, with the sole exception of the abovementioned bonus-malus system, where an update of claims history is performed at the time of the policy renewal. Moreover, as Ayuso et al., notice, "information about driving habits <were> not considered directly, on the grounds that driving style and intensity could not hitherto *be measured objectively*" (Ayuso et al., 2019, 736, emphasis added). By contrast, behavioral data are now considered more trustworthy than demographic, static parameters (Paefgen et al., 2013, 193). Here lies the first conceptual change: in the digital age, the questionnaires are indeed generally perceived as an obsolete, cumbersome and inaccurate process for data collection (Arnoux et al., 2017; Schwartz et al., 2013; Yarkoni, 2010; Youyou et al., 2015).

Big data is not only about the kind and quantity of data, but also about the algorithms that treat them. Cardon et al., (2018) describe how deep learning, as a new family of models, have led the current paradigm shift (in image recognition) at the beginning of the 2010s. Although formally theorized by the end of the 1950s, the strength and potential of neural networks could not materialize before huge data bases were constituted to feed them, together with computer capacities. With other and less trendy machine learning<sup>7</sup> algorithms, neural networks constitute a large family of "predictive analytics" models (Siegel, 2016), that share a couple of specificities when compared to classical statistical analysis. The first is that they eschew any "a priori explicit modeling" of the characteristics of the data (Cardon et al., 2018, 3). The second is that they are far more complex, involving a huge amount of parameters; for Breiman, the simplicity of traditional statistical models has to be traded off for better accuracy (Breiman, 2001, 206-208). Third, instead of reducing the dimensionality by either deleting variables or aggregating them,

---

<sup>7</sup> The first definition of machine learning is attributed to Mitchell (1997). It includes a vast variety of algorithms, deep neural networks being one of them.

neural networks function best with a large quantity of features. Finally, traditional models aimed at reproducing “the true mechanisms” behind the observations (Saporta, 2017, 42–44); new algorithms by contrast are geared towards prediction rather than finding “causal inference,” hence correlations are enough (Charpentier et al., 2018; see also Siegel, 2016, 130-135).

Deep learning models are used today in a variety of domains dealing with signal analysis (image, text, sound). They are capable, based on a very large quantity of observations, of extracting patterns from the data without human intervention: “the key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure” (LeCun et al., 2015, 436). One of the major consequences of this transformation in the perspective of this paper, is that the *work* of quantification, that conditioned the statistical analysis of the previous era, is being bypassed by both the manner in which data is collected (without questionnaires) and the manner in which it is processed. In insurance, the a priori classification that allowed risk measurement might no longer appear technically necessary, or even legitimate.

The volume of information collected at the individual level makes indeed each person significantly different from the others. This is documented in various domains; studying retail recommendation systems, Mackenzie shows how models have moved from a targeting based on broad demographic parameters to the taking into account of the history of purchases of each consumers, with a declared intention to personalize the recommendation (Mackenzie, 2018). Others even invoke “segments of one” (Weed, 2017), leading researchers to conclude that we are in a process of a “re-personalization of pricing” (Moor and Lury, 2018, 501).

In insurance, the hypothesis of homogeneity within a class, paramount for risk measurement, becomes therefore difficult to maintain. In parallel with other domains, this is leading to a *personalization of risk*, in the form of “individual risk scores” (McFall, 2019; Meyers and Hoyweghen, 2018). However paradoxical, the idea of adjusting the premium to the individual risk is not new; it actually surfaced with growing computer capacities and the neoliberal ideology in the 1980s (Frezal and Barry, 2019; Walters, 1981). Walters, in a seminal address to the American Casualty Actuarial Society thus stated in 1981 that insurance is about the transfer of the *individual's* risk to the insurer, without redistribution between insureds (Walters, 1981, 5). Technically, limiting the pooling *to exact same risks*, also called chance or probability solidarity (De Witt and Van Eeghen, 1984; Lehtonen and Liukko, 2015, 2011), with no subsidies between different risk levels, was never achieved.

Applied to insurance the big data paradigm promises to finally achieve this personalization of risk: it would mean looking at each individual in his irreducible differences and assess *his* risk, *as if he was his own class*. As Tselentis

et al., put it: “each driver could be assigned a probability of crash involvement based on his/her driving behavior” (Tselentis et al., 2017, p. 140). Pushed to the extreme, the imaginary of the individualized risk with perfect accuracy conveys the possibility of predicting *individual* claims occurrence. In such a case, the insurer would be able to classify people in a dichotomous manner, separating those that *will have* accidents from those that *won't*. Echoing the pre-statistical era, the imaginary of perfect prediction means the erasure of statistical aggregates. But by deconstructing the pooling process, perfect knowledge would also paradoxically put an end to insurance. More precisely, while insurance was historically built by bringing up collective regularities and acknowledging the opacity of the individual, big data technologies promise to lift this opacity by delivering regularities *between* individuals as look-alikes, without the need to resort to the aggregative viewpoint. But to what extent could big data technologies deliver such a knowledge?

## The Personalization of Risk?

Since 2010, say Cardon et Al., “deep neural networks provoke the same disruption in information communities dealing with signal, voice, speech or text” (Cardon et al., 2018, 3). Likewise, Charpentier et al., mention numerous applications in credit scoring, fraud detection and targeted marketing (Charpentier et al., 2018, 4). As mentioned in the previous part, various devices and the Internet of Things open the way for certain branches of insurance to move from a traditional apprehension of risks based on classification and averages, to “the new paradigm.” Does the access to big data lead to an apprehension of risks without resorting to any a priori classification? The aim of this part is to appreciate the extent of this shift.

Telematics is the oldest connected device in insurance and should therefore have the most mature applications.<sup>8</sup> Besides, contrary to health insurance that is widely regulated against risk individualization (Ewald, 2014; McFall, 2019), automobile insurance regulation, currently at least, gives more freedom to the insurer. This might be due to the repeated promise, by both the industry and public institutions, that these devices, coupled to proper insurance products, could lead to significant reduction in car accidents and fatalities (Husnjak et al., 2015; Ptolemus, 2012; Tselentis et al., 2016, 364).<sup>9</sup> We will therefore focus

---

<sup>8</sup> The first test program for auto insurance with telematics was indeed implemented in 2003 (Ptolemus, 2012, 129).

<sup>9</sup> As early as 2001, the National Highway Safety Traffic Administration (NHSTA) conducted a research involving 100 vehicles with sensors and video cameras to collect data and study factors explaining crashes, drivers behaviors overtime etc... (See National Highway Traffic Safety Administration, 2006).

on telematics and motor insurance, although conceptually at least, the potential shift is the same in the other domains.

This part is thus focused on the way telematics displaced - or not - risk apprehension and pricing in motor insurance. It is based on a reading of predictive analytics papers published on Usage Based Insurance (UBI) and the use of telematics data over the last decade. A lot is said in blogs and insurer's sites, that contributes to fueling the "promise of personalization" (e.g. Perret, 2018; Sandquist, 2019); our choice was instead to focus on actual data analyses and results, together with a couple of reviews (see corpus). We might add here a notice and a disclaimer: our study shows that the disruption actually did not happen. This paper is not, however, an attempt in explaining *why* insurance practice did not change. Although some hypotheses for a future study will be given in the conclusion, the scope of the paper is to show *how* the new data is used, without actually changing existing models. Given the importance of data analysis for insurance and the existence of telematics products for over fifteen years, the number of studies, that amounts to a few dozens, seems scarce. Furthermore, papers coming from actuarial journals were found astonishingly few; more promising was road safety research, where most of the studies mentioned here thus come from. Others have noticed this issue before us, and suggest that, the data being proprietary to insurance companies, their access to researchers remains restricted (Ma et al., 2018; see also Baecke and Bocca, 2017). This would mean that the models exist but are not made public. We will evoke in the conclusion other reasons for the limited number of articles in general, and from the actuarial field in particular. But the main one might be, simply put, that nothing revolutionary has happened yet.

As McFall mentions, the pricing based on apps and Internet of Things is at odds not solely with the conceptual frame of insurance, but also with its infrastructure and working practices (McFall, 2019, 54). It is possible that despite the promise for a personalized price advanced by the UBI providers (Meyers and Hoyweghen, 2018), the actuarial models do not or not fully incorporate the data delivered by the new devices. For Bian et al., "insurers *and* researchers are still trying to find an appropriate path for UBI" (Bian et al. 2018, 21, emphasis added). Here too, as Lury and Wood suggest for other domains, "the technology that facilitates personalized pricing <might be> currently somewhat ahead of its use" (Moor and Lury, 2018, 510).

Since all the studies confirm the predictive power of the variables provided by the devices on accidents, there is however a consensus among researchers that a change needs to occur in the near future. As Verbelen et al., put it:

This potentially high dimensional telematics data, collected on the fly, forces pricing actuaries to change their current practice, both from a business as well as a statistical point of view. *New statistical models have to be developed to adequately set premiums based on an individual policyholder's driving habits and style and the current*

*literature on insurance rating does not adequately address this question* (Verbelen et al., 2018, 2, emphasis added).

In their seminal paper on deep learning, LeCun et al., emphasize the machine's capacity to infer automatically from data the predictive patterns. The application of different algorithms to specific kinds of data is given at the outset: "deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech" (LeCun et al., 2015, 436). By analogy, transforming the statistical approach to automobile insurance via telematics would mean to treat the data as a signal and develop specific algorithms capable of automatically detecting the patterns in the raw data. In such a case, we would argue, the classification would indeed disappear and the individual's behavior would prevail over the group approach. None of the papers examined in this research pretend to do such an epistemological leap. What seems rather to be happening is the step prior to deep learning, described by LeCun et al., as follows:

For decades, constructing a pattern-recognition or machine-learning system required careful engineering and *considerable domain expertise to design a feature extractor* that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input (LeCun et al., 2015, 436).

Actually, looking at an insured as providing data *as signals* rather than one vector of static information is in itself radically new. Hence research currently concentrates on what LeCun et al., have coined "extracting features" with the help of domain experts. The existence of these features are then translated into variables that aggregate the continuous information provided by the device into monthly or yearly averages, easily comparable to the static data available without UBI (e.g Bian et al., 2018).<sup>10</sup> The approach remains therefore in many ways in line with the traditional one:

The added value of involving industry experts in the development of the predictive model is investigated by *augmenting the model with expert-based telematics variables*. These are additional features *that are not automatically extrapolated from the raw data*. Instead, these features are created as a smart combination of metrics from which

---

<sup>10</sup> In this perspective, Denuit et al., contends that: "continuous signals are certainly appealing as many embarked devices produce real measures. *Another approach consists in recording a number of events, or to round a continuous signal in multiples of a natural unit. This makes the mechanism more transparent, at the cost of a negligible loss of accuracy*" (Denuit et al., 2019, 5, emphasis added).

experts expect a significant impact on accident risk (e.g. night trips during the weekend) (Baecke and Bocca, 2017, 72, emphasis added).

Research therefore currently concentrates on the a priori definition of patterns that are then tested for significance. Those patterns are themselves conditioned on the type of information stored by the box (Fourcade and Healy, 2017, 289). As Kitchin (2014, 4-5) put it “data are not simply natural and essential elements that are abstracted from the world in neutral and objective ways and can be accepted at face value (...) Systems are designed to capture certain data.” One might distinguish between products as Pay As You Drive (PAYD) that collect actual mileage, and Pay How You Drive (PHYD) (Tselentis et al., 2016). Some refinement may be introduced in the PAYD data when the mileage is itself split between hours of the day and/or the type of road (Denuit et al., 2019; Verbelen et al., 2018). PHYD products are more recent and incorporate, besides mileage, other parameters claimed to describe one’s “driving style,” and are therefore more personalized than the first.<sup>11</sup> Fruitful also is the distinction proposed by Tselentis et al., between travel and driving behaviour variables: both are considered behavioural and therefore different from traditional demographic data. But the travel behaviour of the driver relates to his “strategic” choices concerning the type of road network and the time of the trips. The driving behaviour of the driver by contrast reflects his “operational” choices, i.e. the manner in which he manipulates his vehicle at real time within the existing traffic conditions (Tselentis et al., 2017).

All the variables are humanly created in this manner: besides days and night trips (and the necessary categorization of time slots implied), one might mention speeding, measured as a proportion of trips above a certain threshold fixed by the researchers (Lahrmann et al., 2012), distribution of trips between urban or rural area - a classification again imposed on raw data (Ayuso et al., 2019; Verbelen et al., 2018). Further personalization occurs with the attempt to characterize driving styles. The data involves pure telematics data such as hard brakes, accelerations and cornering, counted as events above a threshold (usually pre-defined by the box provider). Additional information is sometimes merged with the telematics data base on a trip basis before aggregating the information at the monthly/annual driver level. When trying to model “driving style,” the context of *the trip* becomes indeed relevant: researchers therefore add information concerning real time traffic speed on the same road segment (Hu et al., 2018), or the percent above speed limits (Ma et al., 2018).

---

<sup>11</sup> Technically, the PAYD programs intend to introduce in pricing a measure of the exposure (longer distances supposedly implying more chances of accidents), whereas the PHYD intend to adjust price also to the intensity of risk, independently from exposure.

Far from being an “agnostic data analytics” (Kitchin, 2014, 4), the studies actually reproduce preconceptions of risky behaviours, further tested for significance. In some cases, the work of numerical translation is not immediate; Jin et al., for instance propose to look at the “familiarity” of the driver with their driving routes (Jin et al., 2018). It was quantified by the number of recurring routes taken on overall on a monthly basis. In the same strand of thought, some studies redefine “driving styles” as predefined categories thanks to telematics.<sup>12</sup> The study of driving styles for the improvement of road safety has indeed a long history, going back to 1949 (Sagberg et al., 2015), where descriptive studies were undertaken. Among the relevant traits was “aggressiveness;” it is being redefined with telematics as “risky speeding profiles (irregular, instantaneous and abrupt changes in vehicle speed), improper vehicle position maintenance (quick changes in lateral vehicle position) and inconsistent or excessive acceleration and deceleration (harsh take-off and braking) (Meiring and Myburgh, 2015, 30657). Adding video information (to detect line changes and tailgating other vehicles) at the trip level, Kumtepe et al., train a classifier to define aggressiveness in line with an external observer’s judgment (Kumtepe et al., 2016); again, the measure is not inferred from raw data but added to it as a subjectively created category. Interestingly, what these examples suggest is that quantification (or the human construction of variables) has not disappeared but instead of being imposed upfront in the questions and possible categories of answers, it is built bottom up from the data itself.

Most of the time, the researchers recommend to *add* the new variables to the existing classification (Ayuso et al., 2019; Baecke and Bocca, 2017; Ferreira and Minikel, 2012; Paefgen et al., 2013; Verbelen et al., 2018), as the new variables function best in combination with the traditional ones. Sometimes however (e.g. Ayuso et al., 2019, 737), the hybrid model is seen as temporary. For Weidner et al., in the transitory period the UBI data determines a discount on the traditional tariff (Weidner et al., 2017, 229), which is where we stand now (Meyers and Hoyweghen, 2018). Besides, some mention that telematics variables might become a necessary input to existing models in replacement of other variables that are being removed by regulation:<sup>13</sup>

---

<sup>12</sup> Outstanding in the literature on “driving style” is the study made by Weidner et al., since they generate stochastic driving sequences and derive from them six driving styles based on the sole telematics information (Weidner et al., 2017).

<sup>13</sup> Interestingly, in their study of credit scoring, Fourcade and Healy (2017b: 12) show that the need to replace forbidden variables was the main trigger to move to behavioral variables in this field.

Insurance companies are facing difficult pricing decisions, as several variables commonly used are challenged by regulators. The EU now forbids the use of gender rating. Territory is being challenged in the U.S. as a substitute for race. Insurers are being pressured *to find new variables that predict accidents more accurately and are socially acceptable* (Lemaire et al., 2016, 66, emphasis added)

Taking this path, some researchers thus prove that gender is found redundant when telematics devices are implemented (Verbelen et al., 2018). Yet none of them recommend at the moment to replace the existing models. The intent is *not* to disrupt insurance practices but rather to *refine* the existing segmentation thanks to new parameters; they thus adopt the classificatory logic, and as Paefgen et al., (2013, 193) put it: “ideally, one might derive a one-dimensional aggregated variable that adds only one further dimension to actuarial tables.”

## Conclusion

This paper started by illustrating the paradigm shift introduced by predictive analytics in medicine. Reflecting on this potential revolution, Wilson and Nicholls contend that:

Personalized medicine and personal genomics have been described as paradigm-shifting technologies in medicine, although their pace of implementation may perhaps be better described as a slow revolution in health care. There are significant challenges in moving from traditional genetics, with its focus on monogenic disorders with significant implications for health of a very small proportion of the population, to the development of genetic profiling approaches which are useful for screening, risk assessment, disease prevention, and health promotion. The idea of personalized medicine as fully individualized medicine has still to be realized, and is likely unrealistic. (Wilson and Nicholls 2015, 17)

In many ways, their diagnosis also applies to insurance. The “slow revolution” seems indeed to be taking place by the introduction of additional behavioral and dynamic variables into existing models, therefore making the segmentation more granular. In this perspective, rather than an epistemological leap, the individualization of risk can be seen as an axis of refinement that actually started long before telematics and big data (Lemaire et al., 2016). Such a trend towards more segmentation is not new, except maybe for the kind of data at stake.

While this finding is somehow surprising, many reasons can be advanced to explain it. Understanding the practical causes for the non-occurrence of the shift was out of the scope of this paper and might well be the object of another study. Some hypotheses can be advanced at this stage. The increased segmentation that comes with the “personalization of risk” challenges the business model of insurers. It also bears a reputation risk for the insurer, when for instance a driver faces a rate increase because of “aggressiveness,” without the occurrence of a claim. Other reasons might be similar to those encountered in medicine, and deal with existing infrastructure and practices. Actuaries for instance have no doubt tried over the last decade to assimilate the new predictive techniques (Ollivier, 2017), yet they might be doing so without a full fledged abandonment of existing models and infrastructures, either to maintain their specificity or because they did not find the new ones sufficiently convincing.

This paper focused on the conceptual challenge, that might in itself participate in the explanation: as we have tried to show here, insurance is based on the pooling of risks, with an underlying assumption of homogeneity. The technique of risk classification reflects this anchoring into a group-based approach. Predictive analytics, claiming to replace it by an individual one conceptually jeopardizes the very possibility of insurance. Indeed, at the extreme limit of the axis of segmentation (however realistic this point might be), would be a situation where an individual insured would be *known* to be heading to an accident. More realistically, it would lead to very high rates for the riskier persons, to a point where insurance would become unaffordable to them.

Should the unrealistic scenario occur and crashes become predictable, however, car accidents would not be uncertain events any longer; they would therefore fall out of the scope of insurance, as they would not be “risks” any more. From this viewpoint, the absence of actuarial models willing to consider the radical end of the spectrum is reassuring. The conceptual resilience of insurance so to speak, and the slow pace of change simply reaffirm that insurance is, and will remain, about the collective management of uncertain events, that demands *imperfect* knowledge on the individuals.

## References

- Arnoux, P.H., Xu, A., Boyette, N., Mahmud, J., Akkiraju, R., Sinha, V., 2017. 25 Tweets to Know you: A New Model to Predict Personality with Social Media. AAAI Publications, Eleventh International AAAI Conference on Web and Social Media 472–475.

- Ayuso, M., Guillen, M., Nielsen, J.P., 2019. Improving automobile insurance ratemaking using telematics: incorporating mileage and driver behaviour data. *Transportation* 46, 735–752.
- Baecke, P., Bocca, L., 2017. The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems* 98, 69–79.
- Baker, T., 2002. Risk, Insurance, and (the Social Construction of) Responsibility. University of Connecticut School of Law Articles and Working Papers.
- Bernstein, P.L., 1998. *Against the Gods: The Remarkable Story of Risk*. Wiley, New York.
- Bian, Y., Yang, C., Zhao, J.L., Liang, L., 2018. Good drivers pay less: A study of usage-based vehicle insurance models. *Transportation Research Part A: Policy and Practice* 107, 20–34.
- Billot, R., Bothorel, C., Lenca, P., 2017. Introduction au big data et application à l'assurance, in: Corlosquet-Habart, M., Janssen, J. (Eds.), . ISTE Editions, London, pp. 17–39.
- Breiman, L., 2001. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statist. Sci.* 16, 199–231.
- Cardon, D., Cointet, J.-P., Mazières, A., 2018. La revanche des neurones. *Reseaux* n° 211, 173–220.
- Charpentier, A., Flachaire, E., Ly, A., 2018. *Économétrie & Machine Learning*. HAL working papers hal-01568851.
- Colombo, C., Diamanti, M., 2015. The smallpox vaccine: the dispute between Bernoulli and d'Alembert and the calculus of probabilities. *Lett Mat Int* 2, 185–192.
- Curtis, B., 2002. Foucault on Governmentality and Population: The Impossible Discovery. *Canadian Journal of Sociology / Cahiers canadiens de sociologie* 27, 505–533.
- De Witt, G.W., Van Eeghen, J., 1984. Rate Making and Society's Sense of Fairness. *ASTIN Bulletin* 151–164.
- Denuit, M., Guillen, M., Trufin, J., 2019. Multivariate credibility modelling for usage-based motor insurance pricing with behavioural data. *Annals of Actuarial Science* 1–22.
- Desrosières, A., 2014. *Prouver et gouverner. La découverte*, Paris.
- Desrosières, A., 2008a. *L'argument statistique. II, Gouverner par les nombres*. Presses des mines, Paris.

- Desrosières, A., 2008b. Quételet et la sociologie quantitative. Du piédestal à l'oubli, in: Desrosières, A. (Ed.), *Pour Une Sociologie Historique de La Quantification. L'argument Statistique 1*. Presses de l'école des Mines, Paris, pp. 239–257.
- Desrosières, A., 2008c. *L'argument statistique. I, Pour une sociologie historique de la quantification*. Presses de l'école des Mines, Paris.
- Ericson, R.V., Doyle, A., 2004. *Uncertain Business: Risk, Insurance, and the Limits of Knowledge*. University of Toronto Press.
- Ewald, F., 2014. *Assurance-Prévention-Prédiction. Big data et société – Institut Montparnasse*. URL <http://www.institut-montparnasse.eu/nos-publications/les-inegalites-territoriales-et-sociales-2-12/> (accessed 6.18.18).
- Ewald, F., 1986. *L'Etat Providence*. Grasset.
- Ferreira, J., Minikel, E., 2012. Measuring per Mile Risk for Pay-As-You-Drive Automobile Insurance. *Transportation Research Record* 2297, 97–103.
- Foucault, M., 2009. *Security, Territory, Population. Lectures at the Collège de France. 1977-1978*, eds.M. Senellart, A.I. Davidson, F. Ewald and A. Fontana, trans. G. Burchell. Palgrave Mcmillan, Basingstoke.
- Foucault, M., 1975. *Surveiller et punir*. Gallimard, Paris.
- Fourcade, M., Healy, K., 2017a. Categories All the Way Down. *Historical Social Research* 42, 286–296.
- Fourcade, M., Healy, K., 2017b. Seeing like a Market. *Socio-economic review* 15, 9–29.
- Frezal, S., Barry, L., 2019. Fairness in Uncertainty: Some Limits and Misinterpretations of Actuarial Fairness. *J Bus Ethics*.
- Gilmore, J.N., 2016. Everywear: The quantified self and wearable fitness technologies. *New Media & Society* 18, 2524–2539.
- Gnedenko, B.V., Kolmogorov, A.N., 1954. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, Mass.
- Hacking, I., 1990. *The Taming of Chance*. Cambridge University Press, Cambridge, Mass.
- Haller, J.S., 1988. Industrial accidents--worker compensation laws and the medical response. *West J Med* 148, 341–348.

- Heath, J., 2007. Reasonable Restrictions on Underwriting, in: *Insurance Ethics for a More Ethical World*, Research in Ethical Issues in Organizations. Emerald Group Publishing Limited, pp. 127–159.
- Hu, X., Zhu, X., Ma, Y.-L., Chiu, Y.-C., Tang, Q., 2018. Advancing usage-based insurance – a contextual driving risk modelling and analysis approach. *IET Intelligent Transport Systems* 13, 453–460.
- Husnjak, S., Peraković, D., Forenbacher, I., Mumdziev, M., 2015. Telematics System in Usage Based Motor Insurance. *Procedia Engineering*, 25th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2014 100, 816–825.
- Jin, W., Deng, Y., Jiang, H., Xie, Q., Shen, W., Han, W., 2018. Latent class analysis of accident risks in usage-based insurance: Evidence from Beijing. *Accident Analysis & Prevention* 115, 79–88.
- Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1, 2053951714528481.
- Knight, F.H., 1985. *Risk, uncertainty and profit*. University of Chicago press, Chicago.
- Kulesa, T., 2016. 8 sensors to help you create a smart home - IBM Internet of Things blog [WWW Document]. Internet of Things blog. URL <https://www.ibm.com/blogs/internet-of-things/sensors-smart-home/> (accessed 7.24.19).
- Kumtepe, O., Akar, G.B., Yuncu, E., 2016. Driver aggressiveness detection via multisensory data fusion. *J Image Video Proc.* 2016, 5. <https://doi.org/10.1186/s13640-016-0106-9>
- Lahrman, H., Agerholm, N., Tradisauskas, N., Berthelsen, K.K., Harms, L., 2012. Pay as You Speed, ISA with incentives for not speeding: Results and interpretation of speed data. *Accident Analysis & Prevention, Intelligent Speed Adaptation + Construction Projects* 48, 17–28.
- Le Bras, H., 2000. *Naissance de la mortalité*. Le Seuil, Paris.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444.
- Lehtonen, T.-K., Liukko, J., 2015. Producing Solidarity, Inequality and Exclusion Through Insurance. *Res Publica* 21, 155–169.
- Lehtonen, T.-K., Liukko, J., 2011. The Forms and Limits of Insurance Solidarity. *J Bus Ethics* 103, 33–44.
- Lemaire, J., Park, S.C., Wang, K.C., 2016. The Use of ANnual Mileage as a Rating Variable. *ASTIN Bulletin: The Journal of the IAA* 46, 39–69.

- Lupton, D., 2016. The diverse domains of quantified selves: self-tracking modes and dataveillance. *Economy and Society* 45, 101–122.
- Ma, Y.-L., Zhu, X., Hu, X., Chiu, Y.-C., 2018. The use of context-sensitive insurance telematics data in auto insurance rate making. *Transportation Research Part A: Policy and Practice* 113, 243–258.
- Mackenzie, A., 2018. Personalization and probabilities: Impersonal propensities in online grocery shopping. *Big Data & Society* 5 (1), 1-15.
- Mayer-Schönberger, V., Cukier, K., 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.
- Mazur, D.J., 2016. Analyzing and interpreting “imperfect” Big Data in the 1600s. *Big Data & Society* 3, 2053951715609082.
- McFall, L., 2019. Personalizing solidarity? The role of self-tracking in health insurance pricing. *Economy and Society* 48, 52–76.
- Meiring, G.A.M., Myburgh, H.C., 2015. A Review of Intelligent Driving Style Analysis Systems and Related Artificial Intelligence Algorithms. *Sensors (Basel)* 15, 30653–30682.
- Meyers, G., Hoyweghen, I.V., 2018. Enacting Actuarial Fairness in Insurance: From Fair Discrimination to Behaviour-based Fairness. *Science as Culture* 27, 413–438.
- Mitchell, T.M., 1997. *Machine Learning*, 1 edition. ed. McGraw-Hill Education, New York.
- Moor, L., Lury, C., 2018. Price and the person: markets, discrimination, and personhood. *Journal of Cultural Economy* 11, 501–513.
- National Highway Traffic Safety Administration, 2006. The 100-Car Naturalistic Driving Study, Phase II - Results of the 100-Car Field Experiment [WWW Document]. URL <https://rosap.nhtl.bts.gov/view/dot/37370> (accessed 7.27.19).
- Ollivier, C., 2017. *L’actuaire à la croisée des chemins*. Working paper Chaire Pari.
- Paefgen, J., Staake, T., Thiesse, F., 2013. Evaluation and aggregation of pay-as-you-drive insurance rate factors: A classification analysis approach. *Decision Support Systems* 56, 192–201.
- Pentland, A., 2014. *Social physics: how good ideas spread-the lessons from a new science*. The Penguin Press, New York.

- Perret, P., 2018. Nous passons d'une relation prime-indemnisation a une relation d'accompagnement. *Revue Banque* 25–27.
- Poincaré, H., 1912. *Calcul des probabilités*. Paris, Gauthier-Villars.
- Porter, T.M., 1996. *Trust in Numbers*, Reprint edition. ed. Princeton University Press, Princeton, N.J.
- Ptolemus, C.G., 2012. *Insurance Telematics - from Experimentation to Mass Implementation*.
- Sagberg, F., Selpi, Bianchi Piccinini, G.F., Engström, J., 2015. A Review of Research on Driving Styles and Road Safety. *Hum Factors* 57, 1248–1275.
- Sandquist, E., 2019. Personalizing the insurance customer experience.
- Saporta, G., 2017. Des methodes classiques d'analyses de donnees au big data, in: Corlosquet-Habart, M., Janssen, J. (Eds.), *Le Big Data Pour Les Compagnies d'assurance*. London, pp. 41–55.
- Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M.E.P., Ungar, L.H., 2013. Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE* 8, e73791.
- Siegel, E., 2016. *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*, 2 edition. ed. Wiley, Hoboken, New Jersey.
- Stigler, S., 1986. *The history of statistics. The measurement of uncertainty before 1900*. Harvard university press, Cambridge.
- Tselentis, D.I., Yannis, G., Vlahogianni, E.I., 2017. Innovative motor insurance schemes: A review of current practices and emerging challenges. *Accident Analysis & Prevention* 98, 139–148.
- Tselentis, D.I., Yannis, G., Vlahogianni, E.I., 2016. Innovative Insurance Schemes: Pay as/how You Drive. *Transportation Research Procedia, Transport Research Arena TRA2016* 14, 362–371.
- Verbelen, R., Antonio, K., Claeskens, G., 2018. Unravelling the predictive power of telematics data in car insurance pricing. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 67, 1275–1304.
- Walters, M.A., 1981. Risk Classification Standards. *Proceedings of the Casualty Actuarial Society* 68, 1–23.
- Weed, K., 2017. The future of marketing? Consumer segments of one [WWW Document]. Think with Google. URL

<https://www.thinkwithgoogle.com/consumer-insights/unilever-consumer-marketing-segmentation/> (accessed 6.23.18).

- Weidner, W., Transchel, F.W.G., Weidner, R., 2017. Telematic driving profile classification in car insurance pricing. *Annals of Actuarial Science* 11, 213–236.
- Wilson, B.J., Nicholls, S.G., 2015. The Human Genome Project, and recent advances in personalized genomics. *Risk Manag Healthc Policy* 8, 9–20.
- Yarkoni, T., 2010. Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal in Research on Personality* 44, 363–373.
- Youyou, W., Kosinski, M., Stillwell, D., 2015. Computer-based personality judgments are more accurate than those made by humans. *PNAS* 112, 1036–1040.

# PARI

PROGRAMME DE RECHERCHE  
SUR L'APPRÉHENSION DES RISQUES  
ET DES INCERTITUDES

**PARI, placé sous l'égide de la Fondation Institut Europlace de Finance en partenariat avec l'ENSAE/Excess et Sciences Po, a une double mission de recherche et de diffusion de connaissances.**

Elle s'intéresse aux évolutions du secteur de l'assurance qui fait face à une série de ruptures : financière, réglementaire, technologique. Dans ce nouvel environnement, nos anciens outils d'appréhension des risques seront bientôt obsolètes. PARI a ainsi pour objectifs d'identifier leur champ de pertinence et de comprendre leur émergence et leur utilisation.

**L'impact de ses travaux se concentre sur trois champs :**

- les politiques de régulation prudentielle dans un contexte où Solvabilité 2 bouleverse les mesures de solvabilité et de rentabilité (fin du premier cycle de la chaire);
- les solutions d'assurance, à l'heure où le big data déplace l'assureur vers un rôle préventif, créant des attentes de personnalisation des tarifs et de conseil individualisé ;
- les technologies de data science appliquées à l'assurance, modifiant la conception, l'appréhension et la gestion des risques.

Dans ce cadre, la chaire PARI bénéficie de ressources apportées par Actuaris, la CCR, Generali, Groupama, la MGEN et Thélem.

Elle est co-portée par **Pierre François**, chercheur au CNRS, doyen de l'Ecole Doctorale de Sciences Po et **Laurence Barry**, chercheur à Datastorm, la filiale de valorisation de la recherche de l'ENSAE.

## PARTENAIRES

