

Working paper

32

INTELLIGENCE ARTIFICIELLE ET INDIVIDUALISATION DES GARANTIES EN ASSURANCE : ÉCHEC OU RETARD A L'ALLUMAGE ?

Arthur CHARPENTIER
Université du Québec à Montréal (UQAM)

Xavier VAMPARYS
xavier.vamparys@telecom-paris.fr

Octobre 2023

PARI

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

INTELLIGENCE ARTIFICIELLE ET INDIVIDUALISATION DES GARANTIES EN ASSURANCE : ECHEC OU RETARD A L'ALLUMAGE ?

Arthur CHARPENTIER

Université du Québec à Montréal (UQAM)

Xavier VAMPARYS

xavier.vamparys@telecom-paris.fr

Résumé

Derrière l'engouement suscité par l'utilisation de l'intelligence artificielle dans le secteur de l'assurance se cache une réalité plus nuancée. Prenons par exemple l'assurance automobile et l'assurance santé. Les tentatives d'usage de l'intelligence artificielle à des fins de tarification n'y ont pour le moment pas engendré le « changement de paradigme » pourtant annoncé. Pourquoi ? Plusieurs raisons peuvent être invoquées, qui vont des fondamentaux de l'assurance au choix assumé de certains assureurs de ne pas toucher à la mutualisation des risques, dans un contexte marqué par deux tendances opposées : la recherche par les consommateurs de services et produits toujours plus personnalisés et le refus sociétal de solutions qui laisseraient certains individus au bord du chemin de l'assurabilité.

L'histoire des usages de l'intelligence artificielle (IA) dans le champ économique n'est pas celle d'une succession de réussites, comme on se plaît trop souvent à la raconter. Prenons le cas de l'assurance. Les assureurs semblent être des candidats de choix pour transformer, grâce à l'IA, l'incertitude en connaissance, l'inconnu en risque, le risque en prime ou en cotisation. Les assureurs collectent d'ailleurs depuis bien longtemps de nombreuses données sur leurs clients et les utilisent à des fins de tarification au moyen de modèles statistiques prédictifs. Si on appréhende ces termes de façon imprécise, on pourrait presque dire qu'une forme « primitive » d'IA y a cours depuis longtemps.

Pourtant, plusieurs études récentes, dont l'une au titre évocateur - « La révolution qui n'a pas eu lieu, Télématiques et assurance automobile dans les années 2010 » (François et al., 2022) -, indiquent que les tentatives d'usage de l'IA à des fins de tarification en assurance automobile¹ et en assurance santé² n'ont pas été couronnées de succès. Un article de *l'Argus de l'assurance* (Madeline, 2019, cité par Benoît et al., 2021, p. 1), revue de référence du secteur, indique que le Big Data reste en « *quêtes d'usage* » dans l'assurance en France, à l'état de « *défis* » ou de « *promesses* ».

A contrario, on doit constater que le mouvement d'individualisation des produits et services est désormais définitivement engagé, nonobstant la nature collective de l'assurance, fondée sur la mutualisation des risques. Ce mouvement n'est pas propre au secteur de l'assurance. Dans tous les domaines, le consommateur souhaite bénéficier d'offres personnalisées. Et l'IA est un instrument puissant en matière d'individualisation : elle permet de mieux connaître le consommateur, ses besoins et ses aspirations. Comme l'écrit Minty (2023) « *la personnalisation n'est pas près de disparaître. En effet, il ne s'agit ni d'un phénomène technologique, ni d'un phénomène assurantiel. Il s'agit d'une tendance sociale à laquelle les entreprises s'adaptent (...). [Elles] apprennent à appliquer les technologies pour rester pertinentes et rentables. Rien de nouveau donc* »³. C'est également ce qui fait dire au responsable de l'innovation chez Swiss Re (Meyers, 2018, p. 72) :

« Si vous n'avez pas d'objet connecté pour suivre votre état de santé, il vous sera pratiquement impossible de souscrire une assurance-vie. Si ce n'est pas le cas dans les 5 à 10 prochaines années, je suis convaincu que ce sera le cas dans les 20 prochaines années. Comment puis-je affirmer une chose aussi scandaleuse ? C'est très simple. L'avènement du Big Data entraînera une révolution de l'expérience client similaire à celle que nous avons connue dans les domaines de la communication et de la planification »⁴.

Pour Jawish (2022, p. 107) les sociétés d'assurance sont « *entrées, comme tant d'autres, dans l'ère servicielle, où la satisfaction client règne en maître* ». La réponse apportée par les

¹ Le marché français de la santé et de la prévoyance s'élevait à 40 milliards de cotisations en 2021. V. rapport France Assureurs « Données 2021 », juin 2022.

² Le marché français de l'assurance automobile s'élevait à 24 milliards de cotisations en 2021. V. rapport France Assureurs « Données 2021 », juin 2022.

³ Traduit de l'anglais.

⁴ Traduit de l'anglais.

assureurs aux besoins des assurés se doit donc de tenir compte des particularités de chacun d'entre eux.

Où en est-on donc de l'individualisation par l'IA des couvertures d'assurance santé et d'assurance automobile ? Le choix de ces deux branches d'assurance n'est pas totalement arbitraire. Tout d'abord, parce que les assureurs disposent depuis longtemps de données, même si leur nature – nom, âge, genre, etc. – est différente de celles aujourd'hui collectées, qui tiennent davantage au comportement des assurés. Ensuite, nos véhicules sont équipés – ou peuvent aisément l'être – de boîtiers qui enregistrent de nombreuses données sur notre « style » de conduite : la vitesse, les accélérations, les longueurs de freinage, l'utilisation du téléphone portable ou la vitesse dans les virages⁵. D'autre part, les offres d'assurance « *Pay How You Drive* » (PHYD) ou « *Pay As You Drive* » (PAYD) (Verbelen et al., 2018 ; Cevolini et al., 2022) sont proposées aux assurés depuis près de dix ans maintenant. En outre, des objets connectés (notamment des montres) capables de capter des données d'activité, des mesures biologiques ou des données de santé sont désormais assez largement utilisés par le public, créant de l'espace pour des produits PAYL ou « *Pay As You Live* » (Ernst&Young, 2015), même si concrètement ce type de couverture est plutôt rarement souscrit (Cevolini et al., 2020). Nous disposons désormais d'un recul suffisant sur ces produits pour discerner l'usage qui y est fait de l'IA et son impact sur la tarification. C'est sur cet usage que se sont d'ailleurs positionnées un certain nombre d'*insurtech*. Par ailleurs, les marchés de l'assurance automobile et de la santé sont très concurrentiels et la pression sur les prix y est très forte⁶. Comme l'indique Minty (2023), « *la télématique a été considérée par de nombreux marchés d'assurance de premier plan comme le "meilleur point de départ" pour l'utilisation de technologies de souscription personnalisée* »⁷. Enfin, assurance santé et assurance automobile connectées partagent certaines caractéristiques : comme l'indiquait en 2014 un responsable de l'innovation de Swiss Re, « *la télématique pour les voitures c'est comme les objets connectés pour les personnes* »⁸ (Meyers, 2018, p. 127)⁹. Aussi, les pratiques en assurance automobile influencent fortement celles de l'assurance santé (Meyers, 2018, p. 127).

Nous tenterons d'identifier les principales causes du décalage entre la disruption annoncée et la réalité et ce qui fait dire à Meyers (2018, p. 20) que « *la personnalisation basée sur le comportement est en "cours d'élaboration", mais elle n'est pas (encore ?) stabilisée* »¹⁰. On pourra identifier onze raisons qui constituent autant d'obstacles fondamentaux au changement radical de paradigme qui nous est annoncé, avec des garanties d'assurance qui refléteraient très fidèlement le risque de chacun : le caractère collectif de l'assurance, les convictions de certaines entreprises d'assurance, les

⁵ V. par ex. l'offre Direct Assurance, <https://blog.direct-assurance.fr/innovation/direct-assurance-de-a-a-z-y-comme-youdrive>

⁶ Les tarifs en assurance automobile sont relativement faibles car les produits offerts sont des produits « d'appel ».

⁷ Traduit de l'anglais.

⁸ Traduit de l'anglais.

⁹ Même si comme l'indique cet employé de Swiss Re on parle beaucoup plus librement de ses écarts de conduite (automobile) que de son hygiène de vie (Meyers, 2018, p. 127).

¹⁰ Traduit de l'anglais.

difficultés liées à la collecte de données, l'intérêt modéré des assureurs pour les nouveaux modèles qui permettraient une individualisation et des assurés pour les objets connectés, l'obstacle de l'explicabilité, l'inertie des assureurs, leur faible proactivité, une réglementation sur l'utilisation des données personnelles qui laisse relativement peu de place à l'individualisation de la tarification, une prime qui ne reflète pas exclusivement le profil de risque d'un individu et, enfin, des considérations morales ou sociétales.

1. Le caractère collectif de l'assurance

L'assurance est une activité qui repose fondamentalement sur une vision collective des risques apportés par chacun des assurés. Un assureur ne peut pas déterminer par avance si tel ou tel individu va subir un sinistre. La survenance de ce sinistre est, pour un individu donné, de l'ordre de l'aléatoire, de l'incertitude (Frezal, 2018 ; Landes, 2015). L'assureur peut néanmoins, à partir d'observations passées de sinistralité pour des risques comparables, estimer la distribution de la sinistralité d'un groupe d'assurés, sa moyenne ou sa variance (Cevolini et al. 2022). « *Tant qu'un assuré individuel n'est pas traité comme le membre d'un groupe, il est impossible de connaître sa perte attendue, car, dans la pratique, ce concept est statistique et basé sur des probabilités de groupe* »¹¹ note Abraham (1985, p.423). Pour Barry et al. (2020) l'assurance se définit « *comme la transformation d'une incertitude individuelle inconnue, ou d'un hasard, en un risque global mesurable. Techniquement, il s'agit de la mise en commun de l'incertitude et de l'application de la loi des grands nombres* »¹². Ou encore, pour McFall (2019, p. 56), « *l'assurance est solidaire parce qu'elle transforme les entités en catégories groupées et porteuses de risques. Le risque doit être regroupé parce qu'il ne peut pas être calculé au niveau individuel. Aussi, les profils de risque individuels sont dérivés de l'appartenance à des groupes définis* »¹³. Enfin, sous la plume de von Mises, cité par Charpentier (2022) :

« Lorsque l'on parle de "probabilité de décès", le sens exact de cette expression ne peut être défini que de la manière suivante. Nous ne devons pas penser à un individu, mais à une certaine classe dans son ensemble, par exemple "tous les hommes assurés âgés de quarante et un ans vivant dans un pays donné et n'exerçant pas certaines professions dangereuses". Une probabilité de décès est attachée à la classe d'hommes ou à une autre classe qui peut être définie de manière similaire. Nous ne pouvons rien dire sur la probabilité de décès d'un individu, même si nous connaissons en détail son état de vie et de santé. L'expression "probabilité de décès", lorsqu'elle se réfère à une seule personne, n'a aucune signification pour nous »¹⁴.

En réalité, cette idée d'une approche collective de l'incertitude n'est pas nouvelle. McFall (2011, p. 668) cite ainsi un prospectus d'une compagnie d'assurance datant de 1846 :

« L'assurance-vie en tant que système repose sur le fait que la vie humaine, qui est proverbiallement la plus incertaine de toutes les choses, suit pourtant, dans l'ensemble, une loi fixe ... alors que nous ne pouvons pas dire combien de temps un homme peut vivre, nous sommes capables, grâce à l'étude de la mortalité, de prédire

¹¹ Traduit de l'anglais.

¹² Traduit de l'anglais.

¹³ Traduit de l'anglais.

¹⁴ Traduit de l'anglais.

avec une singulière exactitude combien d'hommes sur un grand nombre mourront chaque année jusqu'à ce que toutes les vies soient éteintes... »¹⁵.

Comme l'écrit Ewald (2012, p. 48) :

« le probable est l'évènement vu du côté du tout, de la banque (du casino), de dieu, de celui qui peut jouer un nombre indéfini de fois : le résultat des tirages tendra toujours plus à rejoindre leur probabilité. Le risque c'est le même évènement vu du côté de celui qui n'a qu'un coup à jouer, qui ne peut savoir ce qui sortira du tirage, mais qui peut calculer ce que dans l'incertitude, il est susceptible de gagner ou de perdre selon la décision qu'il prendra ».

François et al. (2022, p. 1) écrivent que *« l'assureur n'est pas en mesure de savoir si ce client particulier sera effectivement confronté au risque contre lequel il souhaite se couvrir. L'assureur peut savoir en revanche si, en moyenne et au sein d'une population donnée, un individu sur 100, sur 1 000 ou sur 10 000 verra le risque se réaliser ».* Pour Charpentier et al. (2020, p.76), citant Ewald (1986), l'assurance distingue *« entre le dommage que subit tel ou tel individu – c'est affaire de chance ou malchance – et la perte liée au dommage dont l'attribution est, quant à elle, toujours collective et sociale ».*

La prétention de l'IA dans l'appréciation des risques est diamétralement opposée à l'approche probabiliste visée ci-dessus. Il s'agit en effet avec l'IA d'attribuer à chaque individu un « score » dont le calcul ne dépend plus du truchement d'une collectivité ou d'un groupe. Comme l'écrit Barry (2019, p. 18), citant Siegel (2016, p. 16), *« alors que la prévision macroéconomique permet d'estimer le nombre total de cônes de glace qui seront achetés dans le Nebraska le mois prochain, l'analyse prédictive permet de déterminer quels sont les individus qui seront le plus probablement vus avec un cône en main à ce moment-là ».*

Les éléments techniques classiques de l'assurabilité – le caractère aléatoire de l'évènement, la perte maximum, la perte moyenne par occurrence, la période de temps entre deux occurrences pour en citer quelques-uns (Berliner, 1985, p. 325) – disparaissent et sont remplacés par un nombre. La logique n'est plus celle de la régularité à l'intérieur d'un collectif mais celle de la « singularité » (Ewald, 2012, p. 71). Barry et al. (2020, p.2) écrivent que *« les nouvelles techniques, bien que présentant une certaine familiarité avec les mécanismes d'assurance, semblent poser un défi particulier aux conceptions traditionnelles de l'assurance puisqu'elles prétendent prédire le cas individuel plutôt que le groupe »¹⁶.* Pour Barry encore (2020, p. 1), *« là où auparavant on se contentait de catégoriser les personnes dans des classes d'équivalence assez larges, les techniques actuelles cherchent à appliquer à chacun un nombre (...) [à l'écart] d'une catégorisation de type classique ».* On peut d'ailleurs légitimement s'interroger sur le sens statistique même de ce nombre comme le font Barry (2020) ou Charpentier (2023).

A vrai dire, cet objectif est la limite extrême d'un mouvement qui, lui, est ancien : la segmentation des risques (Barry et al., 2020 ; Bénéplanc et al., 2022, p. 142).

¹⁵ Traduit de l'anglais.

¹⁶ Traduit de l'anglais.

Celle-ci consiste à créer des groupes homogènes partageant certaines caractéristiques distinctives. Comme le constatent Barry et al. (2019, p. 16) on voit au cours du XXème siècle un long mouvement de segmentation de plus en plus fine des risques et un déplacement de la « *perspective d'une équité reposant sur la mise en commun de l'aléa (...) [vers] une 'démarche' scientifique visant un idéal de parfait ajustement à l'individu* ». Comme l'écrivent Bamman et al. (2014, p. 153), que cite Barry (2019, p. 17), l'apprentissage-machine « *offre une abondance de techniques de modélisation qui minimisent le besoin d'hypothèses catégorielles* »¹⁷.

Mais, on pourra émettre deux objections : l'individualisation des primes par une connaissance précise de chaque risque relève d'une utopique perfection à laquelle on peut ne pas croire. Tout ce qui peut arriver à un assuré individuel n'est pas fatalement déterminé par avance ou inscrit dans son génome. L'assurance n'est pas – pour le moment ? – une science exacte (Barry, 2019). Certains experts¹⁸ croient en la possibilité d'une connaissance parfaite des risques. Comme l'écrit O'Neil (2016, p. 164) cité par McFall (2020, p. 5) :

*« Aujourd'hui, avec l'évolution de la science des données et des réseaux d'ordinateurs, l'assurance est confrontée à un changement fondamental. Avec toujours plus d'informations disponibles - y compris les données de nos génomes, nos habitudes de sommeil, d'exercice et d'alimentation, ainsi que la qualité de notre conduite - les assureurs calculeront de plus en plus le risque pour l'individu et s'affranchiront des généralités de la masse d'individus »*¹⁹.

Toutefois, dans un univers où on connaît parfaitement le risque, a-t-on encore besoin d'assurance ? L'assurance suppose l'aléa. Or, l'information chasse l'aléa et, partant, l'assurance, selon un effet que certains économistes appellent « Hirshleifer », économiste qui a mis en lumière le lien entre information et assurabilité (Gollier, 2015). Selon Gollier, le « *risque devien[drait] alors inégalité par la révélation (...) totale de l'information* » (2015, p. 199). A l'assurance se substituerait une capacité plus ou moins importante des « assurés » à économiser pour faire face aux sinistres à venir. Le monde de la banque l'emporterait alors sur celui de l'assurance. Pour Maisnier et al. (2022, p. 41-42) :

« La conséquence (en admettant que les modèles soient justes) est que chaque individu paiera la valeur de son propre risque. Ce faisant, l'assureur proposera une baisse du niveau global de ses primes puisque les bons assurés auront une prime plus faible (ce qui attirera d'autres bons risques) et les mauvais seront résiliés (ou incités à partir vers un autre assureur n'ayant pas encore réalisé sa révolution du big data). Cependant, une fois que tous les acteurs se seront adaptés (ou auront disparu), il n'y aura théoriquement plus d'intérêt à l'assurance : les bons risques auront tout intérêt à s'auto-assurer (pour ne pas avoir à payer les frais de l'assureur) et les mauvais risques ne pourront simplement plus être assurés (...).

¹⁷ Traduit de l'anglais.

¹⁸ L. Alexandre par exemple, cité par Ewald (2013).

¹⁹ Traduit de l'anglais.

Ce modèle de segmentation à outrance qui exclut de facto la mutualisation du risque porterait ainsi en germe la mort programmée de la majorité du marché de l'assurance dès lors que les uniques raisons qui pousseraient les gens à s'assurer seraient l'obligation légale (pour la responsabilité civile par exemple) et l'aversion au risque ».

S'ajoute un autre argument, plus technique – ou statistique –, avancé par certains experts : Abraham (1985, p. 423) relève que « *la plupart des sinistres individuels ne sont pas suffisamment solides sur le plan statistique pour justifier une tarification individuelle* ». Appliquer un raisonnement statistique à des échantillons composés d'un seul individu paraît futile, sauf à considérer que l'apprentissage machine est une technologie à part qui n'est pas contrainte par les lois de la statistique.

Enfin, plus on est sur des groupes de taille réduite, moins l'exercice de tarification est robuste (Barry 2019, p. 11). Même avec toutes les variables désirées, un risque imprévu peut se matérialiser qui remet en question l'équilibre financier du groupe restreint. Charpentier et al. (2015) démontrent que l'hypothèse d'une recherche de forte segmentation dans un marché concurrentiel – avec comme objectif pour certains assureurs d'attirer des « bons » assurés aux primes réduites – engendre une situation qu'ils dénomment la « *spirale de la segmentation* » (Charpentier et al., 2015, p.61) - ou « *spirale infernale* » (Davet, 2011, p. 44) – dans laquelle « *quelques rares sociétés sont (potentiellement) à l'équilibre, avec des parts de marché très faibles et une variabilité du résultat importante, et d'autres perdent de l'argent (en moyenne avec des parts de marché beaucoup plus importantes)* ». En pratique, plus le portefeuille est petit plus la variabilité du résultat sera importante, avec comme conséquence pour les assureurs concernés l'obligation réglementaire de mobiliser des capitaux supplémentaires. Zajdenweber (2015), s'appuyant sur des théorèmes mathématiques analysés par Feller (1968), va jusqu'à affirmer que la constitution de portefeuilles très homogènes (et donc fortement « segmentés ») ne permet pas d'obtenir la plus faible variance possible et milite donc pour la constitution de portefeuilles les plus diversifiés possibles. Enfin, comme l'affirme Mayaux, une « *bonne mutualisation suppose d'atteindre une masse critique* » (Mayaux, 2003), ou une « *bonne dilution des risques* » (Mayaux, 2022, n°33).

En somme, Maisnier et al. (2020, p.42) indiquent que :

« L'assurance ne peut fonctionner que sur une population et non sur des risques individuels. Evidemment, le rôle de l'assureur consiste à identifier, sélectionner et diviser les risques. Afin d'avoir une mesure fine du risque il va effectivement chercher à procéder à une individualisation du risque. Mais celle-ci repose sur une individualité moyenne permise par le regroupement de risques similaires et qui n'a de sens que relativement à l'ensemble de la population assurée. Le marché de l'assurance doit donc réussir à trouver un équilibre subtil entre le principe original de la mutualisation des risques et la segmentation fine des assurés permises par les nouvelles technologies ».

Ewald (2012, p. 56) ne dit pas autre chose :

« L'assurance n'a pas vraiment besoin d'une visée prédictive individualisante. Son problème est de constituer des groupes homogènes, son projet est de mutualiser des risques pour effectuer des compensations. Il lui faut identifier pour classer dans le bon groupe. L'identification est au service de la mutualisation. De ce point de vue, le projet prédictif contemporain lui échappe. Il n'est pas assurantiel ».

2. Les convictions de certaines formes d'entreprises d'assurance

Certaines formes d'entreprise d'assurance peuvent faire le choix d'un refus d'une individualisation excessive des couvertures. On pensera d'abord aux mutuelles d'assurance complémentaire (par exemple Macif, GMF, Maif, Matmut, SMABTP). Ces sociétés sont des personnes morales de droit privé ayant un objet non commercial. Elles n'ont donc pas de but lucratif et « *ni les adhérents ni les dirigeants (...) ne recherchent un profit par leur appartenance à la société. Le seul profit dont ils peuvent bénéficier, c'est de cotiser à un moindre coût pour garantir les risques qu'ils font prendre en charge. La société d'assurance mutuelle a pour objet de leur faire réaliser des économies* » (Bigot, 2011, p. 238). Elles ne cherchent donc pas *a priori* à optimiser leurs résultats en individualisant fortement leurs couvertures.

C'est le cas également des mutuelles d'assurance régies par le Code de la mutualité (MGEN par exemple). Parfois dénommées « mutuelles 45 »²⁰ ou « complémentaires-maladie », leur activité se concentre essentiellement autour de l'assurance maladie et de la prévoyance, même si le type d'assurance qu'elles peuvent pratiquer est sensiblement plus large²¹.

Ces mutuelles garantissent à leurs « membres participants » - définis comme les personnes physiques qui bénéficient des prestations de la mutuelle²², en d'autres termes les assurés – le règlement intégral des engagements qu'elles contractent à leur égard²³. Comme dans les cas des sociétés d'assurance mutuelles, les membres participants sont à la fois sociétaires de la mutuelle et assurés, ce qui crée entre eux, mais également entre eux et la société, une solidarité de fait. Une partie importante des mutuelles servaient initialement une population identifiée par l'exercice d'une profession (mutuelle générale de l'Education nationale par exemple), d'une activité (mutuelle des motards par exemple) ou d'une localisation commune (mutuelle de Poitiers par exemple), même si elles se sont depuis ouvertes à un public plus large. On notera que les contrats proposés par les mutuelles, notamment en assurance-santé individuelle, sont historiquement très « solidaires » car pratiquant peu la sélection (prohibée s'agissant de l'état de santé²⁴, mais possible en tenant compte de l'âge par exemple). L'individualisation des garanties d'assurance est donc relativement faible.

²⁰ Date de création de la Sécurité sociale, à l'occasion de laquelle le rôle des mutuelles sera reprécisé.

²¹ Art. L. 111-1 du Code de la mutualité.

²² Article L. 114-1 du Code de la mutualité.

²³ Article L. 211-2 du Code de la mutualité.

²⁴ V. art. L. 112-1, 2ème alinéa du code de la mutualité

Enfin, les institutions de prévoyance (par exemple Malakof Humanis ou Klesia) promeuvent également la solidarité et ne segmentent que faiblement leurs risques.

Sur les vingt premiers assureurs français, cinq ont choisi de faire mention, directement ou implicitement, de la solidarité dans leur raison d'être : CNP Assurances (action pour une « *société inclusive et durable et apport au plus grand nombre des solutions* »)²⁵, Natixis Assurance, via la raison d'être de BPCE et des Caisses d'Épargne (pour ces dernières, « *favoriser la solidarité* »), AG2R (« *solidarités professionnelles et intergénérationnelles* »), Crédit Mutuel Arkea (« *entreprise solidaire, éthique et inclusive* ») et Maif (« *garantir un réel mieux commun* »). Si on ajoute les sociétés qui font figurer la solidarité parmi leurs engagements, ce sont quinze sociétés sur vingt qui déclarent avoir la solidarité au cœur de leur projet d'entreprise²⁶. Autant de sociétés qui devraient placer l'individualisation des couvertures au second rang, à la condition, bien évidemment, qu'elles traduisent leurs engagements en actions concrètes.

La forme mutualiste n'est pas une spécificité française. Selon l'association européenne des assureurs mutualistes et de forme coopérative (AMICE), « *environ la moitié des compagnies d'assurance en Europe sont des mutuelles ou des coopératives, ou leurs filiales, et elles représentent environ 400 millions de sociétaires/assurés et plus de 469 milliards d'euros de primes d'assurance. Un tiers de toutes les assurances en Europe sont souscrites auprès de compagnies d'assurance qui suivent le modèle mutuel/coopératif.* » (AMICE, 2022). Les pays où la forme mutualiste ou coopérative est la plus développée sont les Pays-Bas (60,2% des entités d'assurance), la France (58.1%), l'Autriche (56.1%), la Slovaquie (52,2%) et le Danemark (51,1%). En revanche la forme mutualiste ou coopérative est très minoritaire en Irlande (1,4% de part de marché), Grande-Bretagne (7,4%) et Belgique (13,8%) (AMICE, 2020). Les mutuelles ou coopératives d'assurance sont également très nombreuses aux Etats-Unis et au Canada. Ces entités mutualistes ou coopératives réunissent souvent des assurés selon des critères géographiques ou professionnels. Elles partagent avec leurs équivalents français l'absence de recherche de profits, les bénéfices étant essentiellement réinvestis au profit de leurs membres, et le fait que ces membres sont à la fois sociétaires et assurés de la mutuelle ou de la coopérative. La solidarité y reste donc très forte.

3. Les difficultés liées à la collecte de données utiles à l'IA

L'IA se nourrit de données. Plus leur nombre et leur diversité seront importants, plus on peut espérer que la « mesure » du risque couvert par l'assureur sera juste. Ceci implique de pouvoir obtenir des données sur les assurés. Certaines sont évidentes (genre, âge, adresse, etc...) mais d'autres sont plus difficiles (voire interdites) à rassembler. Dans certaines hypothèses, l'assuré peut trouver avantage

²⁵ On considérera que l'inclusivité et la recherche du bien commun supposent la solidarité. Les deux sociétés concernées (CNP Assurances et Maif) font d'ailleurs figurer la solidarité dans leurs engagements.

²⁶ Crédit Agricole Assurances, Axa, CNP Assurances, Covéa, Allianz, Groupama GAN, Natixis Assurance, Assurances du Crédit Mutuel, AG2R La Mondiale, Abeille Assurance, VyV, Macif, Crédit Mutuel Arkea, MAIF, Apicil.

à les cacher. L'ouverture plus ou moins forte des assurés à la transmission de données les concernant peut également conditionner le succès d'une offre de produits d'assurance. Pour les produits s'adressant à de jeunes conducteurs, le partage de données peut être plus facilement consenti que pour une flotte de véhicules professionnels dont les conducteurs peuvent craindre que les données télématiques soient employées pour les surveiller (François et al., 2022).

Par ailleurs, de nombreuses données susceptibles de présenter un intérêt pour identifier et sélectionner les risques font l'objet d'une protection, notamment celles visées par le Règlement général sur la protection des données²⁷ qui interdit, sous conditions, le traitement de certaines données sensibles, telles que l'origine raciale, les convictions politiques, les données biométriques à fin d'identification ou les données de santé. Même si ces données sont pertinentes, elles ne peuvent pas être utilisées.

Les données génétiques sont un bon exemple de ces données fortement prédictives dont l'utilisation est pourtant prohibée. La menace, telle que décrite par Ewald (2012, p. 8) était la suivante :

« on imaginait que le nouveau savoir génétique allait détruire les mécanismes de solidarité et de mutualisation des risques. La grande menace était supposée devoir venir des assureurs. Leur intérêt pour la connaissance des risques des assurés n'allait pas manquer de provoquer un surcroît de sélection, de discrimination et d'exclusion. On craignait que l'assurance, qui jusqu'alors avait été le grand instrument de la solidarité, ne devienne l'instrument même de sa destruction ».

Mais la menace est retombée à la fois sous le poids des restrictions réglementaires à l'usage de certaines données mais également avec la relativisation de la promesse de données fortement prédictives (Ewald, 2012, p. 8). L'article L. 225-3, 1° du Code pénal prévoit qu'une discrimination se fondant « sur la prise en compte de tests génétiques prédictifs ayant pour objet une maladie qui n'est pas encore déclarée ou une prédisposition génétique à une maladie » est prohibée. Cette interdiction est reprise par l'article L. 133-1 du Code des assurances, par renvoi à l'article L. 1141-1 du Code de la santé publique. Ce qui est de l'ordre de la génétique reste donc du ressort de la solidarité puisqu'elle ne peut constituer un motif de refus d'entrée dans l'assurance, de tarification différenciée, de limitation des garanties, etc. Ce qui est craint ici c'est que sans encadrement (ou prohibition) les données génétiques soient exploitées pour écarter de l'assurance les individus que la génétique prédispose particulièrement à certaines maladies (Lehtonen et al., 2011), ou à tout le moins se voient imposer un niveau de prime qui rendrait l'assurance inaccessible. Le lien de causalité entre prédispositions génétiques et état de santé n'est pas toujours clairement établi (Lehtonen et al., 2011). De plus, le patrimoine génétique est hérité et donc indépendant de tout choix de comportement, ce qui milite, pour des

²⁷ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).

raisons morales ou sociétales, qu'il n'en soit pas tenu compte. Barry et al. (2020, p.6), citant Tselentis et al. (2017, p. 140), indiquent que nous connaissons peut-être un changement en vertu duquel « *d'autres caractéristiques inévitables telles que l'âge, le type de voiture, etc., qui ne reflètent pas nécessairement le risque d'être impliqué dans un accident* »²⁸ ne seront plus prises en compte.

Et même lorsque les données peuvent être utilisées, leur traitement doit faire l'objet de procédures rigoureuses. Comme le notent Jeanningros et al. (2020, p.5), citant le Département des services financiers de l'État de New York (2019), les régulateurs des assurances demandent que « *l'utilisation des sources de données externes, d'algorithmes ou de modèles prédictifs [ne soit] pas injustement discriminatoire. L'assureur doit établir que les sources de données externes, les algorithmes ou les modèles prédictifs sont fondés sur des principes actuariels solides, avec une explication ou une justification valable de toute corrélation ou relation causale* ».

En outre, il est tout à fait possible que l'individualisation des polices d'assurance implique une collecte de données coûteuse (et donc non rentable) pour les assureurs. Comme l'écrit Swedloff (2014, 359), cité par McFall et al. (2020 p. 4) :

« Il pourrait être extraordinairement coûteux d'exploiter des données massives et de générer des classifications des risques plus fines. Chaque assureur pourrait devoir dépenser des sommes considérables pour apporter des améliorations marginales à son système de classification des risques. Ces coûts pourraient être exacerbés par le fait que les assureurs pourraient se sentir obligés de suivre les tendances à la mode. Compte tenu de la couverture médiatique sur les merveilles du big data, les dirigeants d'entreprise peuvent dépenser des sommes exorbitantes même si le nouveau système de classification coûte plus qu'il ne rapporte »²⁹.

On notera enfin que la collecte de données trop nombreuses engendre ses propres difficultés de traitement. Or, l'accroissement du nombre de ces données est une revendication des défenseurs d'une assurance fortement individualisée, comme l'est l'insurtech Lemonade :

A une certaine granularité d'analyse « une grande partie de l'humanité se ressemble, de sorte que vous évaluez et souscrivez un grand nombre de personnes comme s'il s'agissait d'un groupe uniforme et monolithique ; mais nous avons découvert que les données que nous recueillions, ce zoom numérique de cent X dont nous disposons, nous montraient que les groupes que nos concurrents semblaient considérer comme monolithiques étaient en fait constitués de sous-groupes avec une variation de plus de 600 % dans leur probabilité de déposer une demande d'indemnisation. »³⁰

La forte segmentation offerte par certains modèles semble toutefois ne pas convaincre une majorité des assureurs.

²⁸ Traduit de l'anglais.

²⁹ Traduit de l'anglais.

³⁰ Lemonade is 5 Years Old! Here's our strategy and results to date: <https://www.youtube.com/watch?v=j7Q8SyuHWc0&t=25s>, cité par McFall et al., 2020, p. 2.

On pourra remarquer que même à l'heure du Big Data, le nombre de paramètres utilisés par les assureurs reste relativement limité : une quarantaine en assurance-automobile et moins d'une dizaine en assurance santé (Berbain et al., cités par François et al., 2018).

4. L'intérêt modéré des assureurs pour les nouveaux modèles qui permettraient une individualisation...

François et al. (2022, p.4) relèvent que pour Barry et al. (2020) « *les modèles discutés par les actuaires du champ universitaire pour tarifier les contrats automobiles restent très classiques : la recherche sur le sujet ignore très largement les outils du machine learning et continue de s'appuyer sur les modèles linéaires traditionnels de l'actuariat ; elle les enrichit, au plus, de quelques variables comportementales qui n'en améliorent que très marginalement les performances* ». Jeanningros et al. (2020, p. 2) écrivent quant à eux que « *malgré leur visibilité, les données transmises par les objets connectés ont une importance marginale dans le secteur de l'assurance maladie et de l'assurance-vie* ». Comme l'exprime un candidat-actuaire (Bachelet, 2022, p. 5) :

« Plutôt qu'une modélisation directe de la prime pure par du machine learning, nous préférons utiliser ces derniers pour créer une nouvelle variable résumant les interactions. La valorisation de cette variable se fait ensuite en la réinjectant dans les modèles linéaires généralisés existants, réduisant par construction ses erreurs. Entre autres avantages, cette approche permet de conserver l'interprétabilité des coefficients tarifaires, indispensable en tarification automobile ».

Il en va de même pour Bucci (2022, p. 3) :

« la pratique la plus utilisée pour se rapprocher de la performance des modèles de Machine Learning est l'ajout manuel de termes croisés ou d'interaction (i.e. l'âge et la zone de résidence) parmi les variables explicatives dans l'équation tarifaire : on complète cet ajout à l'aide d'un test de significativité, qui permet d'évaluer si l'on doit garder l'interaction dans le modèle »³¹.

Pour Barry et al. (2020, p.8) :

« la plupart du temps, les chercheurs recommandent d'ajouter les nouvelles variables à la classification existante (...) car [ces] variables fonctionnent mieux en combinaison avec les variables traditionnelles (...). Certains soulignent également que les variables issues des outils télématiques peuvent devenir nécessaires aux modèles existants en remplacement d'autres variables supprimées par la réglementation »³².

C'est l'approche également retenue par les assureurs en Italie - pays intéressant car leader mondial en matière d'usage de boîtiers télématiques : 4,8 millions de voitures connectées en 2016, pour 3,3 millions aux Etats-Unis et 600 000 en Grande Bretagne (Swiss Re, 2017) - même si l'usage de données comportementales paraît plus prononcé que dans le cas français. Comme l'indiquent Cevoloni et al. (2022, p. 119),

³¹ Pour Bucci (2022, p. 2), « *Plus de 80% du marché de l'assurance utilise un modèle linéaire généralisé pour prédire la prime pure à partir d'environ 20 variables explicatives* ».

³² Traduit de l'anglais.

« Les professionnels (...) que nous avons interrogés nous ont confirmé que la première étape de cette évaluation est la création de groupes d'assurés différenciés sur la base de variables actuarielles classiques : le sexe et l'âge du conducteur, le type de véhicule conduit, les antécédents du conducteur en matière de sinistres (...) et le lieu de résidence (...). Cette opération constitue la base essentielle d'une première tarification des assurés différenciés par segments selon les méthodes actuarielles classiques »³³.

Ainsi, les assureurs italiens constituent quelques classes de risques. L'usage des données télématiques servent à identifier les comportements qui sont à la source du sinistre. L'assureur accorde ensuite des récompenses ou des pénalités qui permettent d'ajuster le tarif lors du renouvellement du contrat. Comme l'écrivent Cevolini et al. (2022, p. 122),

« Si l'on entend par personnalisation une tarification entièrement basée sur les données personnelles de chaque individu, alors on peut certainement dire que les primes d'assurance numérique ne sont pas personnalisées. (...) La base de la tarification reste le modèle actuariel classique, basé sur des variables indépendantes du comportement individuel, telles que l'âge et le sexe.

Si, en revanche, nous considérons la personnalisation comme une adaptation de la prime pour tenir compte de l'exposition réelle au risque de l'individu, nous pouvons dire que les compagnies d'assurance ont commencé à expérimenter des primes "sur mesure", basées sur le suivi du comportement des assurés. Les personnes appartenant au même segment peuvent payer moins ou plus (parce qu'elles bénéficient ou non d'une réduction) en fonction de leur score réel. En d'autres termes, des personnes appartenant au même segment peuvent payer des primes différentes »³⁴.

Par ailleurs, la performance des « nouveaux modèles » qui autoriseraient l'individualisation des risques ou de la prime fait débat. Dans une étude menée en juin 2018, un candidat-actuaire (Zouggagh, 2018) a pu relever la performance des modèles suivants sur un jeu de données particulier³⁵.

Modèle	Vitesse d'apprentissage	Facilité d'explication de l'algorithme	Facilité de paramétrage	Pouvoir prédictif	Interprétabilité des résultats
GLM	+++++	+++++	+++++	+++	+++++
CART	+++++	+++++	++++	+++	++++
Random Forest	++	+	++	+++++	+
GBM	++	+	+	+++++	+
XGBoost	+	+	+	+	+

La télématique devait servir à affiner la tarification mais les résultats sont modestes. François et al. 2022, p. 22-23) écrivent :

³³ Traduit de l'anglais.

³⁴ Traduit de l'anglais.

³⁵ <http://www.galea-associes.eu/2019/06/tarification-automobile-a-laide-de-modeles-de-machine-learning-et-apport-des-donnees-telematiques/>.

« Affiner le lien causal peut être satisfaisant intellectuellement, mais le véritable enjeu est économique : il consiste à s'appuyer sur cette meilleure compréhension pour segmenter plus efficacement la population, et in fine individualiser les tarifs en conséquence. La segmentation par grandes variables sociographiques permet de découper la population en grandes catégories (relativement) homogènes ; le recours à des variables comportementales doit permettre d'accéder à une granularité beaucoup plus fine. L'individu est âgé, sans doute, mais on voit que certaines des caractéristiques de sa conduite (il freine très tard, par exemple) sont susceptibles de provoquer des accidents. En dépit de son âge, on pourra décider de lui appliquer un tarif élevé. Ces deux objectifs, d'affinement du lien causal et de segmentation accrue, n'ont été que très médiocrement atteints. Le lien que les assureurs parviennent à construire entre la mesure du risque et la sinistralité est très ténu ».

Autre difficulté pour mesurer la performance des modèles construits à partir du machine learning en assurance automobile : la faible fréquence des sinistres. Un représentant d'Axa, cité par Meyers (2018, p. 130) indique :

« Lorsque nous examinons les données historiques moyennes de nos produits d'assurance automobile, nous constatons que les clients subissent un sinistre une fois tous les vingt ans, en moyenne. Cela signifie que, pour vraiment savoir si un produit fonctionne et connaître l'impact de l'historique des sinistres, il faut en fait attendre 20 ans. Ainsi, lorsqu'on nous demande, après cinq ans, quelles sont nos expériences avec cette application... nous pourrions faire quelques déclarations sur la qualité de l'application, mais d'un point de vue actuariel et technique, ce n'est pas grand-chose. ».

Si l'on ajoute l'exigence d'explicabilité des modèles, qui s'imposera aux modèles de machine-learning, notamment en raison du règlement européen sur l'IA à venir, on voit bien que les modèles classiques comme les Modèles linéaires généralisés (« GLM») et les Modèles additifs généralisés (« GAM ») l'emportent sur les modèles d'apprentissage-machine (Bucci, 2022). Pour Fauchet (2022 p. 33), les GAM et GLM « sont toujours privilégiés (...) notamment par [leur] interprétabilité (le fonctionnement de l'algorithme doit être compréhensible par un expert, dans un format interprétable) et l'explicabilité (un néophyte doit comprendre explicitement les éléments responsables de la prise de décisions de l'algorithme) ». Bucci estime ainsi que 80% du marché de l'assurance utilise les GLM « pour prédire la prime à partir d'environ 20 variables explicatives » (2020, p. 4). L'absence de lisibilité des paramètres utilisés pour former le prix ne favorisera par ailleurs pas les démarches de prévention des assurés et plus globalement la sinistralité du portefeuille concerné. Enfin, pour Klein (2022, p. 113), « l'émergence de l'intelligence artificielle (IA) et des techniques d'apprentissage automatique (AA) comme les 'gradient boosting machines' (GBM) ou 'random forest' (...) a ouvert la voie à des gains de vitesse et de performance. Mais l'application de ces techniques classiques d'IA/AA à la tarification se heurte à des limitations majeures en raison de la nature "boîte noire" de ces algorithmes. C'est pourquoi ces types de modèle sont souvent utilisés à des fins exploratoires, mais pas en production, étant donné les risques inacceptables d'antiselection et de réglementation induits ».

Le relatif manque d'intérêt des assureurs pour l'apprentissage-machine est aussi celui des actuaires, acteurs centraux dans le fonctionnement d'une compagnie d'assurance. Le déplacement de la tâche de l'actuaire, comme individu responsable de la tarification, avec l'IA peut être décrit ainsi (Maisnier, 2020) :

« Il s'agit (...) d'un passage d'un univers où l'actuaire utilise les mathématiques pour modéliser (de manière plus ou moins complexe) la réalité à un univers où le data scientist utilise des programmes autonomes pour trouver un modèle dans les données. Par ailleurs, le fonctionnement des algorithmes de machine learning s'affranchit généralement des notions de moyenne et d'espérance sur lesquelles se fondent les approches statistiques classiques. En conséquence, les modèles ne sont plus limités par la loi des grands nombres et peuvent effectuer des analyses à des mailles beaucoup plus fines ».

On voit bien les points de « frottement » entre actuaires et *data scientists*. Pour Maisnier (2020, p. 39), « il semble nécessaire de séparer le champ d'intervention de l'IA et celui de l'actuaire : le premier doit être cantonné à un rôle de suggestion et le second doit pouvoir disposer de la compétence, de l'expertise et de la latitude nécessaire afin de confirmer ou d'invalider le résultat produit par l'algorithme ». Pour Minty (2023), l'approche classique des actuaires n'est pas encore remise en question par le *big data*. Et l'approche de l'actuaire n'est pas faite que de nombre et de formules : « quand les actuaires font un tarif, ils ne se contentent pas de chercher des corrélations, ils cherchent idéalement de possibles variables causales en se racontant des histoires. La narration du modèle, son récit, est presque aussi importante que l'analyse statistique (...). Et les histoires sont souvent écrites sur des stéréotypes » (Charpentier, 2021, p. 100).

5. ... et celui des assurés pour les produits connectés

L'assurance télématique présente *a priori* des avantages pour l'assureur mais aussi pour les assurés, et, plus largement, la société. En assurance automobile, Verbelen et al. (2018, p. 2) écrivent :

« L'assurance télématique incite fortement à modifier les habitudes de conduite et stimule une conduite plus responsable (...). Le retour d'information des utilisateurs sur leur comportement au volant et la gamification de l'assurance basée sur l'utilisation peuvent encore améliorer l'expérience du client en la rendant plus interactive, plus gratifiante et même plus excitante (...). Une conduite moins fréquente et plus sûre est encouragée, ce qui conduit à une amélioration de la sécurité routière et à une réduction des déplacements des véhicules avec moins de congestion, de pollution, de consommation de carburant, de coût de la route et d'accidents »³⁶.

De même, l'assurance connectée en matière de santé pourrait inciter les assurés à pratiquer un sport, adopter une alimentation équilibrée ou connaître les conditions d'un sommeil réparateur. Les objets connectés sur lesquels elle repose peuvent par ailleurs signaler la survenance d'un accident, comme une chute ou un problème cardiaque.

Pourtant en assurance automobile comme en assurance santé, les assurés ne semblent pas convaincus de l'intérêt de l'assurance connectée. En assurance-santé, Jeanningros et al. (2020) ont essayé de comprendre pourquoi le produit d'assurance santé Vitality, distribué dans le monde par divers assureurs, n'avait pas rencontré le succès espéré par l'assureur concepteur du produit, Discovery. Ce programme prévoyait l'attribution d'avantages (cadeaux ou prix réduits auprès de marques comme Starbucks, Apple ou Amazon) comme récompense de certaines activités sportives. En dehors du manque de conviction sur l'intérêt du *self-tracking* pour

³⁶ Traduit de l'anglais.

fixer les prix en assurance-santé, pour les auteurs il y a peu d'attachement des assurés aux marques qui peinent à se faire reconnaître (Jeanningros et al., 2020, p. 9). S'agissant des objets connectés, Jeanningros et al. (2020, p. 10) notent qu'ils « ne déterminent pas le comportement, les [assurés] jouent avec, les oublient et s'en débarrassent »³⁷. Pour Jeanningros et al. (2020, p. 12), le produit Vitality :

« a été largement interprété comme un signe que l'assurance maladie et l'assurance-vie suivaient la voie des systèmes télématiques dans l'assurance automobile, afin d'individualiser les primes en fonction du comportement suivi. Nous pensons qu'un tel résultat est beaucoup plus facile à atteindre en théorie qu'en pratique. La tarification des risques dans le domaine de l'assurance est une pratique technique obscure qui s'inscrit dans un contexte de marché hautement réglementé et concurrentiel. Même si les données d'autosurveillance devaient être modélisées en tant que facteur de tarification, il n'est pas certain que les assureurs seraient suffisamment incités à les utiliser, compte tenu de l'environnement réglementaire, des coûts d'infrastructure élevés et des coûts de réputation que cela implique. Il y a également des raisons d'être sceptique quant à la fiabilité pratique et actuarielle des données d'autosurveillance et des dispositifs et applications utilisés pour les recueillir. Les acteurs du secteur de l'assurance ne savent pas comment évaluer les risques liés à la santé des êtres humains - ils savent comment évaluer les risques liés aux agrégats, aux groupes, aux collectifs, aux pools et aux populations »³⁸.

L'intérêt du programme Vitality résidait *in fine* dans sa capacité à constituer « un moyen d'atteindre - ou de sélectionner - le type de clients que l'assureur souhaite couvrir et de les inciter à se comporter de manière à réduire les coûts de celui-ci »³⁹ (Jeanningros, 2020, p. 12-13). Ce n'est donc pas l'individualisation qui est recherchée ici mais une meilleure sélection des risques.

Cevolini et al (2022, p.123) notent que les clients d'un assureur italien ayant souscrit des assurances « télématiques » n'ont même pas téléchargé l'application permettant d'enregistrer et de transmettre les données. L'assureur a ensuite renoncé à cette offre. Pour un autre assureur italien, et toujours selon Cevolini et al., « ceux qui ont payé pour une police télématique l'ont fait "essentiellement en raison de la réduction plutôt que ... de la possibilité de comprendre leur comportement [de conduite]" »⁴⁰. Ce point est confirmé par une autre étude qui présente la remise d'un rabais sur la prime comme la première raison de souscription de l'assurance télématique en Italie (Swiss Re, 2017).

Autre illustration du relatif manque d'intérêt des assurés pour l'assurance automobile connectée, une étude menée par un assureur belge sur ce marché qui devait réunir au moins 5 000 conducteurs « connectés » pour être statistiquement significative et qui à cette fin se voyaient offrir une remise de 20% de leur prime ne réunit que... 243 volontaires. Le directeur commercial de l'assureur concerné (Meyers et al., 2020, p. 10 ; Meyers, 2018, p. 139 et seq.) tirait les conséquences

³⁷ Traduit de l'anglais

³⁸ Traduit de l'anglais.

³⁹ Traduit de l'anglais.

⁴⁰ Traduit de l'anglais.

suivantes de cet échec : « un produit d'assurance qui lie la prime au comportement au volant ne sera probablement pas proposé de sitôt »⁴¹.

Angle mort des études sur l'assurance télématique, le coût de l'installation du boîtier (ou des objets connectés en assurance santé), lorsque les données ne sont pas captées directement par le téléphone du conducteur, ou des dispositifs intégrés à toutes les voitures⁴². Une étude semble indiquer que ce coût peut incomber à l'automobiliste, pour qui « la valeur perçue des remises sur les primes et des services à valeur ajoutée, tels que la protection contre le vol, est plus élevée que les coûts perçus pour l'installation de la boîte noire et l'abonnement auxdits services »⁴³ (Swiss Re, 2017, p. 12).

Lorsque le coût du boîtier est supporté par l'assureur, la question du coût de revient est également présente même si elle concerne cette fois l'assureur. En 2016, un représentant d'AG Insurance indiquait (Meyers 2018, p. 133-134) :

« En 2007-2008, nous n'avons pas vraiment réalisé un projet test, mais plutôt une étude, une étude préliminaire. Nous ne sommes pas allés jusqu'au bout à l'époque, et la raison principale en était le coût de la technologie, qui était encore très élevé à l'époque (...). Nous avons tous installé une boîte noire (...) dans notre voiture et un dispositif qui émettait des signaux auditifs et visuels lorsque, par exemple, on prenait un virage agressif ou que l'on accélérât de manière agressive. Nous nous sommes ensuite demandé quel était l'effet sur le style de conduite (...). Nous avons conclu que le style de conduite s'était amélioré, mais nous avons décidé d'arrêter le projet parce que nous avions encore des problèmes avec la technologie de temps en temps (...) et que la question des coûts restait d'actualité. Cette technologie reste relativement coûteuse. Je sais qu'il existe aujourd'hui un nombre suffisant d'alternatives »⁴⁴.

Dans sa présentation de l'expérience menée par une société d'assurance belge présentée ci-dessus, Meyers relève que parmi les options technologiques envisageables pour capter des données de conduite les plus fiables, celle dénommée « boîte noire », qui supposait l'intervention d'un technicien automobile, était assez coûteuse (Meyers, 2018, p. 147). L'option des données transmises par téléphone, moins onéreuse, fut finalement choisie.

6. L'obstacle de l'explicabilité

La réglementation européenne à venir⁴⁵ va soumettre un certain nombre de systèmes d'IA à une obligation d'explication sur leur mode de fonctionnement et les résultats qu'ils engendrent. D'une façon générale, le projet de règlement adopté par le Parlement européen le 14 juin 2023 pose la transparence comme un des principes fondamentaux des systèmes d'IA, « transparence » qu'il définit comme le fait que ces systèmes :

⁴¹ Traduit de l'anglais.

⁴² Depuis 2018, toutes les nouvelles voitures construites pour le marché européen doivent être équipées d'un système e-Call qui communique automatiquement la localisation et l'heure d'un accident subi par le véhicule. V. Règlement (UE) 2015/758 du Parlement européen et du Conseil du 29 avril 2015 concernant les exigences en matière de réception par type pour le déploiement du système e-Call embarqué.

⁴³ Traduit de l'anglais.

⁴⁴ Traduit de l'anglais.

⁴⁵ Voir proposition de règlement européen établissant des règles harmonisées concernant l'intelligence artificielle adoptée par le Parlement européen le 14 juin 2023.

« sont développés et utilisés de manière à permettre une traçabilité et une explicabilité appropriées, tout en rendant les personnes conscientes du fait qu'elles communiquent ou interagissent avec un système d'IA, et en informant dûment les utilisateurs des capacités et des limites de ce système, ainsi qu'en informant les personnes concernées de leurs droits »⁴⁶.

S'agissant des systèmes à « haut risque », systèmes autorisés mais les plus fortement encadrés, le projet de règlement prévoit que leur « *conception et (...) développement (...) sont tels que le fonctionnement de ces systèmes est suffisamment transparent pour permettre aux fournisseurs et aux utilisateurs de comprendre raisonnablement le fonctionnement du système* »⁴⁷. Le même article définit la « transparence » comme :

« le fait que, au moment où le système d'IA est mis sur le marché, tous les moyens techniques disponibles conformément à l'état de la technique généralement reconnu sont utilisés pour garantir que les résultats du système d'IA sont interprétables par le fournisseur et l'utilisateur. L'utilisateur est en mesure de comprendre et d'utiliser correctement le système d'IA en connaissant généralement le fonctionnement du système d'IA et les données qu'il traite, ce qui lui permet d'expliquer les décisions prises par le système d'IA à la personne concernée ».

Les systèmes d'IA à risque plus faible mais destinés à interagir avec des personnes physiques doivent également être conçus et développés de manière transparente⁴⁸.

La transparence visée par le projet de règlement porte à la fois sur le système d'IA mais aussi sur les résultats – prédictions, préconisations, etc. – de ces systèmes. Elle concerne donc, avec une intensité variable, les concepteurs, les utilisateurs mais aussi ceux qui font l'objet de ces résultats.

L'obligation de transparence au profit des individus faisant l'objet d'un traitement algorithmique n'est à vrai dire, pas nouvelle. L'article 13 du Règlement général sur la protection des données indique en effet que dans l'hypothèse de « *l'existence d'une prise de décision automatisée, y compris un profilage, (...) des informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée [devront être fournies]* ». Par ailleurs, l'article 4 de la loi du 7 octobre 2016⁴⁹ dispose que toute personne concernée par un traitement algorithmique peut exiger que l'administration concernée produise « *les règles définissant ce traitement ainsi que les principales caractéristiques de sa mise en œuvre* ».

Or, comme le notent François et al. (2022), les données générées par les boîtiers utilisés en assurance automobile prennent la forme de scores « synthétiques » ou « agrégés », le plus souvent calculés par les fournisseurs de boîtiers qui se prêtent difficilement à toute forme d'interprétation :

« Le plus souvent (...) ce ne sont pas les données brutes enregistrées par le dispositif (vitesse d'accélération et décélération, brutalité du freinage ou des virages, etc.) qui sont utilisées directement. Ces données brutes sont utilisées par un

⁴⁶ Art. 4 bis du projet de règlement adopté par le Parlement européen le 14 juin 2023.

⁴⁷ Art. 13 du projet de règlement adopté par le Parlement européen le 14 juin 2023.

⁴⁸ Art. 52 du projet de règlement adopté par le Parlement européen le 14 juin 2023.

⁴⁹ Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

prestataire pour calculer un score synthétique – et c’est ce score synthétique qui est mobilisé par les assureurs pour tenter d’améliorer le calcul du tarif » (p.13).

« Les scores qui sont intégrés dans les modèles de tarification, le plus souvent, synthétisent les informations captées par les dispositifs télématiques. Ils sont calculés par des prestataires extérieurs suivant des procédures souvent opaques, parfois tâtonnantes, et qui en général échappent complètement aux assureurs » (p. 23).

L’opacité des modèles utilisés par les assureurs pour individualiser les couvertures d’assurance est critiquée par O’Neil (2016, p. 164), citée par Meyers, 2018 (p.15) :

« Les assureurs utilisent déjà les données pour nous diviser en petites tribus, pour nous offrir des produits et des services différents à des prix variables. Certains appelleront cela des services personnalisés. Le problème, c’est qu’il ne s’agit pas de services individuels. Les modèles nous placent dans des groupes que nous ne pouvons pas voir, mais dont le comportement paraît ressembler au nôtre. Quelle que soit la qualité de l’analyse, son opacité peut conduire à l’escroquerie »⁵⁰.

Pour Hay (2015, p. 26-27), les scores obtenus par les modèles d’apprentissage machine devront être explicables (à l’assuré ou au régulateur), ce qui paradoxalement fera ré-apparaître le besoin d’un traitement statistique et une démonstration de leur « solidité » qui restera « fonction du nombre d’assurés sous-jacent ». Ainsi, pour Hay, « le big data est gourmand en mutualisation ».

Enfin, ces scores et la tarification qu’ils engendrent peuvent être source de pertes pour l’assureur (le score est « bon » mais l’assuré connaît des sinistres) ou d’incompréhension par l’assuré (le score est « mauvais » mais l’assuré ne connaît aucun sinistre) (François et al., 2022, p. 24).

7. L’inertie des assureurs

Comme indiqué ci-dessus, François et al. (2022) sont partis de l’étude de quatre produits d’assurance automobile pour essayer de comprendre pourquoi la révolution de l’IA sur l’assurance n’a pas eu lieu.

La première raison identifiée par François et al. (2022, p. 4) touche à l’« inertie organisationnelle » des assureurs. Deux mécanismes sont ici à l’œuvre. Le premier « renvoie à l’ensemble des routines organisationnelles qui sont remises en cause par un processus disruptif » (François et al., 2022, p.5). Très concrètement, les personnes supposées mettre en place le mécanisme ne le font pas. Or, en assurance « les routines organisationnelles (...) sont très puissantes » (François et al., 2022, p. 6).

Le second mécanisme tient à la place de certains groupes professionnels qui résistent à la mise en place d’innovations car elles remettent en cause leurs compétences techniques, en les « disqualifiant » (François et al., 2022, p. 6). C’est le cas des actuaires, qui ont une place prépondérante dans les organismes d’assurance et dont les techniques pourraient être rendues obsolètes par le *big data*. En outre, les infrastructures héritées et les pratiques institutionnalisées et enracinées constituent un frein au changement autant qu’une réelle opportunité dans un

⁵⁰ Traduit de l’anglais.

secteur « *mûr pour la rupture* » (McFall et al., 2020, p. 2). L'une des sources de frustration les plus importantes pour les *data scientists* tient à la lourdeur des processus de tarification. Concrètement, « *l'innovation tarifaire est souvent bridée par des contraintes IT* », même lorsqu'il s'agit simplement d'ajouter une variable dans le modèle tarifaire (Fauchet, 2022, p. 4). McFall (2015, p.32) parle des « *défis d'interopérabilité des infrastructures existantes au sein des compagnies d'assurance traditionnelles* ». D'autres auteurs encore insistent sur l'obsolescence des structures informatiques de ces assureurs qui ne leur permettent pas de tirer profit des nouvelles technologies (Giraud, 2022).

8. Des assureurs peu proactifs

François et al. (2022, p. 7) considèrent par ailleurs que « *l'échec des télématiques ne serait pas le symptôme d'une impuissance à adopter une innovation, mais le résultat d'un calcul et d'une stratégie délibérée* ». L'adoption de ces télématiques s'est faite dans le cadre d'« *expérimentations* ». C'est l'ensemble de la chaîne de production du produit d'assurance qui est testée. Mais les résultats de ces tests ne conditionnent pas à eux seuls la décision d'adoption du produit : « *ce qui est déterminant c'est ce que font les concurrents* » (François et al., 2022, p.8). Les initiatives des assureurs dépendent donc des initiatives de leurs concurrents. Ils se préparent aussi en cas de lancement par un concurrent d'un produit. François et al. écrivent (2022, p. 8): « *c'est un marché sur lequel les producteurs s'observent les uns les autres, et où leur comportement dépend de ce qu'ils observent du comportement des autres (...). Les expériences s'inscrivent dans un processus que DiMaggio et Powell (1983) décriraient comme un isomorphisme mimétique : leur marché fonctionne comme une arène où tous les producteurs s'observent et où personne ne bouge car personne ne voit personne bouger* » (p. 8). Dans un secteur où peu d'initiatives structurantes émergent, il n'est donc pas surprenant que les assureurs ne se soient pas encore engagés dans le « virage » de l'individualisation.

Par ailleurs, l'individualisation des produits d'assurance porte en elle un changement auquel certains assureurs peuvent être réticents : le changement de l'activité d'assureur de porteur de risques à celui de conseiller des assurés sur les éléments pouvant être optimisés pour faire baisser leur prime ou leur sinistralité. Comme l'écrit Meyers et al. (2020, p. 12),

« Avec l'application d'étude du style de conduite, les clients des compagnies d'assurance ont été activement encouragés à modifier leur comportement. L'application de style de conduite testée ici est dotée d'infrastructures de retour d'information personnalisées pour responsabiliser les conducteurs quant à leurs "choix" de conduite (...). De plus, avec ces pratiques de coaching, on pourrait assister à un changement du rôle de l'assurance, qui passerait de fournisseur de produits à fournisseur de services, où les assureurs agiraient comme des coaches personnels qui "aideraient les individus à s'aider eux-mêmes" (Tritch, 2007) »⁵¹.

Or, la pertinence économique pour les assureurs de ce changement de rôle reste à démontrer. Quelle est la légitimité d'un assureur pour s'imposer comme « coach » de vie des assurés ? Et selon quel modèle économique ? Les diverses incursions

⁵¹ Traduit de l'anglais.

des assureurs dans le monde des services (assistance, réseaux de soins, etc.) ne semblent pas avoir convaincu.

9. Une réglementation sur l'utilisation de certaines données personnelles qui laisse relativement peu de place à l'individualisation de la tarification

L'individualisation des couvertures d'assurance suppose la collecte de données pertinentes pour sélectionner les risques. Or, en matière de santé, la collecte et l'usage de ces données sont fortement encadrés. En matière d'assurance collective⁵², si le recueil de données de santé ou de données comportementales n'est pas interdit, l'exclusion d'un assuré n'est pas possible. Comme l'indique la Cour de cassation⁵³, citée par Del Sol (2020), « *le principe de non sélection individuelle des risques (...) en matière de prévoyance collective obligatoire prohibe une exclusion qui ne concerne pas la totalité du groupe des personnes assurées* ». De même, aucune individualisation ne peut être pratiquée en matière de tarification. En ce qui concerne l'assurance individuelle, l'art. L. 110-2 du Code de la mutualité dispose que « *les mutuelles (...) ne peuvent en aucun cas recueillir des informations médicales auprès de leurs membres ou des personnes souhaitant bénéficier d'une couverture, ni fixer les cotisations en fonction de l'état de santé* ». Les assureurs qui relèvent du Code des assurances ne sont pas tenus par les mêmes obligations et peuvent de fait pratiquer une tarification en fonction de l'état de santé et donc intégrer cette variable dans leurs modèles. Néanmoins s'ils souhaitent proposer à leurs assurés des contrats « responsables et solidaires », qui bénéficient d'une fiscalité avantageuse, ils doivent pour ces mêmes contrats renoncer à la sélection et à la tarification en fonction de l'état de santé. Or, on estime que 95% des contrats d'assurance santé sont « solidaires » (Del Sol, 2020).

En assurance automobile, on peut noter que l'assureur n'a pas accès aux données sur les infractions et retraits de points du fichier national des permis de conduire, données pourtant probablement fortement prédictives (Gollier, 2015).

10. Une prime qui ne reflète pas exclusivement le profil de risque d'un individu

La prime versée par un assuré ne reflète pas uniquement son « risque ». Des « chargements » sont également intégrés dans le « prix » payé par l'assuré pour couvrir notamment les coûts administratifs et de commercialisation des assureurs et abonder des provisions « techniques » qui répondent à des exigences réglementaires (Zajdenweber, 2006). A ces chargements s'ajoute également une fiscalité significative. Comme l'écrivent McFall et al. (2020, p. 4), « *les perfectionnements techniques en matière de calcul des coûts ne signifient pas nécessairement que les prix refléteront plus précisément le risque, car les décisions en matière de prix impliquent toujours*

⁵² Le contrat collectif est « *un contrat souscrit par une personne morale ou un chef d'entreprise en vue de l'adhésion d'un ensemble de personnes répondant à des conditions définies au contrat* » (v. art. L. 141-1 du Code des assurances). Le cas le plus usuel est celui des salariés d'une entreprise qui souscrit un unique contrat auprès d'une compagnie d'assurance au profit de ces employés.

⁵³ Cass. 2^{ème} civ., 3 février 2011 (n°10-30588)

des considérations d'ordre commercial et de marketing »⁵⁴. Noumedem Djieuzem note (2022, p. 2) qu' :

« en pratique, au montant de prime pure, il faut ajouter des chargements (positif ou négatif), désignés par optimisation layer (OL) pour obtenir la prime commerciale. En appliquant des profils de marge différents sur certains périmètres du portefeuille, l'assureur crée de la solidarité entre ses assurés. Sur des risques différents, l'assureur peut décider de jouer sur l'OL pour faire payer la même prime commerciale à tous les assurés aboutissant ainsi à un portefeuille solidaire. De même, l'assureur peut utiliser l'OL pour financer l'acquisition et la rétention des contrats, notamment les remises à la souscription ».

Si les chargements permettent une re-mutualisation d'assurés à prime pure différente, ils sont parfois utilisés à des fins commerciales qui attirent de plus en plus l'attention des régulateurs. Il en va ainsi de l'utilisation de ces chargements pour financer l'acquisition ou la rétention d'assurés. Le régulateur anglais, la *Financial Conduct Authority* (FCA), s'est ainsi intéressé aux pratiques dites de « *price walking* » en assurance habitation et en assurance automobile. Elles consistent à identifier les assurés les plus susceptibles de renouveler leur contrat pour leur appliquer des primes surélevées et utiliser ce surplus de prime pour accorder des rabais à de nouveaux assurés. Ainsi, les assurés qui résilient régulièrement leur contrat et font jouer la concurrence entre assureurs paient moins que les assurés fidèles, pour un risque équivalent. La FCA observe que cette pratique a été préjudiciable à plus de 6 millions d'assurés et aurait engendré un surplus de cotisations par ces assurés de 1,2 milliard de livres sterling (FCA, 2021, p. 3). A titre d'illustration les nouveaux assurés payaient en moyenne 285 livres sterling contre 370 livres sterling pour les assurés de plus de 5 ans (FCA, 2020, p. 6). La FCA a donc décidé d'interdire cette pratique, les assureurs devant offrir à leurs assurés un niveau de prime qui ne peut être supérieur à celui des nouveaux assurés. Cette mesure est susceptible de générer une économie pour les assurés de 4,2 milliards de livres sterling sur les 10 années à venir (FCA, 2021). On pourrait considérer que la prise en compte de la probabilité de résiliation d'un assuré est une forme d'individualisation des chargements et que son interdiction conduit à une plus grande uniformisation des primes.

D'une façon plus générale, certaines données ne peuvent légalement pas être collectées par un assureur, ce qui rend l'individualisation imparfaite.

11. Quelques considérations morales ou sociétales

François et al. (2022, p. 3), évoquent les travaux de Zelizer (1979) sur la dimension morale et politique de l'activité assurantielle, pour relever des « *préventions morales et politiques* » qui empêcheraient une individualisation des garanties, parfois également par opportunisme (Bénéplanc et al., 2022). Gollier (2015) se demande si une segmentation très forte des garanties n'entraînerait pas des différenciations de tarification d'autant plus insupportables socialement si on découvrait qu'elles sont corrélées à la richesse des individus. En France le régulateur mais aussi l'assureur,

⁵⁴ Traduit de l'anglais.

lorsqu'il est mutualiste, pourraient ne pas accepter qu'il soit touché à la mutualisation, en danger d'être « tuée » (Dixneuf, 2019). D'une façon générale, la crainte de l'exclusion, partielle ou totale, de l'assurance de certains individus est perçue comme inacceptable.

Meyers (2018, p. 117-118) rapporte qu'en novembre 2015 un assureur belge avait distribué 500 montres connectées à certains de ses clients. Ce don a été interprété comme le test d'une tarification fondée sur les habitudes de vie. Le ministre des finances belge a aussitôt réagi en déclarant :

« La fixation des primes est une liberté contractuelle entre l'assureur et l'assuré. Cependant, dès que j'observe que le mécanisme de solidarité est en train de se vider, j'interviens. Les primes et les tarifs doivent être abordables pour chaque citoyen. Du point de vue de la protection de la vie privée, le consommateur doit toujours avoir la possibilité de consentir explicitement au partage de ses données. »⁵⁵.

Il existe par ailleurs un doute sur l'usage qui est fait des données collectées, certains auteurs, comme Zuboff (2019), cité par McFall et al. (2020, p. 3), considérant que « les assureurs utilisent des dispositifs connectés, télématiques et d'autosurveillance pour "modifier les comportements dans le sens d'une rentabilité maximale" »⁵⁶.

De plus, les assureurs hésitent à se lancer dans une pratique vue comme pouvant engendrer des discriminations sur la base de variables plus difficiles à comprendre (McFall, 2020). Certaines de ces variables sont des combinaisons d'autres variables qui n'ont pas de sens « pratique » (par exemple la combinaison âge, genre et domicile).

En outre, se pose la question de la prise en compte d'éléments non choisis par l'assuré (capital génétique, genre, etc...), par opposition à ceux qui résultent de comportements délibérés. Est-il légitime de prendre ces premiers en considération pour tarifier une garantie d'assurance (Steiner, 2018) ? Comme l'indiquait en 2015 le responsable de l'innovation de Swiss Re (Meyers, 2018, p. 124) :

« Par exemple, nous n'avons pas le droit d'utiliser des tests génétiques parce que vous êtes né comme ça. Vous n'avez pas votre mot à dire. Peu importe ce que vous mangez, vous n'avez aucune influence sur votre génétique. À long terme, le régulateur et les groupes d'intérêt des consommateurs mettront probablement un terme à l'utilisation de choses sur lesquelles les gens n'ont aucun contrôle. La seule chose sur laquelle les gens ont un contrôle clair, c'est leur comportement »⁵⁷.

La distinction entre variable admissible et variable interdite tiendrait donc à la différence entre les éléments sur lesquels l'assuré a un contrôle et ceux qui ne sont pas sous son contrôle. La distinction n'est pas toujours évidente même dans des cas qui semblent simples comme le tabagisme, dont l'accoutumance peut en partie s'expliquer par des facteurs psycho-sociaux indépendants de l'individu.

⁵⁵ Traduit de l'anglais.

⁵⁶ Traduit de l'anglais.

⁵⁷ Traduit de l'anglais.

Finalement, le simple risque de réputation fait hésiter certains assureurs. Barry et al. (2020, p. 9) prennent pour exemple le cas d'un conducteur considéré comme « agressif » alors qu'il n'est à l'origine d'aucun accident.

L'individualisation des couvertures d'assurance suppose donc de pouvoir préalablement convaincre les citoyens et les régulateurs de l'équité d'une telle pratique et de sa compatibilité avec les principes qu'ils ont collectivement choisis.

Conclusion

En assurance santé comme en assurance automobile, l'IA n'a à ce jour pas eu l'effet « disruptif » annoncé. L'assurance en 2023, supposément « boostée » par le Big Data et les techniques d'IA les plus récentes, comme le *deep learning*, ressemble beaucoup à celle du début du XXIème, quand ces techniques étaient encore confidentielles. La raison principale de l'écart entre les promesses de l'IA en assurance et la réalité de l'activité assurantielle tient notamment au caractère intrinsèquement collectif de l'assurance, caractéristique diamétralement opposée à l'approche individuelle des risques promise par l'IA. D'autres raisons, liées en particulier à l'inertie des acteurs du secteur, à leur culture ou aux pratiques de ses spécialistes de la donnée, mais également au caractère obscur d'une l'IA souvent perçue comme une insondable « boîte noire », expliquent également cet écart. Ces raisons sont-elles suffisantes pour conclure à un échec de l'IA en assurance ? Rien n'est moins sûr. La recherche par les assurés de produits et services adaptés à leurs besoins, en assurance comme dans bien d'autres domaines, pourrait profondément changer les pratiques du secteur assurantiel.

Dans leur ouvrage à succès mêlant fiction et prospective, Lee et Qiufan (2021) imaginent un monde dans lequel l'IA permettrait d'ajuster en temps réel la couverture et la prime d'un assuré, en fonction par exemple de son activité ou de ses choix de vie. A un individu dont l'IA déterminerait qu'il est susceptible de développer un diabète, l'assureur proposerait une couverture dont la prime dépend de l'activité physique ou des habitudes alimentaires. Un autre verrait sa prime d'assurance santé ou prévoyance évoluer en fonction de la notation sociale des individus qu'il fréquente ou de la dangerosité des quartiers dans lesquels il se déplace. A chaque décision ou action entreprise par un assuré et partagée avec l'assureur, par le biais d'objets connectés toujours plus sophistiqués, correspondraient une couverture et une prime intégralement individualisées. Bien sûr, un monde tel que celui que dessinent Lee et Qiufan ne nous conviendra peut-être pas, pour des raisons éthiques ou sociétales par exemple. Il incombera alors au législateur ou au régulateur d'encadrer les usages de l'IA dans le secteur assurantiel et au-delà. L'IA nous démontre (malheureusement ?) tous les jours que la technologie avance plus vite que le cadre juridique qui devrait lui être applicable. L'insertion en toute dernière minute des systèmes d'IA à finalité générale – ou IA générative – dans le projet de règlement européen sur l'IA adopté par le Parlement le 14 juin 2023 en est la triste illustration.

Bibliographie

AMICE, <https://amice-eu.org> (consulté le 22 août 2023).

Abraham K., « Efficiency and fairness in insurance risk classification », *Virginia Law Review*, 403, 1985.

Bachelet C., « Création d'une variable tarifaire d'interactions en assurance automobile par machine learning », mémoire présenté pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA et l'admission à l'Institut des Actuaire, 2022.

Barry L., « Justice ou justesse ? L'équité de l'assurance », Chaire Pari working paper, avril 2019.

Barry L., Charpentier A., « Personalization as a promise : Can Big Data change the practice of insurance ? » », *Big Data & Society*, Janvier-juin 2020, p. 1-12.

Barry L., « Les scores ont-ils un sens ? », Chaire Pari working paper, décembre 2020.

Bénéplanc G, Charpentier A., Thourot P., *Manuel d'assurance*, PUF, 2022.

Benoît C., Devaux F., Tuffy Y., « Big Data, régulation financière et trajectoires organisationnelles dans l'industrie de l'assurance », working paper, Pari, 2021.

Berliner B., « Large Risk and Limits of Insurability », *The Geneva Papers on Risk and Insurance*, 10, N°37, octobre 1985.

Bigot J. (sous la direction de), *Entreprises et organismes d'assurance*, Tome 1, 3^{ème} édition, 2011.

Bucci S., « Modèles de tarification traditionnels et Machine Learning: vers une nouvelle collaboration en assurance ? », *Variations*, 4 avril 2022.

Cevolini E., Esposito C., « From pool to profile: Social consequences of prediction in insurance », *Big Data & Society*, juillet-décembre 2020.

Cevolini E., Esposito E., «From Actuarial to Behavioural Valuation. The impact of telematics on motor insurance», *Valuation Studies*, 9(1) 2022, p. 109-139.

Charpentier A., Denuit M., Elie R., « Segmentation et mutualisation : les deux faces d'une même pièce ? » : *Risques*, 103, novembre 2015.

Charpentier A., Barry L., Gallic E., « Quel avenir pour les probabilités prédictives en assurance ? », *Annales des Mines – Réalités industrielles*, 2020/1, p. 74-77.

Charpentier A., « Assurance et discrimination : quel rôle pour les actuaire ? », *Risques*, n°127, p. 110, 2021.

Charpentier A., *Assurance : Discrimination, Biais & Equité*, Institut Louis Bachelier, 2022.

Charpentier A., « Statistique bayésienne, data sciences et nouveaux risques », <https://freakonometrics.hypotheses.org/64009>, 2022.

Charpentier A., « Autocalibration », <https://freakonometrics.hypotheses.org/6604>, 2023.

Davet J.-L., « L'assurance santé individuelle : des solidarités à l'épreuve de la segmentation », *Risques*, n°87, 2011, p. 43-50.

Del Sol M., « La réglementation française de l'assurance santé à l'épreuve des objets connectés et des pratiques de profilage », in *Le profilage en ligne : entre libéralisme et régulation*, ouvrage collectif sous la direction de Turgis S., Boizard M. et Benasamoun A., Mare&Martin, 2020.

DiMaggio P.J., Powell W. « The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields », *American Sociological Review*, 1983, 48(2) p. 147–160.

Dixneuf P., « Intelligence artificielle et assurance », *Annales des Mines*, mars 2018, n°1, p. 31 et suiv.

DREES, « Les dépenses de santé en 2021 – édition 2022 – Résultats des comptes de la santé », 15 septembre 2022.

Ernst&Young, *Introducing Pay As You Live (PAYL) Insurance, Insurance that rewards a healthier lifestyle*, août 2015 (<https://fr.scribd.com/document/472549816/EY-introducing-pay-as-you-live-payl-insurance-Copy-docx>).

Ewald F., *L'Etat providence*, Grasset, 1986.

Ewald F., « Assurance, prévention, prédiction... dans l'univers du Big Data », *Rapport pour l'Institut Montparnasse*, 2012.

Ewald F., Thourot P., « Big Data : défis et opportunités pour les assureurs », *Banque & Stratégie*, juin 2013.

Fauchet G., « Le processus de tarification en assurance non-vie : panorama des pratiques du marché », *White Paper Milliman*, novembre 2022.

Feller W., *An Introduction to Probability Theory and Its Applications*, John Wiley, 3ème édition, vol. I, chapitre IX, p. 231, 1968.

Financial Conduct Authority (FCA), « General insurance pricing practices market study Feedback to CP20/19 and final rules Policy Statement », *Policy Statement PS21/5*, mai 2021.

Financial Conduct Authority (FCA), « FCA confirms measures to protect customers from the loyalty penalty in home and motor insurance markets », communiqué de presse 28 mai 2021.

Financial Conduct Authority (FCA), « General insurance pricing practices », Market study, MS18/1.3 septembre 2020.

François P., Voldoire T., « La révolution qui n'a pas eu lieu, télématiques et assurance automobile dans les années 2010 », *working paper*, Chaire Paris, avril 2022.

François P., Barry L., « Les enjeux du Big Data pour l'assurance », *working paper*, Chaire Pari, novembre 2018.

Frezal S., *Quand les statistiques minent la finance et l'économie*, L'Harmattan, 2018.

Giraud F., « L'avenir de l'assurance passe par les écosystèmes », *Risques*, n°130, juin 2022, p. 129-134.

Gollier C., « Dans quel sens la révolution numérique affecte-t-elle l'assurabilité des risques », *Revue d'économie financière*, 2015/2, p. 197-204.

Hay, F.-X., « La mutualisation est-elle soluble dans le *big data* ? », *Risques* n°103, p. 25-30), 2015.

Jawish J., « Société d'assurance : quel niveau de digitalisation adopter pour garder ses clients », *Risques*, n°130, juin 2022, p. 106-110.

Jeanningros H., McFall L., « The value of sharing : Branding and behaviour in a life and health insurance company », *Big Data & Society*, Juillet-Décembre 2022, 1-14.

Klein A.-L., « Libérer le potentiel de la tarification grâce à l'intelligence artificielle transparente », *Risques*, n°130, juin 2022, p. 111-115.

Landes X., « How Fair is Actuarial Fairness ? », *Journal of Business Ethics*, 2015.

Lee K.-F., Qiufan C., *AI 2041 – Ten visions for our future*, Currency, 2021.

Lehtonen T.-K., Liukko J., « The forms and limits of insurance solidarity », *Journal of Business Ethics*, 2011.

Madeline B., « La tech en quête d'usages », *L'Argus de l'assurance*, n° 7588, p. 30-35, 2019.

Maisnier N., Lhenri F., « Cadre éthique de l'utilisation des techniques de data science en actuariat », mémoire présenté devant l'Institut du Risk Management pour validation du cursus à la Formation d'Actuaire de l'Institut du Risk Management et l'admission à l'Institut des actuaires, 2020.

Mayaux L., « Aspects juridiques de l'assurabilité », *Risques*, n°54, avril-juin 2003.

Mayaux L., *Répertoire de droit civil, Assurance*, Dalloz, janvier 2015, mise à jour en septembre 2022.

McFall L., « A 'Good, Average Man' : Calculation and the Limits of Statistics in Enrolling Insurance Customers », *Sociological Review*, 59(4), p. 661-68, novembre 2011.

McFall, L., « Is digital disruption the end of health insurance? Some thoughts on the devising of risk », *economic sociology_the european electronic newsletter*, ISSN 1871-335, Max Planck Institute for the Study of Societies (MPIfG), Vol. 17, Iss. 1, 2015, p. 32-44.

McFall L., Meyers G., Van Hoyweghen I., « Editorial: The personalization of insurance: Data, behaviour and innovation », *Big Data & Society*, Juillet-Décembre 2020, 1-11.

McFall L., « Personalizing solidarity? The role of self-tracking in health insurance pricing », *Economy and Society*, volume 48, février 2019.

Meyers G., Van Hoyweghen, « 'Happy failures': Experimentation with behavior-based personalization in car insurance », *Big Data & Society*, Janvier-juin 2020, 1-14.

Meyers G., *Behaviour-based Personalisation in Health Insurance: a Sociology of a not-yet Market*, 2018, Thèse, KU Leuven.

Meyers G., Van Hoyweghen I. « 'Happy failures : Experimentation with behavior-based personalisation in car insurance, *Big Data & Society*, Janvier-Juin 2020.

Minty D., « Has Telematics Really Changed Insurance? Lessons from Italy », <https://www.ethicsandinsurance.info/has-telematics-really-changed-insurance-lessons-from-italy/>, 3 mai 2023.

Minty D., «Lessons in Digital Innovation from the French Motor Market», <https://www.ethicsandinsurance.info/lessons-in-digital-innovation-from-the-french-motor-market/>, 23 mai 2023.

Noumedem Djieuzem B. D., « De la communauté à l'individu : Quantifier pour comprendre l'impact des pratiques actuarielles sur la mutualisation », mémoire présenté devant l'ENSAE Paris pour l'obtention du diplôme de la filière actuariat et l'admission à l'Institut des Actuaire, 16 mars 2022

Siegel E., *Predictive Analytics : The Power to Predict Who Will Click, Buy, Lie, or Die*, 2^{ème} édition, 2016, Wiley, Hoboken, New Jersey.

Steiner R., « Big Data, mutualisation et exclusion en assurance », *Enjeux numériques*, n°2, juin 2018.

Swedloff R., « Risk Classification Big Data (R)Evolution », *Connecticut Insurance Law Journal*, Vol. 21, No. 1, Automne 2014.

Swiss Re, « Unveiling the Full Potential of Telematics. How connected Insurance Brings Value to Insurers and Consumers : An Italian Case Study », 2017.

Tritch T, « Helping people help themselves, *The New York Times*, 2007.

Tselentis D., Yannick G. et Vlahogianni E., « Innovate insurance schemes: A review of current practices and emerging challenge », *Accident, Analysis and prevention*, 98, p. 139-148.

Verbelen R., Antonio K., Claeskens G., « Unraveling the predictive power of telematics data in car insurance pricing », *Journal of the Royal Statistical Society*, 2018, vol. 67, part 5, p. 1275-1304.

Von Mises, R., *Probability, statistics and truth*, Mcmillan, 1939.

Zajdenweber D., *Economie et Gestion de l'Assurance*, Economica, 2006.

Zajdenweber D., « Quand la sélection augmente le risque », *Risques* n°103, 2015, 54 et suiv.

Zuboff S., *The Age of Surveillance Capitalism*, London, Profile Books, Kindle edition.

Zouggagh F.Z., « Tarification automobile à l'aide de modèles de machine learning et apport des données télématiques », Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du Diplôme d'Actuaire EURIA et de l'admission à l'Institut des Actuaires, 2018, <https://www.institutdesactuaires.com/docs/mem/206a62192470ef6436d38fbcfa408035.pdf>.

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 Addactis, 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

