IRM | institut du risk management

Rapport de projet présenté devant un Jury de Soutenance

Expert ERM

Expert(e) Management des Risques Financiers et Assurantiels

Le 18 novembre 2021

Par : Rémi GERBOUD et Emna FOURATI

Titre : Mise en place d'une gouvernance ERM pour le déploiement de modèles d'*intelligence artificielle* pour une compagnie de réassurance vie.

Confidentialité : ☒ NON ☐ OUI (Durée : ☐ 1an ☐ 2 ans)

La durée de confidentialité expire aux 31 décembre N+1 (1 an) ou N+2 (2 ans)

Les stagiaires s'engagent à ce que les données de l'Entreprise présentées dans le cadre des travaux de la formation (rapport de projet & présentation) respectent les règles relatives à la protection des données à caractère personnel conformément aux dispositions de la Loi informatiques et Liberté n°78-17 du 6 janvier 1978 modifiée par la Loi du 6 août 2004 ainsi que par la loi n° 2018-493 du 20 juin 2018 (RGPD)

Membres présents du jury :

_____

_____

_____

_

_____

_____

Par ma signature j'autorise la publication sur un site de diffusion de documents actuariels du rapport de projet
(après expiration de l'éventuel délai de confidentialité)

Nom : GERBOUD

Prénom : Rémi

Signature du stagiaire

Si binôme :

Nom : FOURATI

Prénom : Emna

Signature du stagiaire

TABLE DES MATIERES

1. INTRODUCTION

Dans le contexte d'un monde de plus en plus digitalisé, les (ré)assureurs ont développé des compétences pour bâtir de nouveaux modèles afin d'être en mesure de proposer des offres pertinentes et des services plus innovants et mieux adaptés à leurs clients.

Plus globalement, les (ré)assureurs ont vocation à avoir un impact positif sur la société. Pour cela, ils peuvent inciter leurs assurés à avoir une meilleure qualité de vie par la prévention, ou encore proposer un accès à l'assurance à davantage de profils de client qui n'y ont pas encore accès. Il s'agit de réduire le fossé des inégalités face à l'accès à l'assurance « protection gap ».

Les entreprises de (ré)assurance ont une responsabilité sociale à assumer à travers la protection sociale. Elle représente la raison fondamentale de leur existence et leur permet de poursuivre un développement durable sur le long-terme. Cet objectif est d'autant plus difficile dans un cadre concurrentiel fort couplé avec des objectifs de profitabilité et de solvabilité.

Les nouvelles technologies et les modèles de type Intelligence Artificielle permettent d'accompagner les enjeux stratégiques des entreprises de (ré)assurance dans l'atteinte de leurs objectifs. Ces modèles peuvent être rapidement déployés dans de nouveaux marchés, notamment émergents. Ils permettent également d'attirer de nouveaux profils de clients et d'augmenter des parts de marché dans des marchés matures.

Cependant, ces algorithmes peuvent dévier de leur objectif initial et conduire à l'émergence de nouveaux risques. Il est donc nécessaire d'encadrer ces transformations digitales à fort potentiel dans les entreprises de (ré)assurance en adoptant une démarche ERM robuste. La gestion et la maîtrise des risques doit-être au centre de ces transformations. La diffusion d'une culture du risque à tous les niveaux de l'entreprise doit-être pensée dès le stade de conception de ces projets.

2. CONTEXTE ET PROBLEMATIQUE : L'INTELLIGENCE ARTIFICIELLE AU CENTRE DES TRANSFORMATIONS ECONOMIQUES

2.1 PRESENTATION DE L'ENTREPRISE REA

2.1.1 CADRE DE L'ETUDE : REA DANS L'ERE DU « TOUT DIGITAL »

Dans le cadre de ce projet, nous considérons une entreprise de réassurance vie fictive, appelée par la suite REA.

Nous proposons d'élaborer une gouvernance reposant sur une démarche ERM afin de permettre au réassureur REA de saisir les opportunités offertes par les modèles d'intelligence artificielle (IA) dans le cadre de sa stratégie de développement tout en s'assurant du contrôle et de la maîtrise de leur finalité.

2.1.2 LA DONNEE EST LA MONNAIE DE DEMAIN[1]

Au regard des évolutions technologiques considérables et des progrès majeurs en termes de collecte et de traitement des données, la société REA développe les outils, les processus, et les analyses pour affiner la compréhension dynamique des leviers de création de valeur.

REA souhaite évoluer dans le domaine des nouvelles technologies avec un cadre ERM bien défini et robuste, ainsi qu'une culture du *risk management* bien développée et diffuse au sein de l'entreprise.

2.1.3 FRONTIERE EFFICIENTE DE L'INDUSTRIE DE L'ASSURANCE

[1] Voir l'illustration en Annexe A

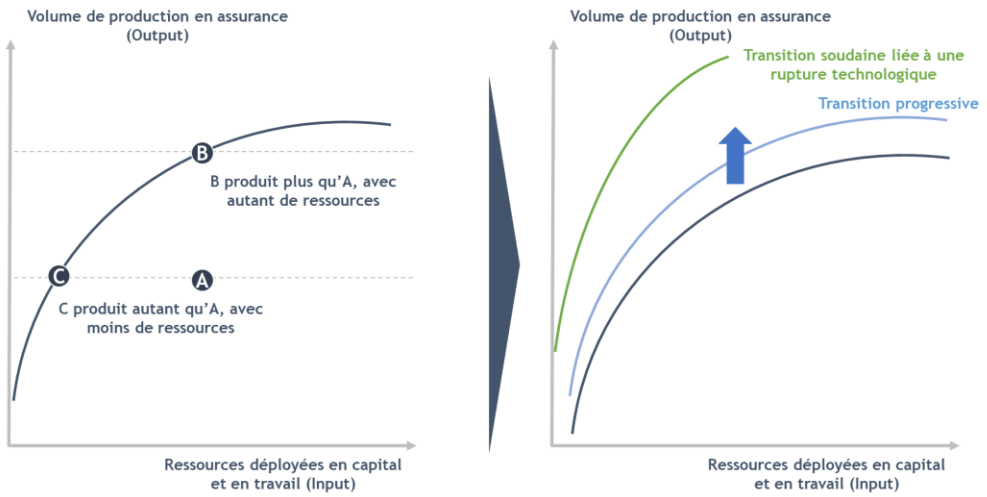
Les développements technologiques entraînent un décalage positif de la frontière d'efficience de l'industrie, permettant de produire davantage de (ré)assurance avec les mêmes ressources en capital et en travail.

Au fur et à mesure de l'intégration de ces développements technologiques dans l'industrie, on constate, progressivement, une production de la (ré)assurance de plus en plus efficace et un déplacement de la frontière efficiente. Ce décalage pourrait même être accéléré de manière soudaine et entraîner une transformation radicale de l'industrie rendant obsolète l'industrie traditionnelle. Les entreprises non préparées face à cette situation auront un coût de transition significatif pour atteindre la nouvelle frontière efficiente.

Aujourd'hui, la transformation et l'investissement dans la technologie constituent un facteur clé du succès dans le positionnement concurrentiel futur.

Le schéma suivant illustre le décalage de la frontière efficiente de l'industrie :

Impact des nouvelles technologies sur les frontières d'efficience



Source : SCOR's new strategic plan Quantum Leap 2019-2021 (SCOR's Investor Day – September 4, 2019)

2.1.4 DEFINITION DE L'APPETENCE AU RISQUE

Les objectifs de l'entreprise se décomposent en trois axes :

1. Garder une notation solide qui permet d'assoir une image de marque, maintenir l'intérêt des assureurs à souscrire à une couverture de réassurance (cf. S&P ERM model).
2. Justifier d'un ratio de solvabilité entre 220% et 230%.
3. Maintenir un RoE (*Return on Equity*) à 10%.

2.2 CHOIX ET MISE EN PLACE DE L'OUTIL

À la suite d'un exercice de séminaire stratégique autour des nouvelles opportunités de développement possibles grâce à l'IA et du management du risque y afférent au sein de REA, une étude a été effectuée afin de lister les cas d'usage pertinents pour l'entreprise[2]. En conséquence, le développement d'un modèle IA d'aide à la souscription a été priorisé. Dans la suite du rapport, nous nous plaçons dans un référentiel de phase d'étude.

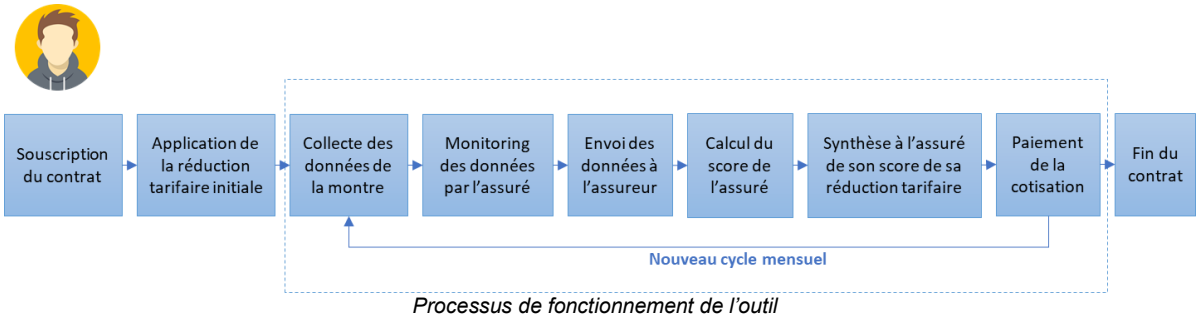
2.2.1 LE PRODUIT

La compagnie REA souhaite mettre en place un outil d'aide à la souscription permettant de mesurer l'état de santé des assurés souhaitant souscrire à son produit temporaire décès. Les données issues de la recherche médicale tendent à établir un lien entre une activité physique plus élevée et un risque de mortalité plus faible[3].

La compagnie REA a développé un modèle de souscription et de suivi du risque permettant de favoriser les assurés qui mènent une vie saine. Les assurés ayant une bonne hygiène de vie se verront récompensés et disposeront d'un rabais financier sur leur prime d'assurance : il s'agit de rembourser aux clients les économies réalisées sur leur risque d'assurance.

2.2.2 L'OUTIL

L'outil repose sur l'utilisation d'une montre connectée portée par l'assuré, et mesurant des paramètres tels que le nombre de pas journaliers, le rythme cardiaque au repos, la qualité du sommeil, ou encore le niveau de stress. Une collecte de ces paramètres permet d'établir une estimation du score santé de l'assuré. Le processus de fonctionnement de l'outil est présenté dans le schéma ci-dessous.



Processus de fonctionnement de l'outil

La réduction tarifaire se verra évaluée en fonction d'un score, principalement calculé à partir de métriques sur lesquelles l'assuré aura un levier d'action. Dans notre exemple, l'assuré sera fortement

[2] Les cas d'usage sont présentés en annexe B.
[3] https://www.cancer.gov/about-cancer/causes-prevention/risk/obesity/physical-activity-fact-sheet

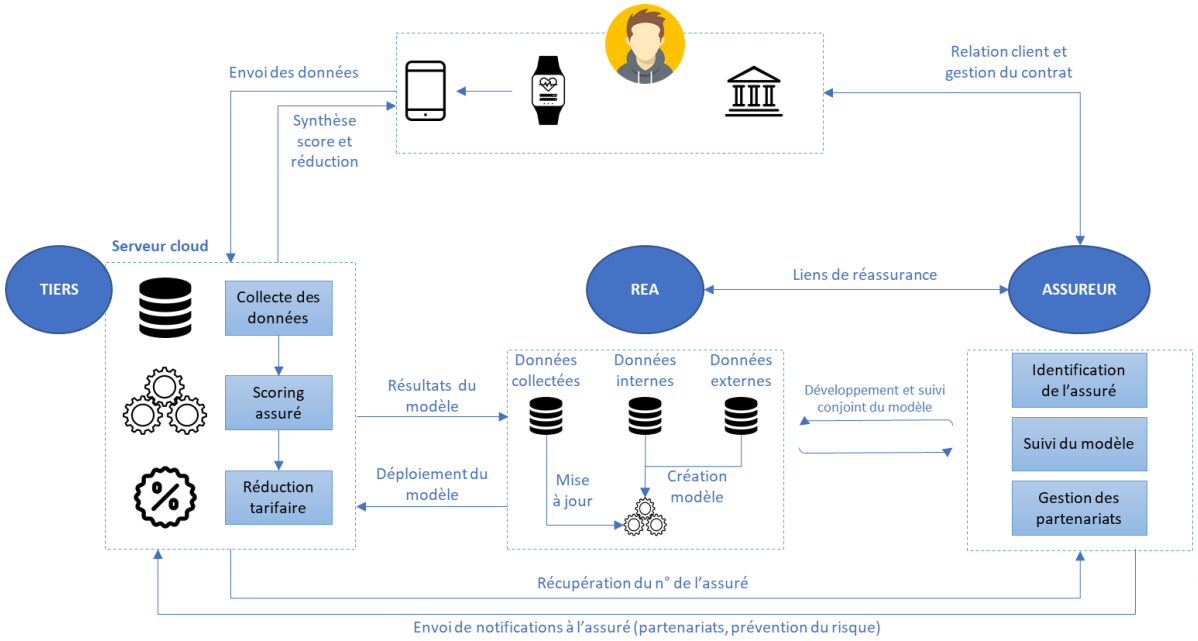
évalué en fonction du nombre de pas journaliers pour lequel des études internes ont démontré un fort pouvoir prédictif de ce paramètre sur le risque d'assurance[4].

2.2.3 LA CARTOGRAPHIE DES INTERACTIONS

L'outil fait intervenir quatre principaux intervenants :

Intervenant	Missions
Le réassureur REA	• Il développe le modèle de prédiction en utilisant des données internes (acquises par son expérience passée) et des données externes (par l'intermédiaire de fournisseurs de données externes). • Il assure le suivi des performances du modèle, et met à jour l'algorithme et les paramètres du modèle lors de revues régulières.
Le client	• Il souscrit le contrat d'assurance et synchronise les données collectées par sa montre connectée avec l'application mobile. • Il reçoit la synthèse de son score santé et de la réduction tarifaire qui lui est appliquée.
L'entreprise de cloud	• Elle met en place une plateforme sur laquelle le modèle de *scoring* est déployé. • Elle collecte les données envoyées par l'application mobile et en assure la sécurité.
Le partenaire assureur	• Il procède à l'identification du client afin d'assurer la relation client et la gestion du contrat d'assurance. • Il négocie les partenariats pour ses clients et définit les notifications à envoyer au client dans le cadre de la prévention du risque ou de campagnes marketing.

Les interactions entre les intervenants sont schématisées ci-dessous.



Interactions entre les intervenants

2.3 DEMARCHE ERM : LE REFERENTIEL COSO

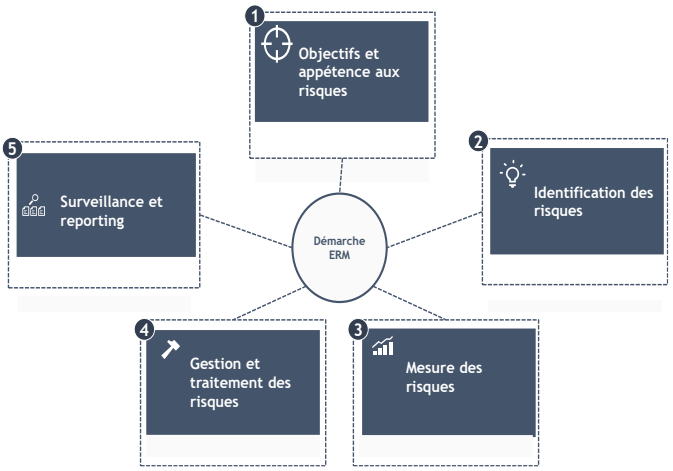
Comme décrit précédemment, REA souhaite évoluer dans le domaine des nouvelles technologies avec un cadre ERM bien défini. Dans ce cadre, nous considérons le référentiel COSO, définissant le management des risques comme un processus mis en œuvre par le conseil d'administration, la direction générale, le management et l'ensemble des collaborateurs de l'organisation. Il est pris en compte dans l'élaboration de la stratégie ainsi que dans toutes les activités de l'organisation. Il est conçu pour identifier les événements potentiels susceptibles d'affecter l'organisation et pour gérer les risques dans

[4] Des précisions sur le fonctionnement de la montre connectée sont fournies en annexe C.

les limites de son appétence pour le risque. Il vise à fournir une assurance raisonnable quant à l'atteinte des objectifs de l'organisation[5].

Ceci se décline concrètement en :

- La **définition des objectifs** et de **l'appétence au Risque** (cf. la définition du cadre en section 2).
- **L'identification** des risques et leur **mesure** (cf. la cartographie des risques en section 3).
- **La gestion et le traitement** des risques, puis la **surveillance et reporting** des risques (cf. la politique de gestion des risques et la mise en place d'une gouvernance en section 4).



3. CARTOGRAPHIE DES RISQUES

Dans le cadre de ce projet, nous proposons une classification des risques en cinq catégories :

1. Les risques liés aux données.
2. Les risques liés aux algorithmes.
3. Les risques d'assurance.
4. Les risques opérationnels.
5. Les risques stratégiques et environnementaux.

Les deux premières catégories fournissent une vision détaillée des risques embarqués dans les modèles de *machine learning* entre les composantes données et algorithmes. Les trois dernières catégories dressent une vision du reste des risques selon les composantes assurances, opérations et stratégiques et environnementaux. A titre d'illustration, une proposition de rattachement des risques présentés ci-dessous avec les normes d'audit IFACI est présentée en annexe D.

La cartographie a été élaborée via une approche *bottom-up* des risques. Il convient de noter que cette cartographie peut fournir une vision biaisée du risque. En effet, les risques peuvent avoir des interdépendances et leur évaluation repose essentiellement sur une vision de jugement d'expert. Ces biais ne seront pas traités dans ce rapport et pourront faire l'objet d'étude ultérieurement.

3.1 IDENTIFICATION DES RISQUES

3.1.1 LES RISQUES LIES AUX DONNEES

Risque	Description du risque
Le risque de disponibilité de la donnée	• Indisponibilité temporaire ou permanente de la donnée alimentant le modèle, empêchant le fonctionnement de l'algorithme. • Transmission tardive de la donnée pour le calcul de l'algorithme à une date précise, notamment pour le calcul des cotisations.
Le risque de sécurité de la donnée	• Modification de la donnée alimentant le modèle par un tiers, impactant le résultat prédit par le modèle. • Donnée frauduleuse ne correspondant pas au profil de l'assuré, la montre connectée étant portée par un tiers. • Erreurs dans les données générées dû à une défaillance de l'objet connecté.

[5] « Le management des risques de l'entreprise - Cadre de Référence » (https://www.coso.org/Documents/COSO-ERM-Executive-Summary-French.pdf)

Risque	Description du risque
Le risque de compliance de la donnée	• Obtention de manière non-éthique ou non-compliant de la donnée alimentant le modèle. • Non-respect des standards éthiques ou règlementaires par un prestataire fournisseur de données.
Le risque de traitement de la donnée	• Mauvaise interprétation des données d'entrées, alimentant le modèle avec de fausses informations. • Traitement inadéquat des données alimentant le modèle (correction des valeurs aberrantes, transformation des variables, réalisation de *feature engineering*, etc.).

3.1.2 LES RISQUES LIES AUX ALGORITHMES

Risque	Description du risque
Le risque de modélisation	• Utilisation de modèles non adaptés pour le risque à modéliser. • L'évolution dans le temps du risque sous-jacent n'est pas capturée par le modèle (drift) • Obsolescence prématurée du modèle vis-à-vis des offres disponibles sur le marché, limitant la pérennité du modèle. • Démutualisation du risque engendrée par une modélisation très granulaire.
Le risque d'explicabilité du modèle	• Le modèle de ML n'est pas explicable (effet boîte noire), rendant impossible d'apporter des explications sur les décisions de l'algorithme (pour les clients, les opérationnels, les auditeurs, le régulateur, etc.). • Impossibilité d'effectuer le traçage des décisions prises par l'algorithme, conduisant à un risque de défaut de preuve. • Risque de ne pas pouvoir expliquer certaines variables d'entrées du modèle générées par des approches automatisées de *feature engineering*.
Le risque de biais	• Données non-représentatives du risque à modéliser, et contenant potentiellement des biais d'échantillonnage. Les biais de la base d'apprentissage seront répliqués et amplifiés par l'algorithme. • Mauvaise estimation des paramètres du modèle à cause d'une mauvaise qualité ou de biais dans les données d'apprentissage.
Le risque de mise à jour de l'algorithme	• Mise à jour des logiciels utilisés pour le modèle, conduisant à une incompatibilité entre logiciels et à une indisponibilité du modèle. • Impossibilité de justifier ou de reproduire les résultats générés par le modèle suite à une mise à jour. • Réentrainement de l'algorithme avec des données biaisées. • Développements additionnels du modèle contenant des bugs non-testés ou non-identifiés.
Le risque d'utilisation de code externe	• Utilisation de code externe (librairies *open source* ou non, interfaces API, etc.), vital pour l'exécution de l'algorithme, et dont l'altération ou la non-disponibilité bloquerait l'exécution de l'algorithme.

3.1.3 LES RISQUES D'ASSURANCE

Risque	Description du risque
Le risque de tarification	• Tarif sous-estimé (réduction tarifaire trop importante) par rapport au prix du risque, ou tarif surestimé et générateur d'anti-sélection. • Inadéquation entre la règle de calcul de la réduction tarifaire contractuelle et le prix du risque.
Le risque de souscription	• Souscriptions de mauvaise qualité, quant aux risques souscrits, malgré leur conformité aux règles de souscription.
Le risque de mortalité	• Augmentation de fréquence des décès et évolution défavorable de la charge sinistre.

3.1.4 LES RISQUES OPERATIONNELS

Risque	Description du risque
Le risque de disponibilité des systèmes	• Indisponibilité passagère de systèmes mis à disposition par un prestataire, par exemple la plateforme *cloud* (panne système, insuffisance, incendie, etc.). • Obsolescence de l'objet connecté (la montre ne fonctionne plus) et nécessité de le remplacer. • Utilisation d'un service fourni par un prestataire externe (plateforme cloud) qui fait défaut. Risque de ne pas pouvoir re-internaliser la solution afin d'assurer la continuité du fonctionnement de l'algorithme.
Le risque humain	• Mauvaise utilisation ou manque d'expertise dans le développement et la mise en œuvre du modèle. • Mauvaise interprétation ou biais d'interprétation des résultats dans le cadre du suivi du modèle. • Homme clé qui concentre toutes les connaissances relatives au modèle. • Erreur dans l'implémentation du modèle.
Le risque cyber	• Acte de malveillance informatique (virus, destruction de fichiers, piratages, etc.) sur les données présentes sur la plateforme cloud, chez l'assureur ou le réassureur. • Paiement d'une rançon pour débloquer les fichiers chiffrés par le *ransomware*. • Vol et divulgation de données (algorithme, données clients, etc.), avec un fort impact sur le risque de réputation.

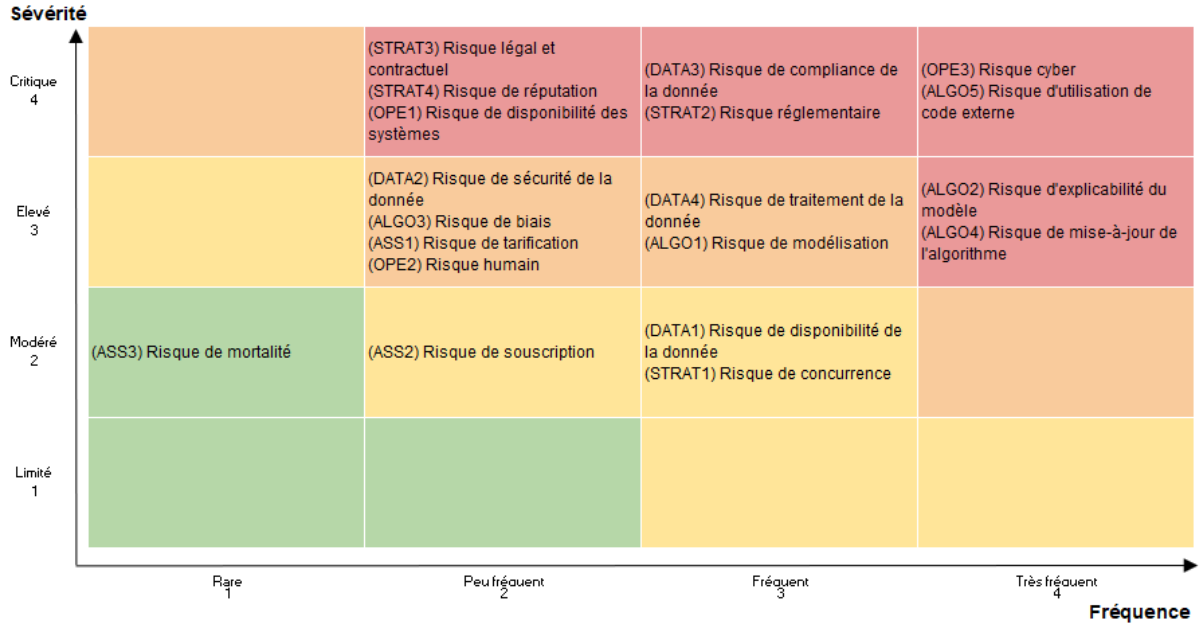
3.1.5 LES RISQUES STRATEGIQUES ET ENVIRONNEMENTAUX

Risque	Description du risque
Le risque de concurrence	• Copie de mon modèle par la concurrence. • La concurrence développe un modèle plus performant, mais avec des standards d'éthique plus faibles que les miens.
Le risque réglementaire	• Risques liés à l'apparition de nouvelles lois ou règlements, et à leur application. • Non-conformité aux exigences de l'ACPR ou de la CNIL (loi française Informatique et Libertés, RGPD) en cas d'audit de la compagnie.
Le risque légal et contractuel	• Responsabilité de l'assureur vis-à-vis de la sécurité des données impliquée par la solution proposée à ses assurés (stockage de données dans le cloud et utilisation d'une montre connectée.
Le risque de réputation	• Risque provenant d'un manquement à la réglementation (mise en demeure, sanction publique). • Risque provenant de la mise en cause par une association de consommateurs, par la presse, ou par un client influent via les réseaux sociaux, quant à un dysfonctionnement de l'algorithme (typiquement un biais).

3.2 EVALUATION DES RISQUES

Les risques étant identifiés, il s'agit maintenant de les évaluer selon leur degré d'importance. Celui-ci est établi selon deux composantes : la fréquence de survenance et le niveau de sévérité. Chaque composante est notée sur une échelle de 1 à 4. Un score global peut également être calculé pour chaque risque comme résultat de la multiplication des deux échelles de fréquence et de sévérité. Cette méthodologie permet de classer les risques selon leur degré d'importance.

Dans notre étude, nous avons adopté une approche qualitative d'évaluation de ces risques et nous avons tenté de les évaluer selon notre jugement. L'évaluation de ces risques n'est bien entendu pas figée, l'objectif étant d'établir une cartographie des risques brute. La représentation visuelle des risques selon ces deux axes peut être réalisée via une *heatmap* :



Dans le graphique ci-dessus, il apparait que les risques les plus critiques sont les risques relatifs à l'environnement stratégique de l'entreprise ainsi que les risques relatifs aux algorithmes, en particulier, le risque d'utilisation de code externe, le risque d'explicabilité du modèle et le risque de mise à jour de l'algorithme.

3.3 RISQUES : PRIORISATION ET MITIGATION

Mesures de mitigations proposées et *Key Risk Indicators* (KRI) associés

Id	Risque	Description du risque	Fréq.	Sévérité	Mesure de mitigation du risque *(Eviter / Transférer / Réduire / Absorber)*	KRI
DATA1	Risque de disponibilité de la donnée	Indisponibilité temporaire ou permanente de la donnée, empêchant le fonctionnement de l'algorithme.	3	2	REDUIRE - si la donnée n'est pas transmise par l'objet connecté, alors la tarification appliquée sera traditionnelle, et aucune réduction ne sera appliquée. - envoi de notifications à l'assuré pour l'inciter à se maintenir dans le programme et continuer à collecter de la donnée.	Fréquence de suivi : mensuel Indicateur : - tableau de bord contenant la proportion de contrats où les données ne sont pas transmises - seuil d'alerte de proportion de contrats avec données non-transmises, déclenchant une campagne de sensibilisation (notifications et rappels de la valeur du produit, campagne marketing, vérification des problèmes techniques de l'objet connecté).
DATA2	Risque de sécurité de la donnée	La donnée transmise ne correspond pas au risque et a été modifiée volontairement (fraude, hacking objet connecté) ou involontairement (objet connecté défaillant)	2	3	REDUIRE - la conception du produit prévoit dans son design des limitations du risque souscrit (limites de capitaux, réduction maximale offerte, rédaction des termes de couverture contractuel, etc.). - restriction de la souscription du produit en réseau traditionnel pour limiter les fraudes. - création d'une signature électronique pour chaque assuré à partir des données collectées afin de s'assurer de l'identité de l'utilisateur (profil utilisateur unique). - mise en place de programme de *bug bounty* pour favoriser la remontée de failles techniques par les utilisateurs.	Fréquence de suivi : mensuel Indicateur : tableau de bord contenant : - le suivi des données collectées pour détecter les dérives vs. l'attendu lors de la modélisation (définis lors de recherches, d'expertises médicales, etc.). Détection des *outliers*. - seuil d'alerte en cas de dérive en dehors de l'intervalle de confiance autour du score santé attendu. - réalisation de tests aléatoires ponctuels sur les données collectées. - nombre d'alertes générées pour non-respect de la signature électronique.
DATA3	Risque de compliance de la donnée	Les données utilisées par le modèle ne respectent pas les standards règlementaires ou éthiques.	3	4	REDUIRE - design du produit de type *privacy by design* dès la phase de conception des modèles (implication du *Data Privacy Officer* dans la conception et la validation du modèle). - procédure de vérification et de validation des données achetées auprès de tiers par le DPO.	Fréquence de suivi : annuel Indicateur : - revue des prestataires fournisseurs de données - audit des modèles pour s'assurer que la procédure a été correctement suivie. - création d'une base d'incidents et d'une adresse mail de contact et pour collecter les alertes remontées par les collaborateurs.
DATA4	Risque de traitement de la donnée	Les données ne sont pas correctement interprétées ou retraitées.	3	3	REDUIRE - documentation des jeux de données utilisés lors de la conception et de la validation du modèle. - nécessité de la complétude de la documentation des jeux de données lors de la validation du modèle. - formation et contrôle des compétences des équipes. - réalisation d'*Analysis of Change* (AOC) lors de changement de traitement de variables.	Fréquence de suivi : annuel (ou en fonction de la feuille de route de l'audit et de la gestion des risques) Indicateur : - audit des modèles pour s'assurer que la procédure a été correctement suivie.
ALGO1	Risque de modélisation	Le modèle n'est pas adapté pour le risque à modéliser. Le modèle n'est pas ou n'est plus pertinent.	3	3	REDUIRE - sélection de métriques de performances (AUC, Gini, etc.) pour évaluer l'efficacité des modèles, et mise en place de seuils minimums de performance pour valider le modèle. - contrôle du sur-apprentissage via des jeux de données séparés. - assurer une veille R&D sur les modèles IA.	Fréquence de suivi : annuel Indicateur : - audit des modèles pour s'assurer que la procédure a été correctement suivie. - base d'alerte sur les évolutions technologiques et les risques opérationnels des modèles (comprenant également une veille des partenariats réalisés par nos prestataires avec la concurrence et éviter qu'ils ne réalisent des modèles plus performants avec la concurrence, rendant notre modèle obsolète).

Id	Risque	Description du risque	Fréq.	Sévérité	Mesure de mitigation du risque *(Eviter / Transférer / Réduire / Absorber)*	KRI
ALGO2	Risque d'explicabilité du modèle	Le modèle n'est pas explicable rendant impossible d'apporter des explications sur les décisions de l'algorithme.	4	3	REDUIRE - définition de méthodes d'explications, avec différents niveaux d'explicabilité (1: observation, 2: justification, 3: approximation, 4: réplication) pour une communication auprès de différentes audiences (client, contrôleur interne, auditeur, régulateur) - cf. démarche ACPR.	Fréquence de suivi : trimestriel Indicateur : - tableau de bord des niveaux d'explicabilité des modèles (fiche technique des modèles)
ALGO3	Risque de biais	Les données d'apprentissage sont non-représentatives du risque à modéliser, et contiennent potentiellement des biais d'échantillonnage.	2	3	REDUIRE - réalisation d'analyses exploratoires (statistiques descriptives, analyse de corrélations des variables) sur les données pré-modélisation.	Fréquence de suivi : trimestriel Indicateur : - tableau de bord des niveaux de traitement de données des modèles (fiche technique des modèles)
ALGO4	Risque de mise à jour de l'algorithme	Mise à jour des logiciels utilisés pour le modèle, conduisant à une incompatibilité entre logiciels et à une indisponibilité du modèle. Réentrainement de l'algorithme avec des données biaisées.	4	3	REDUIRE - mise en place de systèmes de *versioning* logiciels, permettant de restaurer une version fonctionnelle antérieure à la mise à jour. - refuser de mettre à jour le modèle avec de nouvelles données d'entrainement si celles-ci détériorent la qualité du modèle (méthode de Roni).	Fréquence de suivi : trimestriel Indicateur : - tableau de bord des incidents générés lors de mises à jour de briques logicielles
ALGO5	Risque d'utilisation de code externe	Utilisation de code externe dont l'altération ou la non-disponibilité bloquerait l'exécution de l'algorithme.	4	4	REDUIRE - cartographie des librairies utilisées et des dépendances vers des codes externes utilisés par les modèles. -mise en place d'un plan d'action pour réduire la dépendance par ex, internaliser certaines solutions critiques (à différents horizons de temps) ou avoir des modes de fonctionnement (dégradés) en cas de défaillance liée à ces codes externes	Fréquence de suivi : trimestriel Indicateur : - cartographie des librairies et applications tierces utilisées dans les modèles
ASS1	Risque de tarification	Tarif sous-estimé par rapport au prix du risque, ou tarif surestimé et générateur d'anti-sélection.	2	3	REDUIRE - ce risque est essentiellement une conséquence des risques DATA et ALGO : mauvaise estimation due à une erreur de traitement de données, de modélisation, de biais, etc.	*cf. KRI des risques DATA et ALGO.*
ASS2	Risque de souscription	Souscriptions de mauvaise qualité, quant aux risques souscrits, malgré leur conformité aux règles de souscription.	2	2	REDUIRE - la conception du produit prévoit dans son design des limitations du risque souscrit : limites de capitaux, réduction maximale offerte, rédaction des termes de couverture contractuel, etc. (cf. action de mitigation des risques DATA et ALGO).	*cf. KRI des risques DATA et ALGO.*
ASS3	Risque de mortalité	Augmentation de fréquence des décès et évolution défavorable de la charge sinistre.	1	2	REDUIRE (et TRANSFERER) - ce risque est essentiellement une conséquence des risques DATA et ALGO (cf. risque de tarification). - mise en place d'une réassurance si nécessaire (volume de ventes important, etc.)	*cf. KRI des risques DATA et ALGO.*
OPE1	Risque de disponibilité des systèmes	Utilisation d'un service fourni par un prestataire externe (plateforme cloud) qui	2	4	REDUIRE - appliquer le principe de réversibilité des systèmes dès la phase de conception du modèle : une solution de re-internalisation doit être conçue en cas de défaillance de la plateforme Cloud.	Fréquence de suivi : annuel Indicateur : - audit des modèles pour s'assurer que la procédure a été correctement

Id	Risque	Description du risque	Fréq.	Sévérité	Mesure de mitigation du risque *(Eviter / Transférer / Réduire / Absorber)*	KRI
		fait défaut. Risque de ne pas pouvoir re-internaliser la solution.			- réaliser des points réguliers avec les prestataires. - surveiller la solvabilité des prestataires.	suivie. - base d'incidents sur les incidents recensés chez les prestataires. - ratio de solvabilité / niveau d'endettement des prestataires
OPE2	Risque humain	Mauvaise utilisation, manque d'expertise dans le développement du modèle, mauvaise interprétation des résultats. Homme clef.	2	3	REDUIRE - mise en place de plans de recrutement de talents et de programmes d'*on-boarding* pour faire rapidement monter en compétence les nouvelles recrues. - mise en place de programmes de formation pour les équipes en place (actuariat, risk management, audit, contrôle interne, juridique) afin de développer des doubles compétences en IA au sein de chaque équipe.	Fréquence de suivi : annuel Indicateur : - tableau de bord de *turn-over* des équipes (nouvelles recrues, démissions, promotions, certifications obtenues de type ERM, data science pour l'actuariat, etc.).
OPE3	Risque cyber	Acte de malveillance informatique (virus, destruction de fichiers, piratages, etc.), et paiement d'une rançon.	4	4	REDUIRE - mise en place d'un système de gestion de la sécurité d'information pour identifier et traiter les risques. - audit régulier des systèmes d'informations en interne et en externe (*pen-testing* des systèmes) - mise en place de critères basés sur la sécurité des systèmes et de l'information lors du choix du partenaire qui sera retenu pour la plateforme Cloud.	Fréquence de suivi : trimestriel Indicateur : - tableau de bord sur le suivi des risques IT liés à la sécurité d'information. - base d'incidents IT sur les tentatives de piratage (tentatives de *phishing*, attaques DoS, etc.)
STRAT 1	Risque de concurrence	La concurrence copie mon modèle, ou développe un modèle plus performant.	3	2	REDUIRE - mise en place d'équipes (ou de ETP) dédiées à la R&D afin de maintenir l'avance technologique des modèles. - participation à des groupes de travail (par exemple FFSA) avec les concurrents afin d'échanger sur les pratiques de la place pour éviter la concurrence déloyale et garantir les intérêts des assurés.	Fréquence de suivi : trimestriel Indicateur : - suivi des nouvelles offres du marché et mise en place d'un benchmark concurrentiel.
STRAT 2	Risque réglementaire	Risques liés à l'apparition de nouvelles lois ou règlements, et à leur application.	3	4	REDUIRE - ce risque est essentiellement une conséquence du risque DATA3 (risque de compliance de la donnée) s'agissant de la règlementation en vigueur. - mise en place d'une veille règlementaire et participation aux travaux de consultation du régulateur pour anticiper l'apparition de nouvelles lois.	*cf. KRI risque de compliance de la donnée.*
STRAT 3	Risque légal et contractuel	Responsabilité de l'assureur vis-à-vis de la sécurité des données impliquée par la solution proposée à ses assurés.	2	4	REDUIRE - implication des équipes juridiques dès la phase de conception de la solution (création du modèle, sélection du partenaire). - mise en place de groupes de travail réfléchissant à des solutions futures potentielles (création d'un fonds d'indemnisation, achat d'assurances de protection juridique, mise en place de provisionnement pour litige dédié, consultation du régulateur pour clarifier les responsabilités des différentes parties prenantes).	
STRAT 4	Risque de réputation	Mise en cause par une association de consommateurs, par la presse, etc. quant à un dysfonctionnement de l'algorithme.	2	4	REDUIRE - mise en place d'une équipe de gestion des réclamations (mail, téléphone) et formation de l'équipe communication aux remontées des dysfonctionnements de l'algorithme par les clients.	Fréquence de suivi : trimestriel Indicateur : - tableau de bord des indicateurs de réclamations clients (appels reçus, par motif, etc.) et d'indicateurs de sentiment sur la marque d'assurance sur internet (forums, réseaux sociaux, notes *Trustpilot*, etc.)

4. POLITIQUE DE GESTION DES RISQUES

4.1 CADRE REGLEMENTAIRE : REFLEXIONS DE L'ACPR ET CONSULTATION DES ORGANISMES DE PLACE

L'ACPR porte une attention particulière à l'intelligence artificielle depuis 2018. Le pôle Fintech-Innovation a mené des travaux exploratoires via des entretiens et des ateliers techniques à la suite desquels, une note de réflexion a été rédigée[6]. Celle-ci porte en particulier, sur l'état de l'art des techniques d'IA et la réglementation du secteur financier ainsi que leur gouvernance associée.

Les deux axes principaux mis en évidence par cette note sont l'évaluation et la gouvernance des algorithmes IA. Le régulateur a ouvert une consultation publique en 2020 et a rendu une synthèse des réponses des acteurs du marché.

4.1.1 EVALUATION DES ALGORITHMES IA

L'analyse du régulateur a identifié 4 principes interdépendants dans l'évaluation des algorithmes IA :

- Le traitement adéquat des données.
- La performance des algorithmes ML[7].
- La stabilité des algorithmes ML.
- L'explicabilité (c'est-à-dire, la transparence et l'interprétabilité de l'algorithme).

L'attention a été portée particulièrement sur ce critère, tant par le régulateur que par les acteurs de marché dans leurs réponses lors de la consultation publique.

4.1.2 GOUVERNANCE DES ALGORITHMES IA

Le régulateur souligne la nécessité de repenser une gouvernance adéquate qui tient compte des spécificités liées à l'intégration de l'intelligence artificielle dans les processus métiers en finance et en assurance.

Le message clé du régulateur porte sur la nécessité de mettre l'accent sur les aspects de gouvernance, et ce, dès la phase embryonnaire de conception des algorithmes. En particulier, celle-ci doit s'appuyer sur des équipes d'experts, dotés de doubles compétences, leur permettant de couvrir les domaines fonctionnels et techniques. L'analyse menée par l'ACPR sur la gouvernance s'articule autour des six aspects suivants :

- L'intégration dans les processus métiers.
- Les interactions entre humain et algorithme.
- La sécurité et l'externalisation.
- Les processus de validation initiale.
- Les processus de validation continue.
- L'audit.

4.1.3 CONSULTATION PUBLIQUE ET QUESTIONNAIRE AUX ACTEURS DE MARCHE

La synthèse communiquée contient également le questionnaire communiqué aux acteurs du marché lors de la consultation publique de septembre 2020. Dans le cadre de notre étude sur la digitalisation au sein de REA, nous avons analysé la liste des questions et nous avons tenté d'y apporter des

[6] https://acpr.banque-france.fr/gouvernance-des-algorithmes-dintelligence-artificielle-dans-le-secteur-financier
[7] ML : *Machine Learning*.

éléments de réponse. Cette mise en situation nous est apparue importante afin de confronter et de challenger notre démarche ERM ainsi que notre politique de gestion des risques vis-à-vis de l'approche proposée par le régulateur. En particulier, le questionnaire fournit une structure pertinente, d'une part, pour conduire les entretiens avec les équipes d'opérationnels lors de la mise en place des processus de conception et de validation, et d'autre part, pour rédiger la procédure de gouvernance.

Bien que le régulateur ait apporté des clarifications et des éléments de réponse sur des thématiques plus étendues (détaillées en annexe E) que celles considérées par notre étude de cas REA, plusieurs éléments sont apparus transposables à notre problématique.

4.1.4 APPLICATION A NOTRE ETUDE DE CAS REA

Nous avons réalisé l'exercice en plusieurs étapes :

1. Définir une démarche ERM pour les algorithmes IA.
2. Challenger et compléter notre démarche ERM avec l'approche du régulateur, dans le but de :
    a) Définir les bonnes pratiques à mettre au service de notre entreprise REA,
    b) Définir les usages et les indicateurs de risques et de suivi appliqués à notre gouvernance (cf. grille de *scoring* du modèle IA présenté en section 4.2.7).

La réalisation de ce travail sur le document de gouvernance de l'ACPR peut paraître précoce à ce stade, mais celui-ci nous est apparu essentiel. En effet, il permet, d'une part, de saisir les principaux enjeux du point de vue du régulateur, et d'autre part, d'anticiper les évolutions réglementaires ou les potentiels audits portant sur les modèles d'IA.

4.2 PRINCIPES D'EVALUATION DES MODELES IA

Le cycle de vie d'un modèle « classique », ne contenant pas nécessairement d'algorithme de ML, comporte les étapes de conception, d'implémentation et d'utilisation du modèle. De plus, les phases de validation, de suivi et d'ajustement du modèle sont également présentes tout au long du cycle de vie. Un schéma est présenté en annexe F à titre d'illustration.

Cependant, les modèles contenant des algorithmes de ML nécessitent la mise en place d'une gouvernance plus avancée en réponse à la complexification des étapes du cycle de vie du modèle.

4.2.1 PRESENTATION DU CYCLE DE VIE D'UN MODELE IA

L'ACPR propose un diagramme détaillé pour représenter l'ensemble des étapes du cycle de vie d'un modèle contenant des algorithmes de ML en mettant en évidence les éléments importants, en particulier :

- **Les principales étapes** : la conception et l'apprentissage du modèle, la validation du modèle (en tenant compte de quatre critères d'évaluation), le déploiement du modèle, et la maintenance du modèle.
- **Les itérations** liées au (ré)apprentissage et au retour d'expérience lors de l'étape de maintenance du modèle. La distinction entre le contrôle permanent (conception, validation, déploiement et maintenance) et le contrôle périodique (audit interne ou externe).

Le contrôle périodique effectué par l'audit, c'est-à-dire de niveau 3, sera décrit par la suite. Celui-ci fera notamment la distinction entre l'évaluation analytique et l'évaluation empirique des modèles.

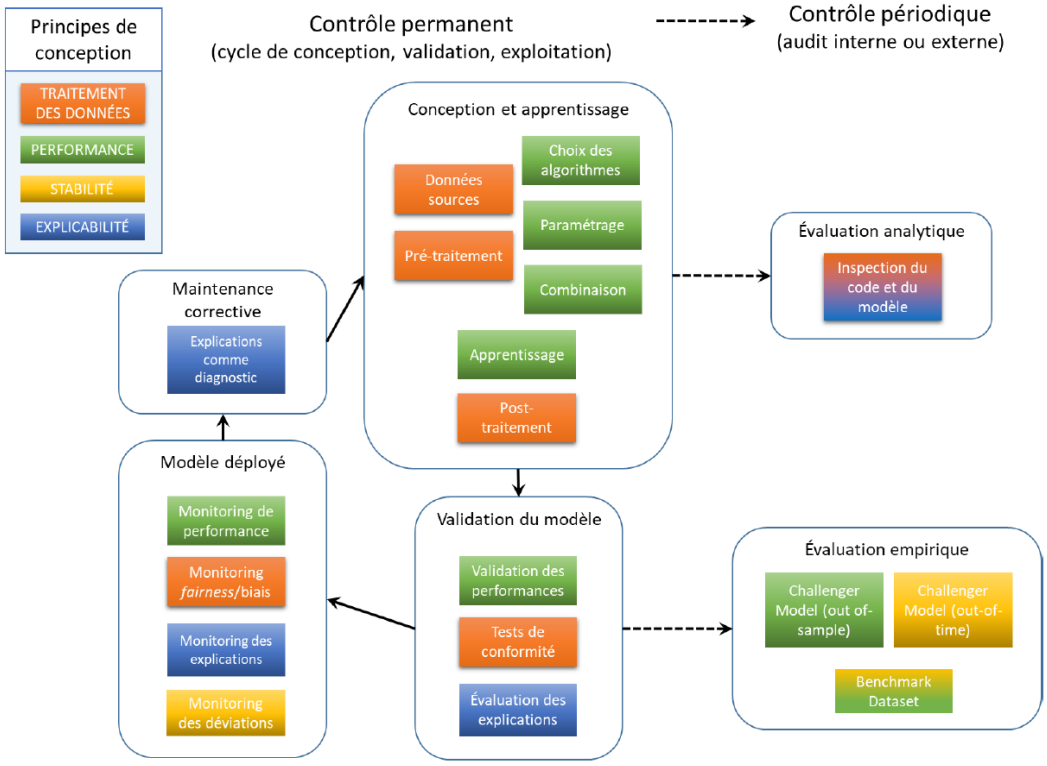
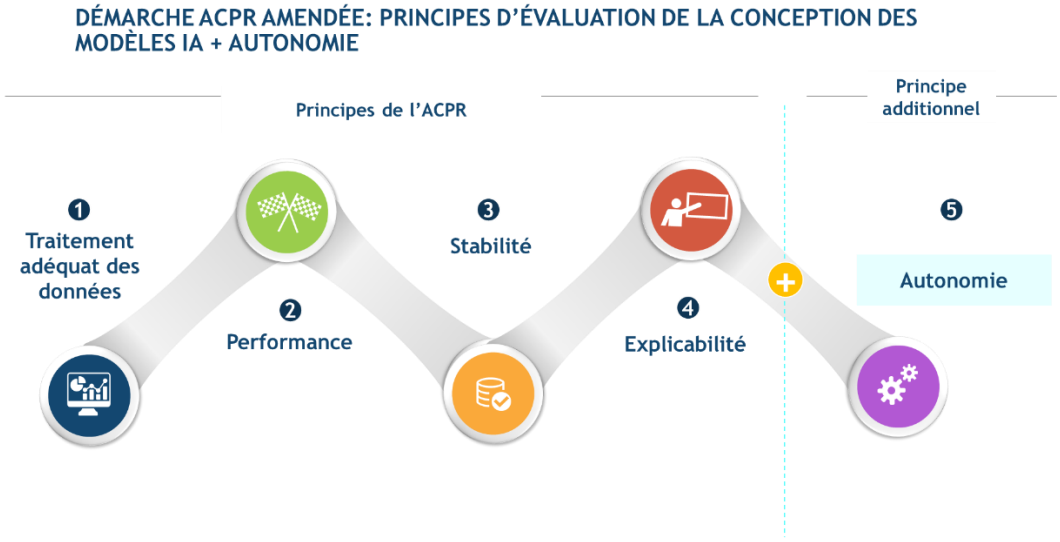


Schéma : Etapes du cycle de vie d'un modèle contenant des algorithmes de ML (ACPR)

Comme évoqué paragraphe 4.1.1, les quatre principes de conception servant à l'évaluation des algorithmes IA par l'ACPR sont les suivants : le traitement des données, la performance, la stabilité et l'explicabilité.

Bien que les principes proposés par l'ACPR soient clairs et couvrent un spectre large, nous estimons qu'un principe supplémentaire mériterait d'être mis en évidence : il s'agit du principe d'autonomie. Ce principe correspond à l'évaluation du degré d'indépendance de l'algorithme vis-à-vis de ressources externes. Cette composante a été identifiée lors de notre démarche et sera détaillée par la suite.

Les quatre principes de conception énoncés par l'ACPR sont directement transposables aux éléments identifiés lors de notre démarche ERM et sont détaillés dans la section suivante. De même, le principe d'autonomie identifié lors de notre démarche ERM, et spécifique pour notre étude de cas REA, est détaillé par la suite.



4.2.2 LE TRAITEMENT ADEQUAT DES DONNEES

Le traitement adéquat des données est un principe essentiel de tout algorithme. D'une part, celui-ci a un impact sur la performance, et, d'autre part, il constitue la première brique à contrôler au regard de la conformité réglementaire et des questions d'éthiques (par exemple, s'assurer de l'équité dans les traitements et de l'absence de biais discriminatoires).

La documentation est un point clé pour l'ensemble des traitements des données ainsi que pour le reste du processus de conception du modèle de ML (par exemple, la documentation du code source des algorithmes, de la performance des modèles, etc.).

Il est également nécessaire de procéder à l'identification et à l'évaluation des risques associés à la conformité réglementaire et à l'éthique. Un plan de remédiation et des techniques de détection doivent être mises en œuvre afin de réduire le risque de biais.

L'analyse du régulateur sur les points de la conformité réglementaire, d'éthique et d'équité est détaillé en annexe G.

Une réflexion a été mené via un groupe de travail au sein de REA afin de déterminer les bonnes pratiques de traitements de la donnée ainsi que le processus de management (cf. annexe G. bis)

4.2.3 LE PRINCIPE DE PERFORMANCE

La performance de l'outil est évaluée par l'intermédiaire :

- De métriques de performance prédictive (par exemple, la valeur de l'AUC[8]) permettant d'évaluer l'efficacité technique de l'algorithme de l'IA.
- De métriques de performance commerciale (sous contraintes réglementaires).

4.2.4 LE PRINCIPE DE STABILITE

La stabilité se caractérise par la robustesse et la résilience du comportement d'un algorithme de ML au cours du temps. Il est important de garantir la stabilité de la qualité et des caractéristiques du modèle tout en assurant un contrôle continu des risques de dérive des modèles déployés en production (correspondant principalement à des risques opérationnels et de conformité). Trois sources principales de dérive sont identifiées (détaillées en annexe H) : les dérives temporelles, les généralisations et le réapprentissage.

4.2.5 LE PRINCIPE D'EXPLICABILITE

Les points d'attention des différents participants au cycle de vie d'un modèle IA (des concepteurs, jusqu'aux utilisateurs finaux en production, les clients ou les auditeurs) portent principalement sur le critère d'explicabilité.

Le principe d'explicabilité peut-être décomposé en trois concepts : la transparence (à l'opposé de la notion de « boîte noire »), l'auditabilité et l'interprétabilité.

L'explicabilité répond à des objectifs de conformité réglementaire, de validation par les équipes ou de compréhension de la qualité des prédictions par les différentes parties prenantes.

Le régulateur a défini un ensemble de caractérisations permettant de fournir une explication de très bonne qualité, et propose quatre niveaux d'explication (observation, justification, approximation et réplication). Il est important de déterminer les facteurs d'influence et le contexte : d'une part, les

[8] AUC: Area under the curve

destinataires des explications, et, d'autre part, la criticité des risques associés. Cette démarche permet de clarifier les attendus, et de déterminer les exigences et la forme des explications. Trois formes d'explications sont identifiées :

- Le client ou consommateur : explication simple.
- Le contrôleur interne : explication fonctionnelle.
- L'auditeur : explication technique.

Le principe d'explicabilité est détaillé en annexe I.

4.2.6 LE PRINCIPE D'AUTONOMIE

Nous définissons le principe d'autonomie comme le degré d'indépendance de l'algorithme par rapport à tout élément externe ou partenaire tiers qui pourrait, en cas de défaut de celui-ci, mettre en difficulté la bonne exécution des tâches réalisées par le modèle. Les modèles étant développés de façon modulaire, la défaillance d'une seule composante peut conduire à la défaillance globale du modèle.

L'analyse de la criticité des modèles dans les processus métiers démontre l'importance de mettre en place un plan de secours en cas de défaillance du modèle. Pour cela, il est nécessaire de définir dès la conception la réversibilité du modèle en cas de défaillance de composantes externes.

Les composantes externes pouvant affecter l'exécution de l'algorithme sont les suivantes :

- Utilisation de ressources externes lors de la conception du modèle (consultants, outils de code, etc.).
- Utilisation de librairies de code externes.
- Utilisation de solutions d'hébergement ou d'exploitation chez un partenaire tiers.

Un modèle faisant intervenir peu de composantes externes permettrait de générer un bon score selon le principe d'autonomie. Cependant, un tel modèle pourrait générer un score détérioré selon les principes de performance, d'explicabilité ou de stabilité dans le cas où la compagne ne dispose pas de ressources internes capables de développer et de maintenir des composantes de qualité en interne (par exemple, des librairies de code). Un arbitrage doit donc être réalisé en fonction de la criticité du modèle dans les processus métiers.

Les pistes de mitigation des risques découlant du principe d'autonomie sont les suivantes : réalisation de process de due diligence en cas de collaboration avec un tiers, formation continue des équipes, revues des librairies et codes externes utilisées ou encore la mise en place de modèles de secours, indépendants, en cas de défaillance du modèle principal.

4.2.7 SCORING DU MODELE IA

Les modèles peuvent par conséquent être évalués selon les principes décrits précédemment. Une note élevée reflète la bonne qualité du modèle suivant la composante considérée. Pour chaque composante, la note finale est obtenue comme moyenne des notes attribuées par les différents intervenants (opérationnels, risk management, audit, DPO) lors de la phase de validation du modèle.

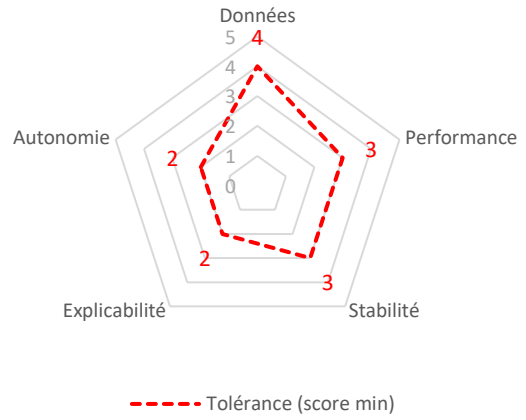
Bien souvent, des arbitrages sont nécessaires car des composantes peuvent évoluer en sens opposés, par exemple les couples performance & autonomie, ou encore explicabilité & performance.

Le graphique en toile d'araignée permet de représenter le modèle selon ces cinq composantes. Il est également possible de déterminer une zone de tolérance pour chaque composante lors de la création et du suivi du modèle en définissant des scores minimums à respecter.

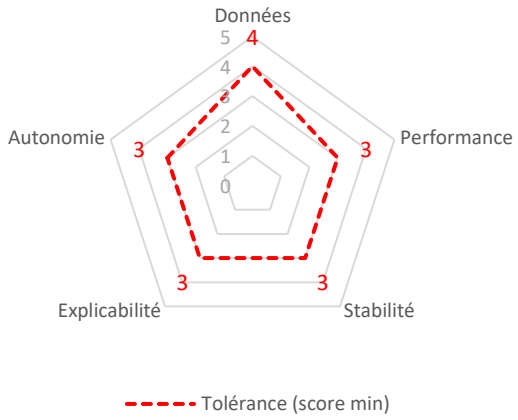
Lors du lancement du projet, la zone de tolérance initialement définie par REA requiert une notation minimale élevée pour le traitement des données, pour la performance et pour la stabilité du modèle. S'agissant du critère d'explicabilité, le niveau minimal requis est bas en raison du niveau de criticité limité de l'algorithme intégré dans le processus métier. Enfin, s'agissant du critère d'autonomie, le niveau minimal requis est également bas en raison de la difficulté actuelle de développer en interne des librairies indépendantes et d'être en mesure de pouvoir réinternaliser intégralement la solution dès le lancement du projet.

Cependant, à moyen terme, l'objectif de REA est d'augmenter les niveaux minimaux requis, en particulier pour les critères d'autonomie et d'explicabilité, après augmentation des compétences en interne. La zone de tolérance est donc évolutive et peut être révisée lors des comités de suivi de modèle.

Zone de tolérance au lancement du projet



Zone de tolérance à moyen terme



4.3 MISE EN PLACE D'UNE GOUVERNANCE

L'adoption de modèles IA dans les processus métiers de l'assurance remanie les attributs de la gouvernance en place. La politique de culture du risque étant déployée, les normes éthiques et professionnelles, les processus et procédures doivent-être revus à la lumière d'un écosystème de plus en plus digitalisé et intégrant l'intelligence artificielle comme outil majeur pour conduire ses opérations quotidiennes.

En particulier, la cartographie des risques doit être régulièrement actualisée et évaluée afin de garantir une vision cohérente et globale du risque. Ce travail a été effectué dans la section précédente et a permis d'identifier les risques et les éléments relatifs à l'atténuation du risque.

L'analyse du régulateur sur la gouvernance dans l'univers digitalisé est pertinente et exhaustive. Elle liste les éléments clés de gouvernance dans le cadre de l'introduction de modèles IA :

- La séparation des fonctions de contrôles et l'évolution de leurs rôles.
- Les procédures opérationnelles.
- L'organisation de la gestion des risques liée à l'IA.

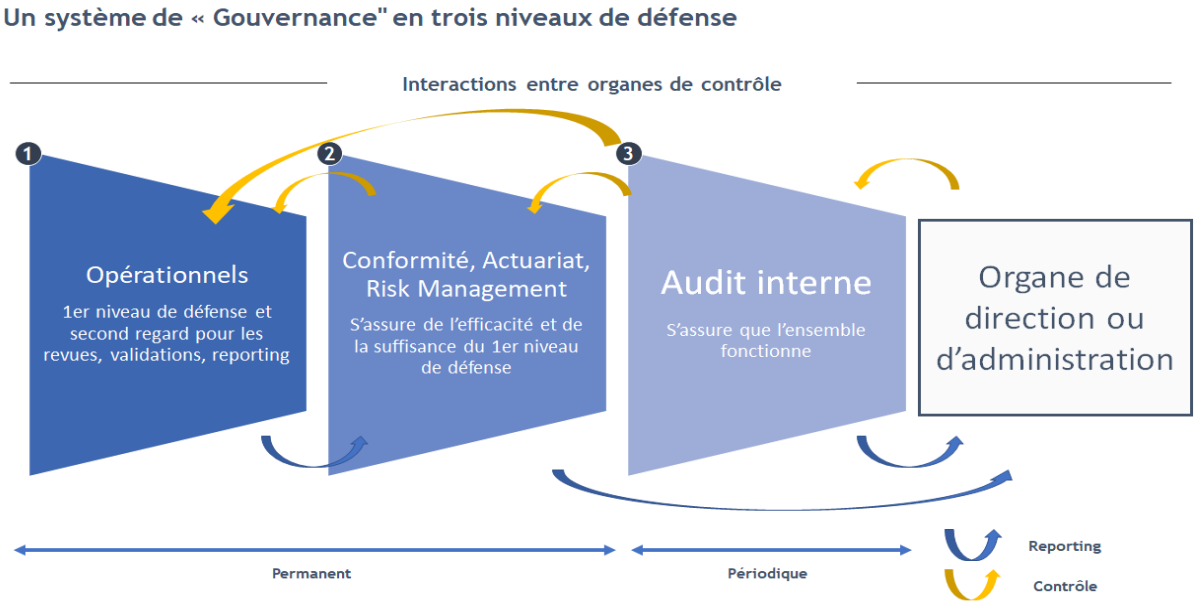
Le régulateur préconise aussi de porter l'attention, dès la phase de conception des algorithmes, sur les aspects de métiers suivants : l'intégration dans les processus, les interactions entre humain et algorithme, la sécurité et l'externalisation, les processus de validation initiale, les processus de validation continue, et l'audit. Dans le cadre de REA, notre objectif est de créer un cadre de gouvernance efficace qui s'inscrit dans le respect de la réglementation et des référentiels.

Nous abordons les points d'attention de l'ACPR, sous un angle spécifique à notre étude de cas REA, notamment lors de la rédaction des procédures opérationnelles. Dans la partie suivante, nous mettons l'accent sur l'importance de la séparation des fonctions soulignée par l'ACPR. Nous nous concentrons par conséquent sur le rôle des fonctions de contrôle et leur organisation.

4.3.1 SEPARATION ET ORGANISATION DES FONCTIONS DE CONTROLES

Afin de permettre une gestion saine et prudente de l'activité, un ensemble de dispositifs est mis en place (organes, fonctions, procédures, délégation de pouvoirs). L'existence de plusieurs niveaux de contrôle permet de répondre au principe du double regard[9]. Le système de contrôle est décomposé en trois niveaux de défense : les deux premiers niveaux sont réalisés en continu, tandis que le troisième niveau est réalisé périodiquement. Toutefois, le troisième niveau de contrôle aura vocation à s'inscrire davantage dans un contrôle continu. En effet, il apparait essentiel d'embarquer les fonctions d'audit lors de la phase de validation du modèle.

La définition de chaque niveau est donnée par l'ACPR dans son document de gouvernance des systèmes IA. Le schéma suivant illustre l'articulation des différentes fonctions du système de gouvernance au sein de notre entreprise REA.



Source : IRM – Gouvernance des risques et rôle du risk manager (Formation ERM - Jean Modry – 11/06/2021)

Le message clé porté par l'ACPR en prérequis est le suivant : une séparation claire des fonctions, garantissant la bonne conduite de la gouvernance, doit s'étendre à la gestion des algorithmes IA au sein de tous les processus métiers (la souscription et le pricing dans notre étude de cas REA) qui requièrent l'intelligence artificielle pour la continuité de leurs opérations.

Organisation des équipes de niveau 2 et de niveau 3

Les équipes de contrôle périodique (audit interne) et de contrôle permanent (contrôle interne et fonction gestion des risques), ayant en charge d'effectuer des vérifications préalables sur des processus métiers impliquant des algorithmes d'IA, doivent impérativement se doter de profils d'experts à la fois techniques et métiers. En effet, le contrôle du comportement de ces algorithmes complexes, requiert des experts à double compétence pour mener une gestion appropriée des risques de conformité induits ou renforcés

[9] On se réfère ici à l'arrêté du 3 novembre 2014 pour les institutions financières.

par une approche basée sur le ML. Ils ont également pour responsabilité d'effectuer des validations croisées, à la fois techniques et fonctionnelles, tout au long du projet, de la conception à la mise en production en assurant une surveillance continue et ponctuelle approfondie.

Les compétences (techniques et expertises) en matière d'intelligence artificielle ne doivent pas se limiter aux équipes d'opérationnels. A l'instar, des organismes de supervision[10], notamment l'ACPR qui a créé un pôle d'expertise dédié à l'IA, pour accomplir l'exercice de ses propres missions.

Une solution envisagée dans notre étude cas est de créer une « *task force* » pour organiser et planifier, en collaboration avec les services de ressources humaines, un programme pour faire monter en compétence les équipes de contrôle. Il s'agit de mettre en place une politique de recrutement et de rétention des talents, inciter les collaborateurs à suivre des formations techniques (comme celles dispensées par l'Institut des Actuaires et l'Institut du Risk Management en Data science), ou encore mettre en place des programmes d'immersion dans les équipes.

Les fonctions de niveau 2, telles que prévues par le code des assurances (article L354-1 R336-1 du code des assurances, cf. annexe J) représentent le premier maillon de la chaine de contrôle impacté par l'intégration de l'IA dans les processus.

Organisation des contrôles et des validations de niveau 2

Il s'agit de définir et de mettre place un processus de validation impliquant les équipes techniques (concevant et validant les modèles) ainsi que les services conformité et risque.

Organisation des contrôles réalisés par l'audit de niveau 3

L'ACPR suggère l'adoption d'une approche à deux volets pour les missions d'audit interne des systèmes basés sur l'IA en finance :

Volet analytique	Volet empirique
Analyses et revues du code logiciel, des données utilisées, de la méthodologie de documentation (si possible standardisée) et des algorithmes et modèles prédictifs.	Utilisation de méthodes explicatives adaptées à l'IA, permettant d'expérimenter des algorithmes en boîte noire, et de justifier des décisions individuelles et/ou le comportement général de l'algorithme. Ces méthodes emploient des données d'évaluation dites de benchmarking, mettent en concurrence le modèle étudié par un modèle dit « challenger » conçu par l'auditeur.

Cette approche fait face à des défis particuliers, en raison de l'étendue du périmètre de la mission d'audit. D'où la nécessité d'avoir des auditeurs avec d'une expertise théorique et pratique en *data science*.

4.3.2 PROCEDURES OPERATIONNELLES

Les procédures opérationnelles doivent être formalisées, adaptées aux différentes activités et régulièrement actualisées, sous forme par exemple de manuels de procédures. Ces procédures servent plusieurs objectifs :

1. Déterminer les différents niveaux de responsabilités, les attributions dévolues et les moyens affectés au fonctionnement des dispositifs de contrôle interne.
2. Décrire les systèmes de mesure, de limitation et de surveillance des risques et le mode d'organisation du dispositif de contrôle de la conformité.
3. Décrire les règles relatives à la sécurité des systèmes d'information et de communication et aux plans d'urgence et de poursuite de l'activité.

[10] A titre d'exemple : le programme de recrutement et de formation des personnels, centré sur l'analyse des données, mis en place par la *Monetary Authority of Singapore*.

Un projet transverse est mis en place pour éditer et rédiger les procédures de gouvernance après intégration des modèles IA dans les processus de REA. En particulier, nous nous concentrons sur la définition des responsabilités (décrite en annexe K) et sur la comitologie.

La comitologie pour la gouvernance des modèles IA repose sur deux principaux comités : le comité stratégique et le comité de pilotage. Le premier a pour objectif de définir la stratégie tout en respectant un cadre d'appétit aux risques, et de prendre les décisions de manière collégiale par les responsables. Le second a pour objectif de suivre le projet et son avancement par les opérationnels y contribuant.

Par ailleurs, il convient également de faire levier avec les comités existants de l'entreprise, notamment le comité des risques et le comité de *data privacy* déjà en place dans la compagnie. Il s'agit d'éviter de multiplier les comités pour une efficacité optimale. Le tableau suivant illustre une proposition de comitologie adaptée à l'importance du projet (taille x criticité) que nous entreprenons dans le cadre cette étude :

	Comité stratégique	Comité de pilotage
Fréquence	Trimestrielle	Mensuelle
Objet	**Définit la stratégie** et **prend les décisions**, comprenant, mais sans s'y limiter : • La validation du budget pour le projet. • La définition des KPI de succès du projet. • La validation du modèle. • Le suivi des principaux KRI (se situant en zone rouge de la *heatmap*) et de leur plan de mitigation. • L'arrêt du modèle. • La validation de modifications significatives du modèle.	**Réalise le suivi des opérations** du projet et l'avancement du modèle, comprenant, mais sans s'y limiter : • L'état d'avancement du modèle. • Le suivi des KRI des risques du projet. • L'identification des dérives du modèles (dépassement des seuils d'alerte) • La mise à jour du modèle si celle-ci n'engendre pas de modification significative ni de surcout de développement. Une escalade au comité stratégique est automatiquement effectuée en cas de : • De matérialisation de risque relatif au légal ou à la protection des données. • De dépassement des seuils d'alerte.
Membres obligatoires	• Président : sponsor du projet (membre du comité exécutif) • Chef de projet • Responsable projet technique (*product owner*) • Responsable de la gestion des risques • Responsable actuariat • *Data Privacy Officer* (DPO) • Responsable commercial	• Chef de projet • Responsable projet technique (*product owner*) • Membre de l'équipe Data Science • Membre de l'équipe actuariat • Membre de l'IT • *Data Privacy Officer* (DPO)
Membres optionnels (en fonction de l'ordre du jour)	• Responsable de l'audit • Responsable Data Science • Responsable IT	• Membre de la gestion des risques
Documentation	Les minutes sont rédigées par le chef de projet et sont diffusées aux participants.	Les minutes sont rédigées par le chef de projet et sont diffusées aux participants.

Le comité de pilotage permet la préparation du comité stratégique en rédigeant notamment une synthèse des risques clés. Il identifie également les points bloquants pour lesquels des décisions doivent être prises en comité stratégique.

5. STRESS TESTS ET FALL BACK PLAN

Dans cette section, nous simulons un scénario de stress avec la mise en place d'un plan de remédiation (*fall back plan*). Le mécanisme de secours doit être spécifique à la remédiation d'un incident, d'un dysfonctionnement majeur ou d'une panne d'un composant d'IA, allant jusqu'aux mécanismes de continuité d'activité. La complexité, la sophistication de ce mécanisme dépendra de la manière avec laquelle le modèle de Machine Learning aura été conçu et intégré aux process. Si l'intégration dans les processus initiaux s'est faite de façon suffisamment modulaire et robuste, le plan de remédiation peut simplement consister à revenir au processus initial le temps de réparer la défaillance. A l'inverse, si

l'introduction des composantes ML a plus largement modifié la chaîne de traitements, le plan de secours sera nécessairement plus sophistiqué (et souvent plus complexe à mettre en œuvre car il doit lui-même avoir été validé).

Description du scénario de stress considéré : notre prestataire mettant à disposition la plateforme *cloud* vient de nous informer qu'il s'est fait pirater. Selon lui, les attaquants auraient été présents dans leurs systèmes depuis plusieurs mois et viennent de rendre publiques des bases de données. Celles-ci correspondent aux données collectées par les montres connectées des assurés ayant transité par la plateforme *cloud*. De plus, il semblerait que des fichiers infectés provenant du prestataire *cloud* nous aient été transférés dans le cadre du protocole d'échange des données.

Plan d'action à court terme

- Isoler et arrêter les systèmes critiques en relation avec la plateforme *cloud* (déjà identifiés au préalable).
- Déclencher le plan de continuité de l'activité :
  - Mettre en place une cellule de crise impliquant la direction des systèmes informatiques, le *data privacy officer*, la direction juridique, la direction communication, la direction générale, la direction commerciale et la direction du risque.
  - Réaliser une analyse antivirale approfondie des systèmes d'information et restaurer une sauvegarde de secours saine.
  - Déployer une version dégradée du système (en utilisant un autre prestataire *cloud* identifié au préalable, ou en ayant une solution interne de secours à disposition).
- Signaler la cyberattaque à la CNIL dans un délai de 72h maximum, déclarer l'incident aux autorités de contrôle, et déposer plainte contre les auteurs du piratage.
- Mettre en place une adresse de contact qui sera mise à disposition des assurés.
- Préparer une communication, interne à destination des employés, et externe à destination des assurés concernés par la fuite de données.

Plan d'action à moyen terme

- Surveiller quotidiennement le nombre de contacts des assurés (messages, appels téléphoniques), les messages publiés sur les forums et les articles nous mentionnant.
- Surveiller les plaintes déposées à notre encontre et étudier des solutions de dédommagement.
- Réaliser un audit chez le prestataire pour identifier les défaillances ayant permis le piratage.
- Revoir la procédure de sélection des prestataires, les protocoles de stockage et d'échange de données, ainsi que la politique d'audit de nos systèmes d'information.
- Réaliser des campagnes de communication sur les actions menées afin de rebâtir la confiance auprès des assurés.
- Intensifier les campagnes contre le *fishing* à destination des collaborateurs en interne.
- Etudier la possibilité de souscrire une assurance cyber contre ce type d'événements.
- Programmer des exercices de *bunkering* avec les dirigeants afin d'améliorer leur maîtrise de gestion de crise et de réduire les délais de réaction.

6. CONCLUSION

Les bouleversements technologiques liés au développement de l'intelligence artificielle, au déploiement de plateformes *cloud* et d'objets connectés, ont fait émerger de nouvelles opportunités de développement pour le secteur assurantiel. La crise sanitaire a notamment révélé une accélération de la digitalisation dans toutes les industries. En conséquence, les développements technologiques entraînent un décalage positif de la frontière d'efficience de l'industrie, permettant de produire davantage de réassurance avec les mêmes ressources en capital et en travail. Les acteurs de l'assurance doivent saisir cette opportunité, d'une part en interne, en faisant levier sur leurs données collectées, et d'autre part en externe, en nouant des alliances stratégiques avec d'autres acteurs.

L'industrie de l'assurance se situe donc à un tournant majeur, dont l'action ou l'inaction de ses acteurs sera déterminante pour la pérennité de leur modèle économique et de leurs activités.

Notre étude de cas REA a évalué l'impact de l'utilisation d'un modèle d'intelligence artificielle pour la tarification et l'aide à la souscription. Il s'agit d'un cas d'usage parmi un éventail de possibilités plus large, pour lequel l'approche ERM s'est révélée particulièrement adaptée. En effet, celle-ci assure le pilotage de notre projet, stratégique pour l'entreprise, par la gestion des risques. En particulier, la démarche ERM a permis de définir les objectifs de l'entreprise et son appétence au risque, d'identifier les risques et de les mesurer (cartographie des risques et KRI), de gérer et de traiter ces risques, et enfin de les surveiller et de les suivre dans le temps (politique de gestion des risques et mise en place d'une gouvernance). Bien que notre étude se soit concentrée sur un outil, l'approche ERM décrite reste transposable et généralisable à d'autres outils faisant intervenir des composantes d'IA.

La gestion des risques par l'approche ERM constitue un facteur clé pour garantir la conformité des modèles. En effet, l'intégration des modèles IA dans les processus métiers doit s'accompagner par la mise en œuvre de mesures de gouvernance proportionnées et de process adéquats. Le cadre de gouvernance doit être robuste pour faire face aux risques émergents, notamment les risques de cyber attaque et les risques de divulgation de données à caractère privé. La surveillance tout au long du cycle de vie des systèmes d'IA doit être définie et mesurée, notamment par l'attribution des rôles, la définition claire des responsabilités des diverses parties prenantes (opérationnelles et de contrôles), et leur formation continue. Les fonctions de contrôle jouent un rôle décisif dans la maîtrise et dans le pilotage des risques liés aux algorithmes d'IA. Elles doivent se doter de compétences duales dans les domaines de la *data science* et des techniques d'IA ainsi que dans la gouvernance et la sécurité de ces systèmes. Les modèles d'IA doivent être régulièrement contrôlés et évalués selon des principes bien définis. En particulier, ceux-ci doivent être explicables, autonomes, stables, performants, et assurer un traitement adéquat des données. Les modèles doivent disposer d'un niveau d'autonomie suffisant leur permettant d'assurer la continuité de l'activité en cas de défaillance d'un tiers. Il est donc nécessaire d'anticiper et de prévenir une dépendance vitale de l'algorithme à une source externe.

Le traitement et la gestion des données sont essentiels et constituent un prérequis au bon fonctionnement et à la pertinence des modèles. Des efforts particuliers doivent être fournis pour structurer les données et en faciliter leur usage. Par ailleurs, il est important d'œuvrer à l'élimination des biais dans les données de (ré)apprentissage, de conserver rigoureusement une historisation des différentes versions (*versioning*) et de documenter les méthodologies de traitement des données.

En complément de notre démarche ERM, nous avons challengé celle-ci avec l'approche proposée par l'ACPR dans une perspective de veille réglementaire. Plus globalement, d'autres organismes émettent des recommandations sur la maîtrise des algorithmes d'IA et sur les problématiques d'éthique et d'équité. A titre d'exemple, la CNIL veut s'emparer du sujet du contrôle des algorithmes IA dans la continuité du RGPD, et rappelle les principes de transparence, de compréhension et d'explicabilité des modèles IA identifiés lors de notre démarche ERM. De son côté, l'EIOPA a notamment publié un rapport sur les principes de gouvernance de l'IA, et y traite les problématiques d'éthique et de confiance.

Doter les algorithmes d'IA d'une base éthique solide, au travers d'une charte d'éthique ou d'un comité d'éthique à court terme, constituera un élément central au fur et à mesure que ces algorithmes auront la charge d'opérations critiques. Dans ce cadre, il sera important de traiter ce risque d'éthique en conservant l'approche ERM afin de permettre l'exécution de la stratégie et de l'atteinte des objectifs de notre entreprise.

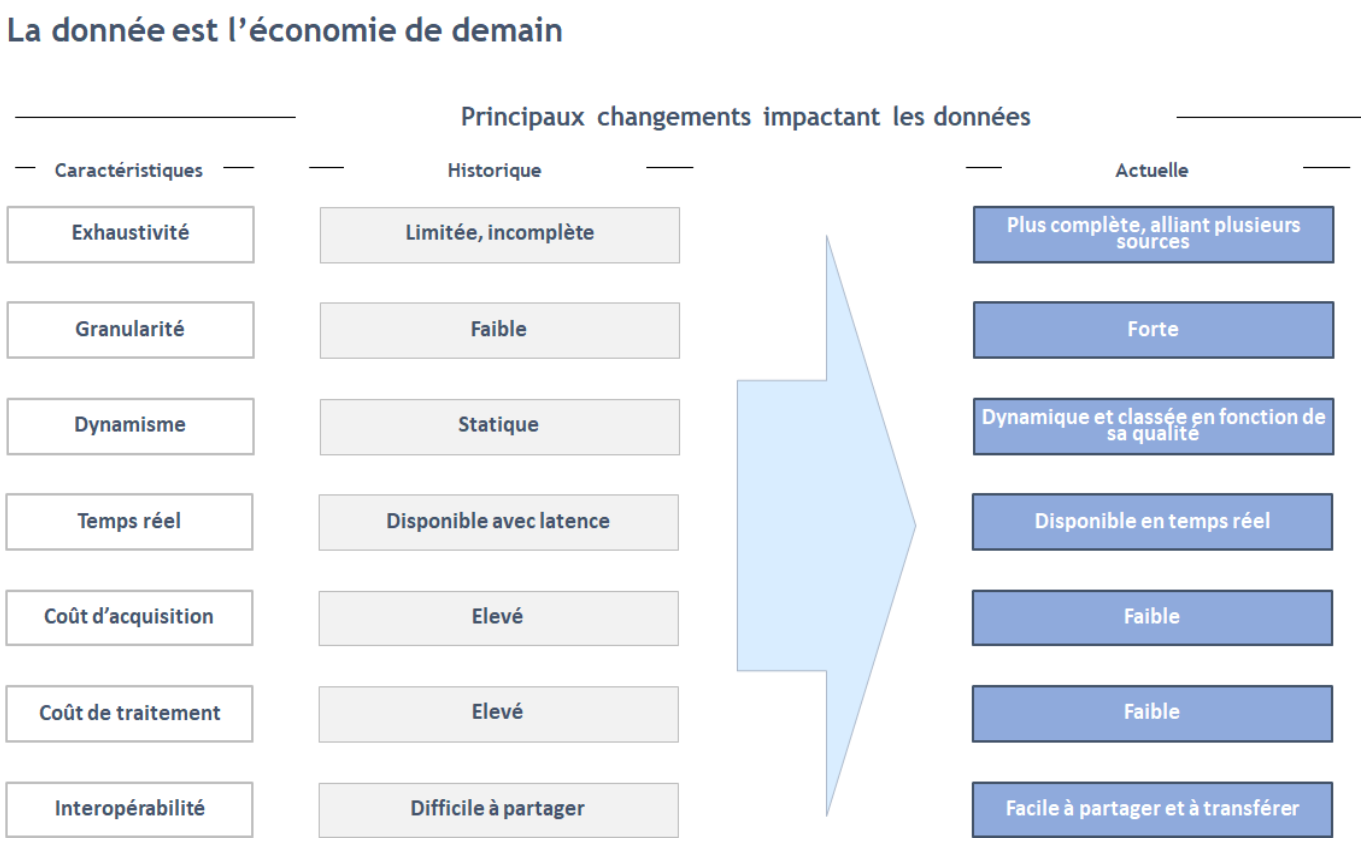
BIBLIOGRAPHIE

- **« Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier »**, Document de réflexion ACPR, juin 2020. https://acpr.banque-france.fr/gouvernance-des-algorithmes-dintelligence-artificielle-dans-le-secteur-financier
- **« L'auditeur interne et les algorithmes d'IA »**, IFACI, mars 2021. https://docs.ifaci.com/wp-content/uploads/2021/03/Lauditeur-interne-et-les-algorithmes-dIA.pdf
- **« Artificial intelligence governance principles: towards ethical and trustworthy artificial intelligence in the European insurance sector »**, EIOPA, juin 2021. https://www.eiopa.europa.eu/sites/default/files/publications/reports/eiopa-ai-governance-principles-june-2021.pdf
- **« SCOR's new strategic plan Quantum Leap 2019/2021 »**, SCOR Investor Day, 2019. https://www.scor.com/en/press-release/scor-launches-its-new-strategic-plan-quantum-leap
- **« Biological Age Model (BAM) – Using wearable data to empower healthier lives »**, SCOR https://www.scor.com/en/biological-age-model-bam
- **« Algo & risque »**, David Dubois et Voahirana Ranaivozanany, Journée de Deauville à Paris, septembre 2021.
- **« Intelligence artificielle : quelles évolutions pour les profils de risques des entreprises ? »**, Deloitte, février 2019. https://www2.deloitte.com/fr/fr/pages/risque-compliance-et-controle-interne/articles/intelligence-artificielle-quelles-evolutions-pour-profils-de-risques-des-entreprises.html
- **« Algorithmes : prévenir l'automatisation des discriminations »**, Défenseur des droits, mai 2020. https://www.defenseurdesdroits.fr/fr/rapports/2020/05/algorithmes-prevenir-lautomatisation-des-discriminations
- **« IRM – Gouvernance des risques et rôle du risk manager »**, Jean Modry, CRO/LCO Hannover Re, Formation ERM, juin 2021.
- **« Le risque de modèle – mise en pratique dans un contexte de réassurance vie »**, Gurvan Le Rhun et Elsa Renouf, Mémoire ERM, 2013.

Articles

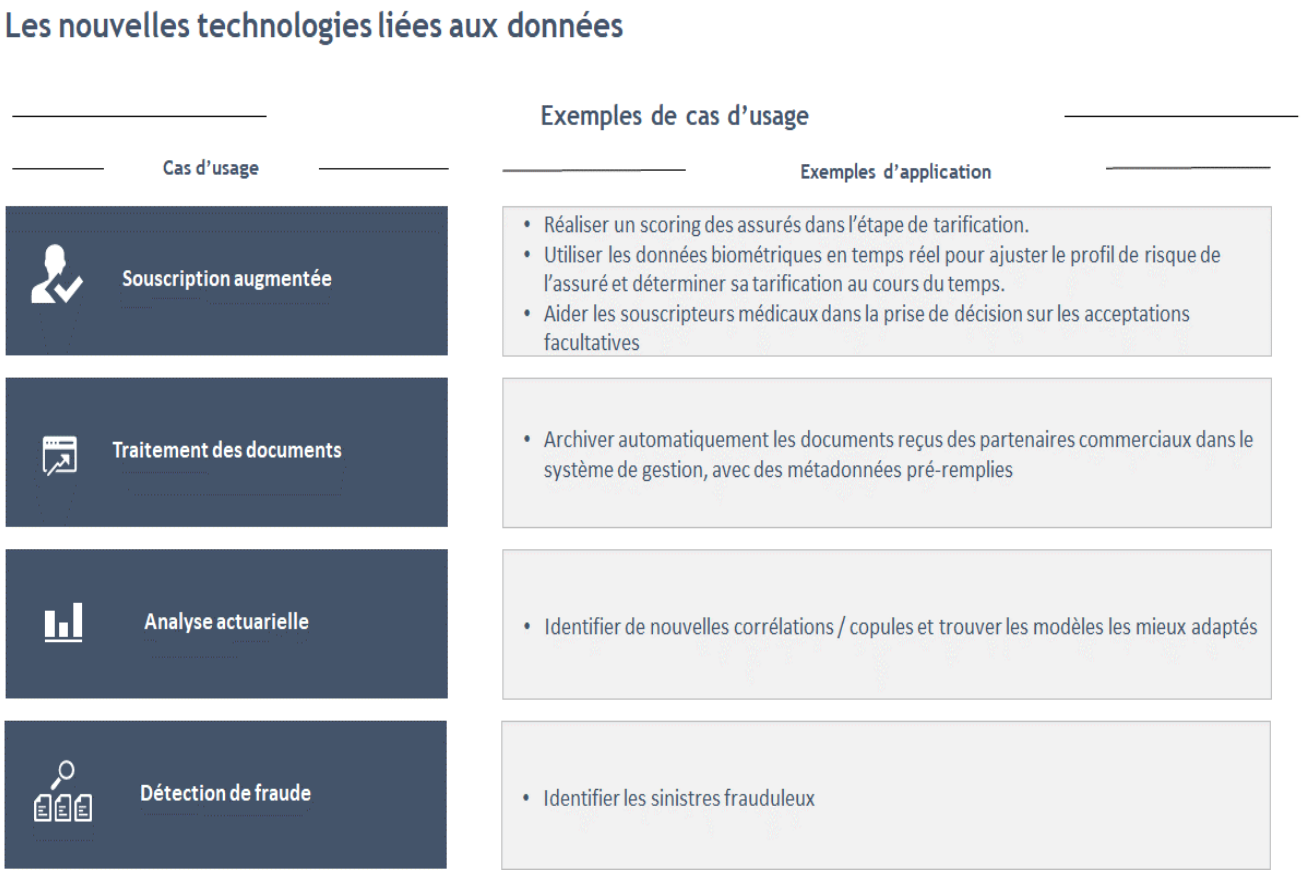
- **« La Commission européenne veut placer l'IA sous contrôle des autorités »**, la revue du digital, février 2020. https://www.larevuedudigital.com/la-commission-europeenne-veut-placer-lia-sous-controle-des-autorites/
- **« La mise en conformité sur l'IA pourrait fortement pénaliser les entreprises européennes »**, la revue du digital, juillet 2021. https://www.larevuedudigital.com/la-mise-en-conformite-sur-lia-pourrait-largement-plomber-les-entreprises-europeennes/
- **« La Cnil tient à contrôler elle-même l'intelligence artificielle »**, la revue du digital, septembre 2021. https://www.larevuedudigital.com/la-cnil-tient-a-controler-elle-meme-lintelligence-artificielle/
- **« Décryptage de DORA : qu'est-ce que cela signifie pour la résilience des organisations financières ? »**, RiskInsight, janvier 2021. https://www.riskinsight-wavestone.com/2021/01/decryptage-de-dora-quest-ce-que-cela-signifie-pour-la-resilience-des-organisations-financieres/

ANNEXE A – LA DONNEE EST L'ECONOMIE DE DEMAIN



Source : SCOR's new strategic plan Quantum Leap 2019-2021 (SCOR's Investor Day – September 4, 2019)

ANNEXE B – CAS D'USAGE DE MODELES IA



Source : SCOR's new strategic plan Quantum Leap 2019-2021 (SCOR's Investor Day – September 4, 2019)

ANNEXE C – DETAIL DE FONCTIONNEMENT DE LA MONTRE CONNECTEE

- Si l'assuré ne porte plus sa montre connectée, alors celui-ci ne se verra plus accorder de réduction tarifaire.
- Si l'assuré ne transmet pas ses données à l'assureur (absence de synchronisation de la montre avec l'application avant la fin du mois par exemple), alors celui-ci ne se verra pas accorder de réduction tarifaire.
- Si l'assuré génère un mauvais score, mais que celui-ci continue de transmettre ses données, alors sa réduction tarifaire sera moins importante mais sera non-nulle. En effet, l'objectif est d'inciter les assurés à porter leur montre et de les encourager à améliorer leur rythme de vie.

ANNEXE D – CLASSIFICATION DES RISQUES SELON LA NORME IFACI

Nous proposons ci-dessous une classification des risques précédemment selon la norme d'audit IFACI. Celle-ci est donnée à titre illustratif et dans une perspective de *best practice* lors d'audits de modèles de *machine learning*.

Risque	Famille	Niveau 2	Niveau 3
Risque de disponibilité de la donnée	(R3) Risques opérationnels	(R303) Dysfonctionnements de l'activité et des systèmes	(R303-11) Systèmes - Données
Risque de sécurité de la donnée	(R3) Risques opérationnels	(R307) Fraude externe	(R307-06) Sécurité des systèmes - Malveillance informatique
			(R307-05) Usurpation de compte / d'identité
		(R303) Dysfonctionnements de l'activité et des systèmes	(R303-11) Systèmes - Données
Risque de compliance de la donnée	(R3) Risques opérationnels	(R301) Clients / tiers, produits et pratiques commerciales	(R301-03) Conformité, diffusion d'informations et devoir fiduciaire - Protection des données personnelles
Risque de traitement de la donnée	(R3) Risques opérationnels	(R302) Exécution, livraison et gestion des processus	(R302-01) Saisie, exécution et suivi des transactions - Erreur
Risque de modélisation	(R3) Risques opérationnels	(R303) Dysfonctionnements de l'activité et des systèmes	(R303-11) Systèmes - Recette
			(R303-10) Systèmes – Pérennité
Risque d'explicabilité du modèle	(R3) Risques opérationnels	(R302) Exécution, livraison et gestion des processus	(R302-09) Saisie, exécution et suivi des transactions - Piste d'audit
Risque de biais	(R3) Risques opérationnels	(R302) Exécution, livraison et gestion des processus	(R302-07) Saisie, exécution et suivi des transactions - Paramétrage
Risque de mise à jour de l'algorithme	(R3) Risques opérationnels	(R302) Exécution, livraison et gestion des processus	(R302-04) Risques d'interface inter-services
			(R302-09) Saisie, exécution et suivi des transactions - Piste d'audit
		(R303) Dysfonctionnements de l'activité et des systèmes	(R303-09) Systèmes – Régression
			(R303-02) Systèmes - Développement
Risque d'utilisation de code externe	(R3) Risques opérationnels	(R303) Dysfonctionnements de l'activité et des systèmes	(R303-05) Systèmes - Disponibilité des systèmes
Risque de tarification	(R2) Risques assurances	(R201) Technique	
Risque de souscription	(R2) Risques assurances	(R202) Souscription	
Risque de mortalité	(R2) Risques assurances	(R203) Sinistralité non-vie / Prestations vie	
Risque de disponibilité des systèmes	(R3) Risques opérationnels	(R303) Dysfonctionnements de l'activité et des systèmes	(R303-05) Systèmes - Disponibilité des systèmes
Risque humain	(R3) Risques opérationnels	(R304) Pratiques en matière d'emploi et de sécurité sur le lieu de travail	
Risque cyber	(R3) Risques opérationnels	(R307) Sécurité des systèmes – Données	(R307-07) Vol et divulgation de données
			(R307-06) Sécurité des systèmes - Malveillance informatique
Risque de concurrence	(R4) Risque stratégiques et environnementaux	(R401) Marché de l'assurance	
Risque législatif et réglementaire	(R4) Risque stratégiques et environnementaux	(R406) Législatifs, réglementaires et judiciaires	
Risque de réputation	(R4) Risque stratégiques et environnementaux	(R405) Réputation	

ANNEXE E – GOUVERNANCE SELON LE REGULATEUR

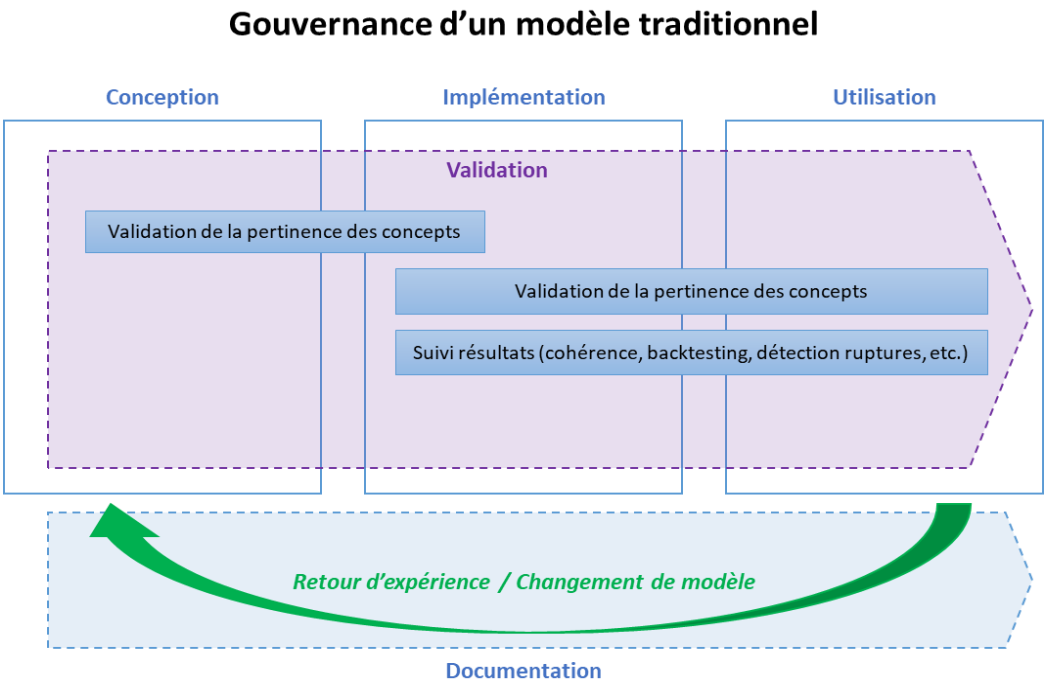
Les annexes du document de l'ACPR constituent un référentiel et une base d'informations sur certains sujets techniques et fonctionnels. Elles contiennent en particulier, les retours d'expérience issus des ateliers et des entretiens menés avec certains acteurs du marché.

Le point d'attention principal des travaux exploratoires de l'ACPR porte sur trois thèmes :

- Thème 1 : Lutte contre le blanchiment et le financement du terrorisme (LCB-FT)
- Thème 2 : Modèles internes en banque et assurance La question centrale sur ce thème était d'étudier comment et à quelles conditions l'IA peut être utilisée dans les modèles internes.
- Thème 3 : Protection de la clientèle

A titre d'exemple, « *en assurance, le processus de vente de contrats d'assurance est soumis à une réglementation propre, impliquant entre autres un devoir de conseil et des exigences de motivation personnalisée le cas échéant. À l'inverse, la segmentation de la clientèle ex ante en assurance repose essentiellement sur des objectifs d'efficacité, sans la même exigence d'explicabilité associée.* »

ANNEXE F – SCHEMA DU CYCLE DE VIE SIMPLIFIE D'UN MODELE TRADITIONNEL



Source : Le risque de modèle – Mise en pratique dans u ncontexte de réassurance vie (Mémoire ERM G. Le Rhun & E. Renouf – 2013)

ANNEXE G – LE TRAITEMENT ADEQUAT DES DONNEES

Conformité réglementaire

- La conformité aux réglementations relatives à la protection de la vie privée ou des données personnelles, à commencer par le RGPD ;
- La conformité relative à la Réglementation Sectorielle - l'ACPR cite un cas d'usage relatif au secteur assurantiel : *« Par exemple dans le domaine de l'assurance, l'interdiction d'orienter le processus de vente en fonction de la capacité à payer : l'offre doit au moins être cohérente avec les exigences et besoins du client, et non dictée par une possibilité d'optimisation du chiffre de vente de produits d'assurance* ». En effet, La DDA (Directive sur la distribution d'assurances) impose des principes d'équité comme ceux énoncés dans la partie suivante.

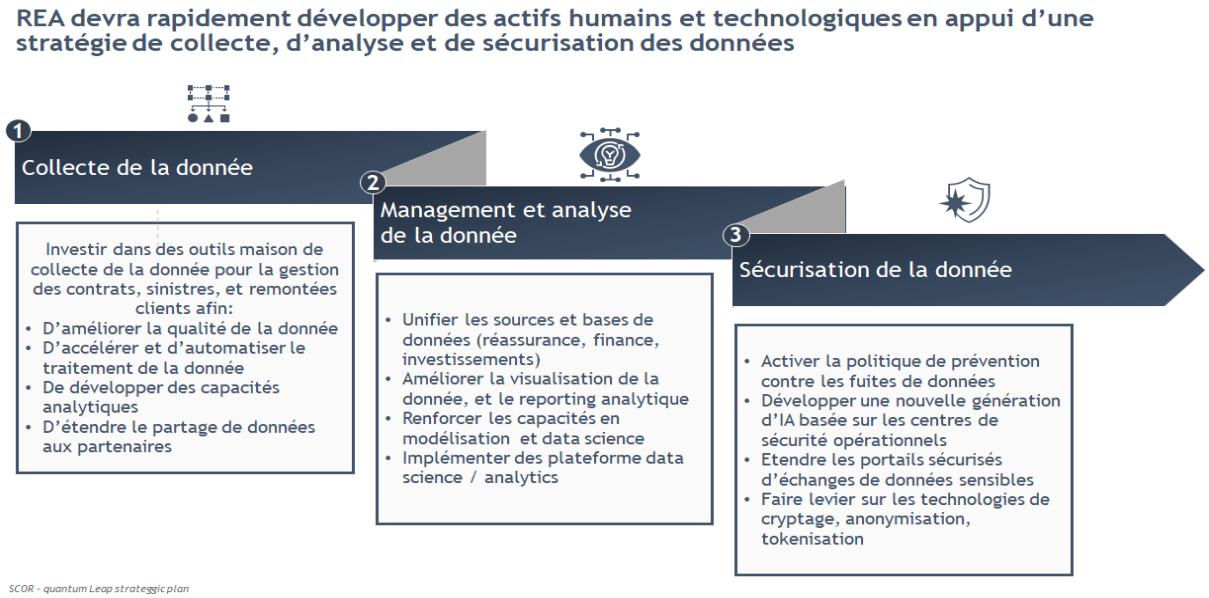
Éthique et équité

- Une illustration est donnée dans cette partie à travers les recommandations en matière d'éthique publiées par le groupe d'expertise en IA de la Commission européenne (*European Commission High-Level Expert Group on A*I, 2019) :
    1. action humaine et contrôle humain ;
    2. robustesse technique et sécurité ;
    3. respect de la vie privée et gouvernance des données ;
    4. transparence ;
    5. diversité, non-discrimination et équité ;
    6. bien-être sociétal et environnemental ;
    7. responsabilité.
- L'ACPR énonce : « *Ces recommandations montrent le large spectre des enjeux liés à l'éthique et à l'équité en IA. Or un point d'attention en matière d'équité algorithmique concerne l'étude des biais, notamment à caractère discriminatoire, qui constitue un domaine de recherche actuellement très actif. Schématiquement, il s'agit :*
    - *tout d'abord de bien définir les biais de nature problématique – qu'ils soient biais de classification ou de prédiction, ou biais statistiques non souhaités déjà présents dans les données – et les métriques permettant de caractériser et quantifier ces biais, y compris par des méthodes explicatives (Kamishima, 2012) ;*
    - *de déterminer dans quelle mesure les biais présents dans les données sont reflétés, voire renforcés, par les algorithmes d'IA ; »*
- L'ACPR précise : « *Toutefois les travaux exploratoires réalisés par l'ACPR, même complétés par une étude plus générale du secteur financier, ont montré que seuls quelques acteurs du secteur financier avaient commencé à aborder la question de la détection et remédiation des biais de modèles. L'accent est pour l'instant mis sur la validation interne des solutions ainsi que sur leur conformité réglementaire, sans pousser l'analyse de l'équité algorithmique plus loin qu'elle ne l'était avec les méthodes traditionnelles – et notamment en ignorant souvent le renforcement potentiel des biais inhérents aux données. Cela ne fait toutefois que refléter le manque relatif de maturité de l'IA, ainsi que la priorité donnée jusqu'à présent aux processus non critiques (y compris en termes de risques éthiques et d'équité) : on peut dès lors prévoir que l'incorporation croissante d'IA en finance bénéficiera de la recherche en cours sur ces sujets.* »

ANNEXE G BIS – REFLEXIONS DE REA SUR LES BONNES PRATIQUES DE TRAITEMENTS DE LA DONNEE

Réflexion sur les enjeux et impacts sur l'entreprise : l'intégration d'une stratégie misant sur un développement de ces nouvelles technologies soulève un ensemble de questions stratégiques, notamment les suivantes :

Collecte	• Quelles données collecter (choix stratégique pour ne pas être submergé) ? • Via quels capteurs / réseaux d'acquisition, notamment lorsqu'on est loin du client final ? • Comment s'assurer de leur fiabilité, qualité, résilience dès les phases amont ?
Traitement	• Comment rendre la donnée exploitable de manière efficace, donc utile au métier, et efficiente, avec une meilleure automatisation des process intermédiaires ?
Exploitation	• Comment utiliser les technologies de data science ainsi que ces nouvelles données collectées pour mieux évaluer certains risques ? • Etendre la couverture à de nouveaux segments de clients et / ou arriver à couvrir de nouvelles typologies de risque ? • Ajuster le pricing de certains contrats ? • Améliorer la satisfaction client (ex : meilleurs niveaux de services) ? • Améliorer la performance et rentabilité globale du portefeuille.
Visualisation / communication	• Comment s'assurer d'un meilleur partage d'information au sein et en dehors de l'entreprise ? • Comment diffuser une culture « *data-driven* » de la gestion du risque au sein de l'entreprise, à ses clients et parties prenantes ? • Comment faire levier sur cette communauté informée pour mieux gérer certains risques
Résilience et autonomie	• Comment s'assurer du bon déroulé de l'ensemble des opérations du cycle de vie de manière « industrielle / autonome » (ex : Process et gouvernance) ? • Comment garantir la gestion des risques contre des risques qui lui sont spécifiques : Sécurité des données, dispositifs la servant (ex : Cybersécurité) ? • Comment s'assurer de la fiabilité, de l'interprétabilité de l'IA en mode « run »?



Source : SCOR's new strategic plan Quantum Leap 2019-2021 (SCOR's Investor Day – September 4, 2019)

ANNEXE H – PRINCIPE DE STABILITE

Dérives temporelles : la dérive temporelle est souvent liée à la base de données d'apprentissage, note le régulateur, et peut être détectée au moyen de méthodes de *monitoring* et de lancement d'alertes (mise en place d'un outil de détection en priorité).

Généralisations : le modèle doit pouvoir être généralisable. Cette caractéristique doit-être vérifiée dès les premières phases de conception et de paramétrage. Les défauts en la matière peuvent-être détectés notamment lors de la phase de validation (via du *out-of-time testing, out-of-distribution testing*). Il reste néanmoins impératif de soumettre l'algorithme à un *monitoring* continu durant ses phases de production.

Réapprentissage : périodique ou quasi-permanent, il peut conduire à l'impossibilité de reproduire des décisions prises initialement par l'algorithme. Dans ce cas, le régulateur attribue un défaut de déterminisme au système et dans ce cas-là, la décision ne peut être reproduite et il est impossible d'être conforme aux exigences réglementaires telles que les droits à l'information et à l'opposition prévus par le RGPD et aucune explication ne peut être donnée le cas échéant. L'archivage et l'historisation de l'ensemble des versions évolutives du modèle d'IA en production peut-être dans ce cas-là une réponse adéquate pour pallier ce défaut.

ANNEXE I – PRINCIPE D'EXPLICABILITE

Zoom sur la démarche d'évaluation du principe d'explicabilité selon l'ACPR – l'explication algorithmique vise généralement à répondre aux questions suivantes :

- Quelles sont les causes d'une décision ou prédiction donnée ?
- Quelle est l'incertitude inhérente au modèle ?
- L'algorithme fait-il les mêmes erreurs que l'humain ?
- Au-delà de la prédiction du modèle, quelle autre information est utile (par exemple pour assister l'humain dans la prise de décision finale) ?

Les **objectifs** sont donc multiples, car dépendants des parties prenantes :

- Rassurer les experts métiers et les équipes en charge de la conformité ;
- Faciliter la validation du modèle par les équipes de conception et de validation ;
- Garantir la confiance des individus impactés par les décisions ou prédictions de l'algorithme.

Caractérisation : une explication idéale posséderait les qualités suivantes :

- **Précise** : elle décrit aussi précisément que possible le cas considéré (pour une explication locale) et le fonctionnement exact de l'algorithme (qu'elle soit locale ou globale) ;
- **Complète** : elle couvre l'ensemble des motifs et caractéristiques de la ou des prédictions en question ; - compréhensible : elle ne nécessite pas d'effort exorbitant pour être correctement comprise par l'audience à qui elle est destinée ;
- **Succincte** : elle est assez concise pour être assimilée en un temps raisonnable, en fonction des contraintes de temps ou de productivité du processus où elle s'inscrit ;
- **Actionnable** : elle permet une ou plusieurs actions de la part d'un humain, par exemple infirmer la prédiction en question ;
- **Robuste** : elle demeure valable et utile lorsque les données sont changeantes et bruitées ;
- **Réutilisable** : elle peut être personnalisée selon le type d'audience. Bien entendu, certains de ces objectifs sont dans la pratique souvent mutuellement irréconciliables. En outre et comme détaillé par la suite, ils devront être mis en balance avec les autres principes – notamment celui de performance. Aussi ces objectifs serviront plutôt de critères de comparaison entre les explications fournies par différentes méthodes afin de choisir la méthode la plus appropriée à un cas d'usage bien spécifique.

ANNEXE J – ARTICLES L354-1 ET R336-1 DU CODE DES ASSURANCES

Article L354-1 du Code des assurances

« Les entreprises d'assurance et de réassurance mettent en place un système de gouvernance garantissant une gestion saine et prudente de leur activité et faisant l'objet d'un réexamen interne régulier. Ce système de gouvernance repose sur une séparation claire des responsabilités et comprend un dispositif efficace de transmission des informations. Il est proportionné à la nature, à l'ampleur et à la complexité des opérations de l'entreprise.

Ce système de gouvernance comprend les fonctions clés suivantes : la fonction de gestion des risques, la fonction de vérification de la conformité, la fonction d'audit interne et la fonction actuarielle.

Les entreprises élaborent des politiques écrites relatives, au moins, à la gestion des risques, au contrôle interne, à l'audit interne et, le cas échéant, à l'externalisation mentionnée à l'article *L. 310-3*. Elles veillent à ce que ces politiques soient mises en œuvre.

Les entreprises prennent des dispositions permettant d'assurer la continuité et la régularité dans l'exercice de leurs activités, ce qui inclut l'élaboration de plans d'urgence. Elles mettent en œuvre, à cette fin, des dispositifs, des ressources et des procédures appropriés et proportionnés.

Un décret en Conseil d'Etat précise les conditions d'application du présent article. »

Article R336-1 du Code des assurances

« Les entreprises mentionnées à l'article *L. 310-3-2* sont tenues de mettre en place un dispositif permanent de contrôle interne.

Le conseil d'administration ou le conseil de surveillance approuve, au moins une fois par an, un rapport sur le contrôle interne, qui est transmis à l'Autorité de contrôle prudentiel et de résolution.

La première partie de ce rapport détaille les conditions de préparation et d'organisation des travaux du conseil d'administration ou du conseil de surveillance et, le cas échéant, les limitations apportées par le conseil d'administration aux pouvoirs du directeur général dans l'exercice de ses fonctions.

Toutefois, les entreprises dont les titres financiers sont admis aux négociations sur un marché réglementé ne sont pas tenues de fournir ces éléments lorsqu'elles transmettent à l'Autorité de contrôle prudentiel et de résolution le rapport mentionné, selon les cas, à l'article *L. 225-37* ou à l'article *L. 225-68* du code de commerce.

La seconde partie de ce rapport détaille :

a) Les objectifs, la méthodologie, la position et l'organisation générale du contrôle interne au sein de l'entreprise, les mesures prises pour assurer l'indépendance et l'efficacité du contrôle interne et notamment la compétence et l'expérience des équipes chargées de le mettre en œuvre, ainsi que les suites données aux recommandations des personnes ou instances chargées du contrôle interne ;

b) Les procédures permettant de vérifier, d'une part, que les activités de l'entreprise sont menées selon les politiques et stratégies établies par les organes dirigeants, d'autre part, la conformité des opérations d'assurance ou de réassurance aux dispositions législatives et réglementaires ;

c) Les méthodes utilisées pour assurer la mesure, l'évaluation et le contrôle des placements, concernant en particulier l'évaluation de la qualité des actifs et de la gestion actif-passif, le suivi des opérations sur instruments financiers à terme et l'appréciation des performances et des marges des intermédiaires financiers utilisés ;

d) Le dispositif interne de contrôle de la gestion des placements, ce qui inclut la répartition interne des responsabilités au sein du personnel, les personnes chargées d'effectuer les transactions ne pouvant être également chargées de leur suivi, les délégations de pouvoir, la diffusion de l'information, les procédures internes de contrôle ou d'audit ;

e) Les procédures et dispositifs permettant d'identifier, d'évaluer, de gérer et de contrôler les risques liés aux engagements de l'entreprise et de détenir des capitaux suffisants pour ces risques, ainsi que les méthodes utilisées pour vérifier la conformité des pratiques en matière d'acceptation et de tarification du risque, de cession en réassurance et de provisionnement des engagements réglementés à la politique de l'entreprise dans ces domaines, définie dans les rapports mentionnés à l'article *L. 336-1* et à l'article *R. 336-5* ;

f) Les mesures prises pour assurer le suivi de la gestion des sinistres, le suivi des filiales, la maîtrise des activités externalisées et des modes de commercialisation des produits de l'entreprise et les risques qui pourraient en résulter. »

ANNEXE K – ROLES ET RESPONSABILITES : LA MATRICE RACI

Matrice RACI

Etapes	Description	Responsables Métiers / Actuariat	Responsable Projet Technique (*Product Owner*)	DPO	**Risk Management**	Responsable Data Scientist (Expertise revue modèle & métiers)	Data Scientist	Data Engineer	Responsable Actuariat
1. Comprendre et définir	- Définir la problématique et les objectifs. - Comprendre les cas d'usage et prioriser en fonction de la compréhension des enjeux business. - Recenser les données existantes.	A	R		I	C			
2. Infrastructure et préparation de la donnée	- Collecter : identifier les différentes sources de données. - Centraliser et uniformisation du lieu de stockage. - Fiabiliser et normaliser afin d'optimiser la performance des algorithmes à modéliser et éviter les dérives et les biais.		A		I		I	R	
2.Bis *Data Privacy*	- Garantir l'intégrité de la donnée.			A					
3. Développement du modèle	- Concevoir le modèle. - Réaliser un *Minimum Viable Product* (MVP). - Améliorer le modèle de base. - Documenter et mise en place de *versioning* (algorithme & données).	I	A		I		R		
4. Validation initiale du modèle	- Réaliser une première validation (recette, phases de tests avec l'implication progressive de plusieurs niveaux d'opérationnels). - Analyser en détail les résultats. - Valider en comités et obtenir le *sign-off* pour le déploiement.		I		**A**	R			
5. Déploiement	- Mettre en production le modèle. - Analyser les résultats.		A		I			R	
5. Revue et Validation Continue	- Tester et réaliser des contrôles itératifs. - Obtenir un *sign-off* à chaque changement significatif. - Communiquer et former à chaque modification ou amélioration.		I		**A**	R			
6. Maintenance	*En continu* - Définir un plan de maintenance entre les différentes parties prenantes (Métriques : Technique, Métiers, Services, user guidelines, documentation et formations). - Définir une architecture modulable et prévoir de faciliter les audits				I			R	A
6. Monitoring en continu	- Piloter le modèle en continu afin d'éviter son obsolescence.				I			R	
7. Risk & Plan de remédiation	- Mettre-à-jour de la cartographie des risques et remettre à niveau les plans de remédiation correspondants. - Réaliser des stress tests et des mises en situation. - Effectuer une veille (réglementaire, juridique, technique).				**A**				

R: Responsible **A: Accountable** **C: Consulted** **I: Informed**