



Association
Henri Capitant

Journées internationales sud-coréennes

L'Intelligence artificielle

Rapport français

Partie 1 - L'I.A et la responsabilité civile

Rapporteur national : Fabrice LEDUC, Professeur émérite de l'université de Tours

Pour donner un contexte, la définition des “**systèmes d'IA**” utilisée dans ce questionnaire suivra celle définie par l'Article 3(1) de la Loi sur l'intelligence artificielle de l'UE récemment adoptée.

« **système d'IA**»: un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels;

Aussi, pour avoir une image plus claire en répondant aux questions, veuillez garder à l'esprit les scénarios hypothétiques suivants :

Scénario A - Erreur de diagnostic médical par IA (Préjudice corporel)

Un système de diagnostic alimenté par l'IA ne détecte pas un cancer traitable à un stade précoce sur le scan d'un patient, entraînant un retard de traitement et des complications de santé importantes.

Scénario B - Dysfonctionnement d'un système d'irrigation contrôlé par IA (Préjudice matériel)

Un système d'IA gérant la distribution d'eau dans une grande exploitation agricole dysfonctionne, inondant plusieurs champs et détruisant les cultures. Le dysfonctionnement est attribué à une erreur dans l'interprétation par l'IA des données des capteurs d'humidité du sol.

Scénario C - Erreur de jugement d'un conseiller financier IA (Préjudice économique)

Un algorithme d'investissement IA recommande une stratégie à haut risque basée sur une interprétation erronée des tendances du marché, entraînant des pertes financières substantielles pour ses clients.

Scénario D - Vidéo *deepfake* générée par IA (Préjudice moral)

Un système d'IA crée une vidéo très convaincante mais fautive d'une personne se livrant à un comportement scandaleux. La diffusion virale de la vidéo cause de graves dommages à la réputation et un traumatisme émotionnel à la personne représentée.

Scénario E - Collision de véhicule autonome

Une voiture autonome interprète mal les données des capteurs dans des conditions météorologiques inhabituelles, la faisant dévier dans la circulation en sens inverse et provoquant un accident impliquant plusieurs véhicules avec des blessés.

Vous pouvez vous référer aux acteurs suivants lors de l'élaboration de vos réponses. Selon les faits spécifiques, il peut y avoir un chevauchement dans les rôles énumérés ci-dessous :

Développeurs, qui conçoivent, programment et testent les systèmes d'IA.

Fabricants, qui produisent et commercialisent les systèmes d'IA ou les produits intégrant l'IA.

Opérateurs, qui déploient et gèrent les systèmes d'IA dans des contextes opérationnels.

Utilisateurs, qui interagissent directement avec les systèmes d'IA ou utilisent des produits basés sur l'IA.

Victimes, qui subissent un préjudice résultant des actions ou décisions d'un système d'IA.

Fournisseurs de données, qui fournissent les données utilisées pour former et alimenter les systèmes d'IA.

Propriétaires, qui possèdent légalement les systèmes d'IA.

I. QUESTIONS GENERALES SUR LES REGIMES JURIDIQUES, LES CADRES ET LES DEFINITIONS

1. Cadres juridiques existants

a) Quels cadres juridiques (comprenant des accords ou des normes internationaux) existants dans votre juridiction sont actuellement appliqués aux cas de responsabilité civile liés à l'IA ?

La question semble avoir pour objet de recenser les textes, nationaux ou internationaux, applicables en France, sur lesquels pourrait se fonder l'indemnisation des dommages causés par l'IA.

- Textes supranationaux :
 - directive européenne 85/374 du 25 juillet 1985 relative à la responsabilité du fait des produits défectueux, remplacée par la directive n° 2024/2853 du 23 octobre 2024, qui devra être transposée en droit interne avant le 9 décembre 2026 ;
 - règlement européen 2024/1689 du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle : cet instrument ne traite pas directement de la responsabilité civile du fait de l'IA mais intéresse tout de même celle-ci indirectement en ce qu'il édicte un certain nombre d'obligations à la charge des acteurs de l'IA, dont le non-respect sera de nature à entraîner la mise en œuvre de la responsabilité civile de ceux-ci.
- Code civil : plusieurs régimes de responsabilité civile figurant dans le Code civil peuvent servir à indemniser les victimes de dommages causés par l'IA :
 - articles 1231 et s. : responsabilité contractuelle ;
 - article 1240 et 1241 : responsabilité extracontractuelle pour faute ;
 - article 1242, al. 1er : responsabilité extracontractuelle du fait des choses ;
 - articles 1245 et s. : responsabilité du fait des produits défectueux (transposition en droit interne de la directive du 25 juillet 1985).
- Textes non codifiés : loi n° 85-677 du 5 juillet 1985 relative à l'indemnisation des victimes d'accidents de la circulation, qui a vocation à s'appliquer également aux accidents de la circulation dans lesquels sont impliqués des véhicules autonomes.

b) Existe-t-il des lois ou des réglementations spécifiques à l'IA en place ou en cours d'élaboration dans votre juridiction ?

La réponse est négative.

A l'échelon européen, le règlement 2024/1689 du 13 juin 2024 sur l'IA ne traite pas de l'indemnisation des dommages causés par l'IA (cf. *supra* a).

Dans une résolution du 20 octobre 2020, le Parlement européen avait proposé la

création d'un régime de responsabilité civile propre à l'IA (pour une présentation de ce régime, V. *infra* II.1 c). Cependant, ce régime n'a pas été consacré.

A l'échelon national, aucun régime de réparation propre à l'IA n'existe en droit positif ni n'est en cours d'élaboration.

En droit français, l'indemnisation des victimes de l'IA passe donc nécessairement par l'utilisation de régimes de réparation préexistants qui n'ont pas été spécialement conçus pour traiter les dommages causés par l'IA (*supra* a).

c) Comment ces cadres abordent-ils les défis uniques posés par les systèmes d'IA ?

Pour répondre à cette question, il faut commencer par identifier les défis que l'IA lance au droit de la responsabilité civile (1) avant d'évoquer la stratégie choisie par le droit français pour les relever (2).

1) Identification des défis soulevés par l'IA

1.1) Défi relevant de la politique juridique

Confronté à une technologie très innovante et en constante évolution, le droit de la responsabilité civile doit rechercher un point d'équilibre entre deux impératifs antagonistes : d'un côté, protéger suffisamment les victimes potentielles de l'IA ; de l'autre, ne pas entraver l'innovation.

1.2) Défis relevant de la technique juridique

❖ Premier défi technique : appréhender l'immatériel.

Si les systèmes d'IA « incarnés » s'incorporent dans un objet matériel (cf. robots chirurgicaux, véhicules autonomes...), d'autres, les systèmes d'IA « désincarnés », sont, quant à eux, dépourvus de tout support physique et demeurent purement immatériels (cf. agents conversationnels, systèmes d'IA utilisés dans le secteur bancaire et financier pour la réalisation d'investissements ou l'évaluation de la capacité d'emprunt des clients...). Or, le droit de la responsabilité civile, qui n'a pas été pensé pour s'occuper des méfaits de l'immatériel pur, est a priori mal armé pour traiter les dommages causés par les systèmes d'IA désincarnés.

❖ Second défi technique : surmonter les obstacles probatoires.

Les victimes de dommages causés par l'IA vont se trouver confrontées à des obstacles probatoires à la fois considérables et inédits.

Tout d'abord, le grand nombre de participants à l'élaboration d'un système d'IA (concepteur de l'algorithme, programmeur qui l'implémente, *data scientist* qui gère les données, entraîneur d'IA qui conçoit et supervise les processus d'apprentissage...) complique l'identification du responsable.

Ensuite et surtout, les caractères propres des systèmes d'IA contemporains sont de

nature à placer les victimes dans une situation d'inaptitude à la preuve.

Pour le comprendre, il convient de rappeler que le modèle d'IA aujourd'hui dominant est l'IA dite « connexionniste » ou « neuronale ». Celle-ci repose sur un algorithme d'apprentissage basé sur des réseaux de neurones artificiels, inspirés de la structure du cerveau humain. Un système d'IA connexionniste n'est pas conçu pour exécuter directement une tâche pré-programmée mais bien plutôt pour apprendre par lui-même à réaliser une tâche. L'IA connexionniste est ainsi auto-apprenante (le fameux *machine learning*, dont la forme la plus en vue est aujourd'hui le *deep learning*) : à partir d'une gigantesque masse de données (*big data*), le système apprend à réaliser une tâche en cherchant des récurrences ou corrélations entre les données, sans instructions préprogrammées. Un système d'IA connexionniste raisonne ainsi par induction probabiliste à partir des données mises en corrélation statistique, de sorte que le résultat auquel il parvient est exprimé sous forme de probabilité. De là découlent deux caractéristiques notables des systèmes d'IA connexionnistes.

La première tient à une certaine *autonomie de fonctionnement*. Les systèmes d'IA connexionnistes reposent sur un algorithme « non-déterministe » : dans la réalisation des tâches qui leur sont confiées, ces systèmes ont la capacité d'effectuer des choix non déterminés à l'avance par un être humain (qu'il s'agisse du concepteur de l'IA ou de son utilisateur). Il s'ensuit que leur comportement n'est pas prévisible, au point qu'à partir d'une même entrée, un système d'IA connexionniste peut, d'une fois sur l'autre, produire des résultats différents.

La seconde caractéristique des systèmes d'IA connexionnistes, qui est étroitement corrélée à la première, réside dans leur *opacité de fonctionnement* (que les scientifiques nomment volontiers « l'effet boîte noire ») : le mécanisme interne de fonctionnement d'un système d'IA qui le mène de la masse des données d'entrée qui l'alimente jusqu'au résultat qu'il va produire (recommandation, décision, production de contenu...) demeure mystérieux, même pour le concepteur du système. Voudrait-on le disséquer *ex post* qu'il faudrait reconstituer un cheminement faisant intervenir des millions de neurones artificiels et des milliards de connexions, ce qui n'est guère praticable au regard du temps et de l'énergie que cela mobiliserait. Aussi bien, en dépit des recherches contemporaines tendant à améliorer l'explicabilité algorithmique afin de mieux appréhender la logique de fonctionnement des systèmes d'IA connexionnistes, l'opacité de ceux-ci est-elle loin d'être levée.

L'autonomie et l'opacité de fonctionnement des systèmes d'IA connexionnistes sont, en réalité, systémiques, dans la mesure où elles procèdent de la conjonction de plusieurs facteurs techniques qui sont inhérents à ces systèmes : nature même de l'algorithme, qui est non-déterministe ; auto-apprentissage du système d'IA (qui peut se poursuivre au-delà de la date de mise sur le marché : l'apprentissage en continu, lorsqu'il est implémenté dans le système, permet à celui-ci d'évoluer par lui-même au fil du temps, au-delà de sa programmation initiale).

Or, cette autonomie et cette opacité systémiques de fonctionnement auront inévitablement pour effet de rendre beaucoup plus difficile la preuve, par la victime, du fait générateur de responsabilité et de son lien causal avec le dommage.

2) Stratégie choisie pour relever les défis soulevés par l'IA

Pour affronter les défis que pose l'IA, le droit français a choisi de privilégier une

approche prudente, pragmatique et progressive.

Dans un premier temps, on s'en tient aux ressources qu'offre le droit positif, en se contentant, pour indemniser les dommages causés par l'IA, d'utiliser les régimes de réparation déjà existants. Au regard des difficultés tout-à-fait inédites que soulève l'IA, telles qu'on vient de les esquisser sommairement, certains aménagements devront tout de même être apportés aux régimes existants afin de leur permettre de traiter avec une efficacité suffisante les dommages provoqués par l'IA. Pour l'heure, à vrai dire, seule la responsabilité du fait des produits défectueux fait l'objet d'un effort législatif d'adaptation au particularisme de l'IA, dont il n'est au demeurant pas certain qu'il soit suffisant (infra II.7 c, d et e).

Ce n'est que si, à l'usage et au regard des avancées futures de l'IA, les instruments existants devaient révéler leur insuffisance qu'il conviendrait, dans un second temps, d'envisager la création d'instruments nouveaux, propres à l'IA.

2. Définition juridique et classification

a) Comment votre juridiction définit-elle ou classe-t-elle juridiquement les systèmes d'IA ?

Le Comité consultatif national d'éthique a défini l'intelligence artificielle comme une « construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique » (Rapport de synthèse sur les états généraux de la bioéthique, 2018, p. 110). Cette définition est toutefois dépourvue de portée normative.

Il n'existe pas de définition « officielle » de l'intelligence artificielle formulée par les pouvoirs publics français. Mais la définition qu'énonce le règlement européen sur l'IA du 13 juin 2024 (qu'a rappelée le rapporteur général en ouverture de ce questionnaire) s'impose en droit interne puisqu'elle est de source supra-législative.

Le droit français classe les systèmes d'IA dans la catégorie des choses. Il ne s'est, en effet, pas rangé à l'avis (minoritaire) de ceux qui prônent une personnification des systèmes d'IA. La majorité de la doctrine française considère, en effet, qu'attribuer la personnalité juridique aux systèmes d'IA afin de les rendre personnellement responsables des dommages qu'ils causent serait contestable sur le plan théorique et inopportun sur le plan pratique.

b) Existe-t-il une classification des différents types d'IA ou des niveaux de risque qu'elles posent ?

A l'échelon européen, le règlement sur l'IA du 13 juin 2024 établit une classification des systèmes d'IA en considération du degré de risques qu'ils présentent. Trois catégories sont ainsi distinguées.

La première est celle des systèmes d'IA à risques inacceptables, dont l'article 5 dresse une liste limitative. Ils sont purement et simplement prohibés.

La deuxième catégorie correspond aux systèmes d'IA à haut risque (article 6 et annexes 1 et 3). Présentant des risques substantiels pour la santé, la sécurité et les droits fondamentaux, ils sont étroitement encadrés et doivent respecter de nombreuses exigences énoncées par les articles 8 à 15.

La troisième catégorie est constituée par les systèmes d'IA à bas risque, qui ne sont soumis qu'à des exigences limitées (incitation à adopter un code de bonne pratique ; exigences de transparence).

A ce triptyque, le règlement ajoute les systèmes d'IA à usage général, qui se caractérisent par la « capacité à répondre à diverses finalités » (art. 3, §6), pour lesquels sont formulés certaines exigences spécifiques (articles 53 et s.)

Cependant, cette classification des systèmes d'IA n'intéresse pas directement le droit de la responsabilité civile, puisque le règlement sur l'IA ne traite pas de la réparation des dommages causés par l'IA (même si elle peut néanmoins avoir quelques répercussions indirectes en ce domaine, qu'on relèvera, chemin faisant, dans les réponses au questionnaire). On peut donc dire que la classification des systèmes d'IA établie à l'échelon européen ne structure pas la réparation des dommages causés par les systèmes d'IA.

A cet égard, le droit français adopte fondamentalement une *approche indifférenciée* : les conditions et modalités de l'indemnisation des victimes de l'IA ne dépendent pas du type de système d'IA à l'origine du dommage.

II. ASPECTS SPECIFIQUES DE LA RESPONSABILITE CIVILE EN MATIERE D'IA

1. Fondements de la responsabilité civile

a) Quels sont les principaux fondements de la responsabilité civile dans les affaires liées à l'IA dans votre juridiction ?

Il n'existe pas, en droit français, de régime de réparation propre aux dommages causés par l'IA (*supra* I.1.b). En conséquence, les victimes de l'IA doivent fonder leur demande de réparation sur des régimes préexistants, dont l'objet premier n'est pas de traiter les dommages que peut occasionner cette nouvelle technologie. Dans la mosaïque des régimes de réparation existant en droit positif, plusieurs peuvent, à des degrés divers, servir à indemniser les dommages causés par l'IA. Certains ont vocation à jouer un rôle principal, tandis que d'autres ne peuvent prétendre qu'à un rôle secondaire.

Deux régimes existants sont appelés à jouer un rôle principal dans la réparation des dommages causés par l'IA, en ce qu'ils apparaissent plus efficaces que les autres pour indemniser les victimes de cette technologie. Il s'agit d'une part, de la responsabilité

du fait des produits défectueux (qui peut *a priori* servir à indemniser les dommages causés par tout système d'IA, quel qu'il soit) ; d'autre part, du régime d'indemnisation des victimes d'accidents de la circulation (qui pourra être utilisé par les victimes d'accidents de la circulation dans lesquels sont impliqués des voitures autonomes). Ces deux régimes présentent les traits communs suivants : d'abord, ils transcendent la distinction des responsabilités contractuelle et extracontractuelle, en ce qu'ils s'appliquent indistinctement à toute victime, qu'elle soit tiers ou cocontractant ; ensuite, ils ont un caractère exclusif ou quasi-exclusif, en ce que, dès lors que leurs conditions d'application sont réunies, ils écartent tous les autres régimes ou du moins certains d'entre eux ; enfin, il s'agit de régimes de responsabilité objective, de plein droit.

D'autres régimes de réparation existants peuvent, à titre secondaire, servir à indemniser les victimes de l'IA.

Il s'agit, d'abord, des régimes de responsabilité civile de droit commun, qui possèdent une vocation subsidiaire générale, en ce qu'ils représentent l'ultime ressource des victimes qui ne disposent pas d'une autre voie de droit pour obtenir satisfaction : responsabilité extracontractuelle pour faute ; responsabilité extracontractuelle du fait des choses ; responsabilité contractuelle.

Il s'agit, ensuite, de quelques régimes spéciaux qui, à titre exceptionnel, peuvent, dans des circonstances très particulières, servir à indemniser certaines victimes de l'IA (ces régimes seront ultérieurement évoqués, au fil des questions).

b) Comment les fondements traditionnels de la responsabilité civile (par exemple, responsabilité civile contractuelle/délictuelle, responsabilité du fait des produits) pourraient-ils devoir être adaptés pour les systèmes d'IA ?

Les systèmes d'IA aujourd'hui dominants sont les systèmes d'IA dits « connexionnistes » qui présentent un particularisme marqué, tenant essentiellement à leur autonomie et à leur opacité systémiques de fonctionnement (*supra* I.1.c).

Ces propriétés tout-à-fait originales des systèmes d'IA ont pour effet de placer les victimes dans une *situation d'inaptitude à la preuve* (du fait générateur de responsabilité et du lien de causalité) : dès lors que le concepteur du système d'IA lui-même n'est pas capable de reconstituer le processus qu'a suivi le système pour parvenir au résultat qui s'est avéré dommageable pour la victime, comment celle-ci pourrait-elle ne pas ployer sous le fardeau de la preuve ?

C'est pourquoi une adaptation des régimes de réparation existants s'avère nécessaire si l'on veut qu'ils puissent traiter efficacement les dommages causés par l'IA. Dans cette perspective, il faudrait :

- corriger l'asymétrie dans l'aptitude à la preuve existant entre la victime et le défendeur, en imposant à celui-ci de divulguer à celle-là les éléments de preuve dont il dispose ;
- recourir à la technique présomptive, en admettant, à certaines conditions, des

- présomptions de fait générateur de responsabilité et/ou de lien de causalité ;
- adapter les causes d'exonération susceptibles d'être invoquées par le défendeur au caractère évolutif des systèmes d'IA afin que le responsable potentiel ne puisse pas s'exonérer trop aisément en soutenant le dommage causé par le système d'IA ne saurait lui être imputé parce qu'il résulte d'une auto-évolution ultérieure de celui-ci qui s'est émancipé par lui-même de sa configuration initiale.

Un tel effort d'adaptation a d'ores et déjà été entamé. Ainsi, la directive européenne n° 2024/2853 du 23 octobre 2024 se propose de mieux adapter la responsabilité du fait des produits défectueux à l'intelligence artificielle. Elle s'appliquera aux produits mis sur le marché après le 9 décembre 2026. La teneur précise des ajustements apportés à la responsabilité du fait des produits défectueux et la congruence de ceux-ci avec le besoin de protection des victimes de l'IA seront examinées plus loin (*infra* II.7 c, d et e).

En résumé, pour que les victimes de dommages causés par l'IA soient traitées de façon satisfaisante, il est nécessaire d'apporter des retouches aux régimes de réparation existants afin de les adapter au particularisme très prononcé des systèmes d'IA. Or, on le verra, il n'est pas sûr que les initiatives qui ont été prises en ce sens s'avèrent suffisantes.

c) Prévoyez-vous l'introduction de nouveaux fondements de responsabilité spécifiques à l'IA ?

Des velléités d'innovation se sont manifestées à l'échelon européen, qui n'ont cependant pas abouti.

En 2017, le Parlement européen a exploré la piste de la personnification des systèmes d'IA. Il s'agirait de traiter ceux-ci non plus comme des choses mais comme des personnes juridiques, des « personnes électroniques ». Une résolution du Parlement européen du 16 février 2017 a ainsi envisagé « la création, à terme, d'une personnalité spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques responsables, tenus de réparer tout dommage causé à un tiers ». La reconnaissance de la personnalité électronique aurait pour effet de rendre les systèmes d'IA personnellement responsables des dommages qu'ils causent. Cette idée, passablement fantasque, d'une responsabilité personnelle de systèmes d'IA personnifiés a fort heureusement été, par la suite, abandonnée.

Étant admis que les systèmes d'IA sont des choses – et non des personnes juridiques –, le Parlement européen a ensuite proposé, dans une résolution du 20 octobre 2020, la création d'un nouveau régime spécial de responsabilité du fait de l'IA (*infra* II.2 g). Mais celui-ci n'a pas davantage été consacré.

A l'échelon interne, le législateur français n'a jamais envisagé de créer un régime de responsabilité propre à l'IA. On notera ainsi que ni le projet de réforme du droit de la responsabilité civile du 13 mars 2017 émanant du ministère de la justice ni la proposition sénatoriale de réforme de la responsabilité civile du 20 juillet 2020 ne

s'intéressent aux dommages causés par l'IA.

2. Fait générateur

a) Comment le concept de faute ou d'action fautive est-il défini dans votre système juridique, tant dans le contexte contractuel qu'extra-contractuel ?

En matière contractuelle, la faute consiste en un manquement du débiteur à une obligation créée par le contrat.

Lorsque l'obligation contractuelle est une obligation de résultat (le débiteur a promis un résultat précis), il suffit, pour établir la faute contractuelle, de constater que le résultat promis n'a pas été procuré au créancier. La caractérisation de la faute contractuelle résulte alors d'un simple constat objectif.

Lorsque l'obligation contractuelle est une obligation de moyens (le débiteur a promis sa diligence en vue d'un objectif dont la réalisation n'est cependant pas garantie), il faut, pour établir la faute contractuelle, prouver que le débiteur n'a pas fourni la diligence qu'on est en droit d'attendre d'une personne raisonnable placée dans les mêmes circonstances. La caractérisation de la faute contractuelle nécessite donc, dans ce cas, une appréciation judiciaire du comportement du débiteur.

En matière extracontractuelle, l'article 1240 du Code civil pose un principe général de responsabilité pour faute : « *Tout fait quelconque de l'homme, qui cause à autrui un dommage, oblige celui par la faute duquel il est arrivé à le réparer* ». Ce principe général de responsabilité extracontractuelle pour faute est fondé sur un concept de faute général et abstrait, que cependant le législateur ne définit pas.

Étymologiquement, le mot « faute » vient du latin *fallere*, qui signifie défaillir. La faute civile extracontractuelle consiste donc, en première approche, en un comportement défaillant quel qu'il soit (qu'il s'agisse d'une action ou d'une abstention, que le comportement soit gravement défaillant ou que la défaillance soit bénigne). Plus précisément, la faute civile extracontractuelle est un comportement défaillant en ce qu'elle viole une règle de conduite.

La règle de conduite violée peut être précisément imposée par une loi ou par un règlement. La faute constitue alors un comportement *illicite*. En pareil cas, pour caractériser la faute, il n'est pas nécessaire d'analyser le comportement du sujet ; le seul constat de la méconnaissance de la règle textuelle suffit à établir la faute. Aucun problème d'appréciation de la faute ne se pose donc.

La règle de conduite violée peut être le devoir général de prudence et de diligence (également dénommé devoir général de conduite). La faute s'entend alors d'un comportement *anormal*. Se pose, dans ce cas, un problème d'appréciation de la faute. La caractérisation de celle-ci nécessite, en effet, un jugement de valeur porté par le juge sur le comportement du sujet. Celui-ci doit être comparé au comportement qu'aurait adopté, dans les mêmes circonstances, le modèle de référence de la personne raisonnable.

Il est à noter, qu'en droit français, la faute civile est une notion purement objective dans la mesure où elle ne requiert pas que son auteur soit doué de discernement.

Même un sujet non-discernant (jeune enfant, personne souffrant de troubles mentaux) peut se voir imputer une faute civile.

b) Dans quelle mesure la notion de faute, qu'elle soit fondée sur l'intention ou la négligence, peut-elle être appliquée aux systèmes d'IA pour évaluer leur responsabilité dans des situations dommageables ?

Lorsqu'un dommage est causé par un système d'IA, la faute à l'origine du dommage ne peut pas, à proprement parler, être une faute du système d'IA lui-même. En effet, la faute civile s'entend nécessairement de la faute d'une personne. Or, les systèmes d'IA sont des choses et non des personnes juridiques (*supra* II.1c). On ne peut donc, en toute rigueur, parler de faute du système d'IA.

La faute ne peut être que la faute d'une personne, commise en amont et expliquant l'action dommageable du système d'IA. Tout système d'IA ayant, par hypothèse, été conçu par des humains et ayant vocation à être utilisé par des humains, c'est donc la faute de ces personnes qu'il s'agira de caractériser.

A priori, la très grande plasticité de la notion de faute civile, définie génériquement comme un comportement défaillant, quelle qu'en soit la nature ou la gravité, permet de l'appliquer indifféremment à un participant à l'élaboration du système d'IA (mauvaise conception de l'algorithme, défaut de mise à jour du logiciel, constitution incorrecte des données d'entrée, mauvais entraînement du système...), à un utilisateur (qui, par exemple, n'a pas respecté la notice d'utilisation), à un pirate (qui se livre à une cyberattaque sur un système d'IA dans le but de commettre un acte malveillant au moyen de celui-ci)...

Pourtant, en pratique, la responsabilité pour faute ne devrait jouer qu'un *rôle assez résiduel* dans l'indemnisation des dommages causés par l'IA. Cela tient à ce qu'en ce domaine, la mise en œuvre de la responsabilité pour faute se heurte à divers obstacles, tantôt de fond tantôt probatoires.

1) Obstacles de fond

Lorsque le défendeur peut être poursuivi sur le fondement de la responsabilité du fait des produits défectueux, il est acquis, en droit positif, que la victime conserve la faculté d'intenter contre lui, si elle le préfère, une action en responsabilité pour faute de droit commun. Une condition a toutefois été posée par la jurisprudence à la mise en œuvre de la responsabilité pour faute en pareil cas : la faute alléguée à l'encontre du défendeur doit être distincte du défaut de sécurité (CJCE, 25 avr. 2002, n° C-183/00). Le problème est que la notion de faute détachable du défaut de sécurité n'est pas clairement définie. Voilà qui pourrait constituer un premier obstacle, de fond, auquel pourrait se heurter l'action en responsabilité pour faute.

Lorsque la victime agit en responsabilité pour faute contre l'utilisateur d'un système d'IA, son action peut également se heurter à un obstacle de fond, qui résulte de la confiance que les utilisateurs seront naturellement enclins à accorder à un système

d'IA dont la fiabilité est reconnue. On peut prendre, pour illustration, l'exemple des professionnels de santé. Ceux-ci sont de plus en plus appelés à utiliser l'IA dans l'accomplissement des actes de prévention, de diagnostics et de soins. Or, le principe est que la responsabilité d'un professionnel de santé ne peut être engagée que sur la base d'une faute prouvée. La responsabilité pour faute trouve donc en ce domaine un champ d'application privilégié. Le recours à l'IA est cependant de nature à modifier l'appréciation de la faute médicale par le juge. En présence d'un système d'IA médical dont la fiabilité élevée est avérée, le juge pourrait, en effet, être enclin à considérer que le médecin ayant suivi le diagnostic de l'IA qui s'est révélé erroné n'a pas commis de faute, au motif que le modèle-type du bon praticien aurait fait de même.

2) *Obstacles probatoires*

En premier lieu, la victime qui agit en responsabilité pour faute contre un *participant à l'élaboration du système d'IA* peinera, dans nombre de cas, à identifier la faute et son auteur. Cette difficulté procède de la conjonction de deux facteurs.

Le premier tient à la multiplicité des acteurs intervenant dans l'élaboration d'un système d'IA. Contre qui la victime doit-elle agir ? Contre celui qui a conçu l'algorithme (qui a pu commettre une faute de conception) ? Contre celui qui l'a implémenté (qui a pu commettre une faute de codage) ? Contre celui qui a collecté les données (qui peuvent être incomplètes ou biaisées) ? Contre celui qui a contrôlé l'apprentissage de l'IA (qui a pu commettre une faute dans l'entraînement du système) ? Contre celui qui a intégré l'IA dans le robot (qui a pu commettre une faute d'assemblage, à l'origine d'une mauvaise réaction du support matériel à l'instruction donnée par le logiciel) ?

Le second facteur réside dans l'opacité des systèmes d'IA connexionnistes (*supra* I.1 c) : le mécanisme interne de fonctionnement d'un tel système d'IA, qui le mène des données d'entrées jusqu'au résultat produit en sortie, échappe même à ses concepteurs (le fameux effet « boîte noire »). Comment, dans ces conditions, la victime pourrait-elle ne pas éprouver de sérieuses difficultés pour rapporter la preuve d'une faute humaine à l'origine de son dommage ?

En second lieu, la victime qui agit en responsabilité pour faute contre un *utilisateur du système d'IA* risque, par exemple, d'éprouver également quelque difficulté à établir que le défendeur a utilisé le système d'IA sans se conformer à la notice d'utilisation ou qu'il a introduit dans le système des données non pertinentes ou incomplètes avant de formuler une requête ou encore qu'il a exercé un contrôle insuffisant sur le fonctionnement du système d'IA...

c) **Comment définiriez-vous le devoir de diligence pour les différentes parties prenantes de l'écosystème de l'IA (par exemple, développeurs, fabricants, opérateurs, utilisateurs) ?**

On a vu (II.2.a) que la faute civile consiste soit en la violation d'une règle de conduite précise édictée par la loi (comportement illicite) soit en la violation du devoir général de prudence et de diligence, tel qu'apprécié par le juge (comportement anormal).

Sur le premier aspect (comportement illicite), le règlement européen du 13 juin 2024 sur l'IA formule de nombreuses exigences à l'égard des participants à l'élaboration et à la commercialisation des systèmes d'IA à haut risque, ainsi qu'à l'égard des déployeurs (utilisateurs de systèmes d'IA dans le cadre de leur activité professionnelle). Ces exigences se traduisent concrètement par l'édiction de diverses obligations mises à la charge des fournisseurs (articles 16 et s.), des mandataires (article 22), des importateurs (article 23), des distributeurs (article 24) et des déployeurs (articles 26 et s.). Le non-respect de ces exigences formulées par le législateur européen pour chacune de ces parties prenantes à l'écosystème de l'IA constituera *ipso facto* une faute civile de nature à engager la responsabilité civile de chacune de ces catégories d'acteurs de l'IA (à condition, bien sûr, qu'elle soit en relation causale avec le dommage subi par la victime).

Sur le second aspect (comportement anormal), le juge caractérisera la faute en comparant le comportement de l'acteur de l'IA, défendeur à l'action en responsabilité, à celui qu'aurait adopté, dans les mêmes circonstances, le type-modèle de la catégorie dont il relève. Ainsi, le juge comparera le comportement qu'a eu le fournisseur de l'IA à celui qu'aurait eu, dans les mêmes circonstances, le modèle de référence du fournisseur normalement prudent et diligent ; il comparera le comportement qu'a eu l'utilisateur de l'IA à celui qu'aurait eu, dans les mêmes circonstances, le modèle de référence de l'utilisateur normalement prudent et diligent...

Il est à noter que les codes de bonne pratique ou de bonne conduite, que les instances de l'Union européenne invitent les acteurs de l'IA à adopter, pourront aider le juge à déterminer le comportement du modèle de référence auquel sera comparé celui du sujet jugé.

d) Dans le cas des systèmes d'IA autonomes ou d'auto-apprentissage, comment le droit devrait-il aborder les situations où l'IA prend des décisions ou effectue des actions qui n'étaient pas explicitement prévues par ses créateurs ?

Cette question étant insérée dans un ensemble de questions relatives à la responsabilité pour faute, c'est au regard de celle-ci qu'on l'appréhendera.

Que l'établissement de la faute du défendeur puisse se heurter à l'autonomie de fonctionnement des systèmes d'IA trouve une illustration dans un arrêt rendu par la Cour de cassation, le 19 juin 2013 (Cass. 1^{re} civ., 19 juin 2013, n° 12-17591). Dans cette affaire, une société d'assurance réclamait à Google des dommages et intérêts car, lorsqu'un internaute saisissait le nom de la société demanderesse sur le moteur de recherche Google, l'algorithme de celui-ci faisait apparaître, au troisième rang des suggestions proposées, le terme diffamatoire et injurieux d'« escroc » accolé au nom de la société. Pour écarter la responsabilité de la société Google, la Cour de cassation met en avant l'autonomie de fonctionnement du moteur de recherche : « *la fonctionnalité aboutissant au rapprochement critiqué est le fruit d'un processus purement automatique dans son fonctionnement et aléatoire dans ses résultats, de*

sorte que l'affichage des « mots clés » qui en résulte est exclusif de toute volonté de l'exploitant du moteur de recherche d'émettre les propos en cause ou de leur conférer une signification autonome au-delà de leur simple juxtaposition et de leur seule fonction d'aide à la recherche ». Même si cette décision a été rendue, non sur le fondement du régime général de responsabilité pour faute, mais sur celui de la loi spéciale du 29 juillet 1881 qui vise spécifiquement la diffamation et l'injure, il n'en demeure pas moins qu'elle invite, au-delà du cas d'espèce, à s'interroger sur l'incidence de l'autonomie de fonctionnement des systèmes d'IA sur la mise en œuvre de la responsabilité pour faute.

Le fait qu'un système d'IA ait, dans la réalisation des tâches qui lui sont confiées, la capacité d'effectuer des choix qui n'ont pas été prédéterminés par un être humain ne suffit pas à exclure purement et simplement tout recours à la responsabilité pour faute en ce domaine. L'autonomie de fonctionnement des systèmes d'IA n'est, en effet, pas immanente. Elle résulte d'une action humaine en amont qui a conçu le système pour qu'il fonctionne ainsi. Or, des fautes ont très bien pu être commises à ce stade (mauvaise conception de l'algorithme, mauvaise constitution des données, mauvais entraînement du système...). En revanche, ce qui est incontestable est que cette autonomie des systèmes d'IA (jointe à l'opacité de leur fonctionnement) est de nature à rendre extrêmement difficile la preuve par la victime d'une faute imputable à un acteur de l'IA (*supra* II.2.b).

Les instances de l'Union européenne ont parfaitement pris conscience des difficultés que ne manqueront pas de rencontrer les victimes de l'IA pour prouver la faute du défendeur. Une proposition de directive du 28 septembre 2022 s'était ainsi donnée pour objectif de faciliter la mise en œuvre de la responsabilité extracontractuelle pour faute en remédiant aux difficultés probatoires auxquelles les victimes de l'IA peuvent se trouver confrontées (au moyen d'une part, d'un procédé d'injonction de divulgation de preuves en cas de dommage causé par un système d'IA à haut risque et d'autre part, d'une présomption de faute dans le cas où le défendeur ne défère par à l'injonction judiciaire). Mais cette proposition de directive du 28 septembre 2022 a été retirée récemment par la Commission européenne (cf. Programme de travail 2025, Annexe 4, point 32). L'obstacle probatoire que l'autonomie (et l'opacité) de fonctionnement des systèmes d'IA représente pour les victimes souhaitant introduire une action en responsabilité pour faute à l'encontre d'un acteur de l'IA demeure donc entier.

e) Comment le concept de faute pourrait-il s'appliquer dans les cas où le préjudice résulte d'une interaction complexe entre plusieurs systèmes d'IA ou entre des systèmes d'IA et des acteurs humains ?

Deux hypothèses peuvent être distinguées.

Première hypothèse : la victime parvient à rapporter la preuve qu'une faute commise par un acteur de l'IA (concepteur, distributeur, utilisateur du système d'IA...) a concouru, avec d'autres facteurs à l'action dommageable du système d'IA.

Le fait que le préjudice subi puisse, outre cette faute, avoir d'autres antécédents

(action d'un autre système d'IA ou d'un autre acteur humain) est indifférent car, en droit français, dès lors que la responsabilité pour faute d'une personne est engagée, celle-ci est tenue de réparer intégralement le dommage subi par la victime, même si sa faute n'est pas seule à l'origine de celui-ci.

Seconde hypothèse : la victime ne parvient pas à déterminer, parmi plusieurs acteurs de l'IA précisément identifiés, lequel d'entre eux a commis la faute à l'origine de l'action dommageable du système d'IA. En pareil cas, le juge peut-il retenir une responsabilité collective de tous pour faute ?

La réponse est négative. La Cour de cassation décide, en effet, que, lorsque l'action de la victime est fondée sur la responsabilité pour faute, la condamnation collective de l'ensemble des responsables potentiels est subordonnée à la preuve d'une faute commise par chacun (V. Cass. 2^e civ., 3 novembre 2016, n° 15-25348). Dans la situation envisagée, la victime du dommage causé par un système d'IA ne pourra donc pas obtenir réparation de son dommage.

f) Dans quelle mesure la conformité aux normes de l'industrie, aux meilleures pratiques ou aux réglementations spécifiques à l'IA devrait-elle influencer la détermination du caractère fautif d'une action d'un système d'IA ?

Le fait, pour un acteur de l'IA, de respecter la réglementation légale spécifique à ce domaine (c'est-à-dire, au premier chef, le règlement du 13 juin 2024 sur l'IA), les normes techniques et autres règles de l'art ainsi que les codes de bonne pratique applicables en matière d'IA, milite *a priori* en faveur d'une absence de faute de sa part.

Cependant, un acte qui n'est contraire ni aux normes légales ou réglementaires ni aux normes techniques ou déontologiques, peut néanmoins être fautif si, au regard des circonstances particulières de l'espèce, il apparaît qu'il n'est pas conforme à ce qu'aurait fait, dans les mêmes circonstances, une personne normalement prudente et diligente (cf. Cass. 2^e civ., 31 mai 1965, D. 1965, 579).

g) Le préjudice causé par les systèmes d'IA est-il mieux encadré par la responsabilité stricte ou la responsabilité pour risque ? Quelle est la situation législative ou la discussion doctrinale autour de cette question ?

Au regard des obstacles auxquels se heurte l'application de la responsabilité pour faute en matière d'IA (*supra* II.2 b), il n'est pas douteux que la réparation des dommages causés par les systèmes d'IA sera mieux assurée par des régimes de responsabilité objective, de plein droit.

Le droit positif l'atteste : les régimes de responsabilité qui sont actuellement appelés à jouer un rôle prépondérant dans la réparation des dommages causés par l'IA, parce qu'ils sont mieux que les autres à même d'indemniser effectivement les victimes sont, on l'a vu, le régime de responsabilité du fait des produits défectueux et le régime d'indemnisation des victimes d'accidents de la circulation (*supra* II.1 a) ; or, ce sont là deux régimes de responsabilité de plein droit.

Cela dit, la responsabilité de plein droit entretient avec l'IA un rapport complexe. En ce domaine, un point d'équilibre doit être trouvé entre d'une part, la protection des victimes et d'autre part, la protection de l'innovation. Or, une responsabilité de plein droit du fait de l'IA, si elle sert la protection des victimes pourrait, lorsqu'elle incombe aux participants à l'élaboration de l'IA, nuire, par sa rigueur, à leur compétitivité face à la concurrence étrangère. Un contrepois doit donc être introduit, qui ne peut guère consister qu'à rompre avec le principe de réparation intégrale, en limitant les préjudices réparables.

Cette recherche d'un point d'équilibre entre protection des victimes et protection de l'innovation se manifeste, en droit positif, dans le cadre de la responsabilité du fait des produits défectueux : cette responsabilité de plein droit, qui pèse sur les producteurs, ne prend pas en charge les préjudices économiques purs et les préjudices moraux purs (*infra* II.7 c).

Cette recherche d'un point d'équilibre entre protection des victimes et protection de l'innovation se manifeste encore plus nettement si, quittant le droit positif, on s'inscrit dans une perspective prospective. On l'a vu, le droit positif fait, dans l'immédiat, le pari de s'en tenir aux régimes de réparation existants pour indemniser les dommages causés par l'IA (*supra* I.1 b). Mais si, à l'avenir, les instruments existants devaient finalement s'avérer insuffisants – ce qu'on ne peut pas totalement exclure –, il faudrait bien alors faire preuve d'inventivité pour améliorer le sort des victimes de l'IA, en créant un nouveau régime de réparation propre aux dommages causés par l'IA. Une réflexion prospective a d'ores et déjà été engagée en ce sens.

Le point de départ de cette réflexion prospective réside dans le constat suivant : la plupart des régimes de responsabilité civile existants susceptibles d'être utilisés pour réparer les dommages causés par l'IA ont ceci de commun qu'ils requièrent une anormalité du fait générateur (la faute est un comportement anormal, le fait de la chose suppose un état ou un comportement anormal de celle-ci, le défaut de sécurité d'un produit tient à sa dangerosité anormale). Or, c'est là que se situe, pour les victimes, le point névralgique. La preuve de l'anormalité du fait générateur imputé au défendeur est, en effet, contrecarrée par la nature même de l'IA. D'abord, il est inhérent à l'IA de pouvoir causer des dommages en dehors de toute anormalité de fonctionnement : « *la plus sophistiquée des intelligences artificielles, respectant les canons les plus élevés du secteur d'activité, n'en sera pas moins susceptible d'emporter des dommages qu'il sera impossible d'anticiper. Ceux-ci relèvent d'un aléa propre à la technologie* » (S. Méribet, Vers un droit de l'intelligence artificielle, Dalloz, coll. « Nouvelle Bibliothèque de thèse », Vol. 197, p. 486, n° 524). Ensuite, la preuve de l'anormalité du fait générateur à l'origine du dommage causé par l'IA se trouve compliquée par des facteurs intrinsèques à cette technologie, au premier rang desquels figurent l'autonomie et l'opacité des systèmes d'IA (*supra* I.1 c). Si une carence devait ultérieurement se manifester dans l'indemnisation des dommages causés par l'IA, c'est là que s'en trouverait la cause principale : elle tiendrait à ce que les régimes de responsabilité civile existants empêchent trop souvent les victimes d'obtenir réparation parce qu'elles échouent à établir

l'anormalité du fait générateur. C'est donc cet obstacle qu'il faudrait *de lege ferenda* surmonter.

Si l'obstacle principal à l'indemnisation des victimes de l'IA tient à l'exigence d'anormalité du fait générateur, la seule manière de le lever serait d'abandonner celle-ci. Il s'agirait donc de mettre en place un nouveau régime de responsabilité qui reposerait sur le *simple fait causal de l'IA* voire sur sa simple implication dans le dommage, en dehors de toute considération d'anormalité ou de défectuosité.

Mais l'automatisme total d'une telle responsabilité, si elle est évidemment favorable aux victimes, pourrait, en revanche, s'avérer d'une rigueur excessive à l'endroit des responsables potentiels et, par là-même, entraver le développement de l'IA. Aussi devrait-elle nécessairement être contrebalancée par des mesures qui en restreignent la portée : si, du côté des responsables, on rompt avec l'exigence d'anormalité du fait générateur, il faudrait, en contrepois, du côté des victimes, restreindre l'indemnisation. L'équilibre général du nouveau régime de responsabilité du fait de l'IA serait donc schématiquement le suivant : une responsabilité purement objective, déclenchée par la seule participation de l'IA à la production du dommage, indépendamment de toute considération d'anormalité ou de défectuosité, mais n'offrant aux victimes qu'une réparation limitée.

Sur cette trame générale, ont d'ores et déjà été tissés les contours de ce que pourrait être un nouveau régime de responsabilité propre à l'IA.

En premier lieu, le Parlement européen a, dans une résolution du 20 octobre 2020 (JOUE 2021, n° C 404/05), proposé la création d'un régime de responsabilité spécifique du fait de l'IA. Selon cette proposition, qui n'a à ce jour pas encore été suivie d'effet, la responsabilité incomberait aux « opérateurs » de systèmes d'IA, définis comme les personnes physiques ou morales exerçant un contrôle sur le système d'IA et ayant, par là-même, une influence sur la réalisation du risque de dommage.

Si les opérateurs de systèmes d'IA qui ne sont pas à haut risque ne seraient soumis qu'à une responsabilité pour faute présumée, en revanche les opérateurs de systèmes d'IA à haut risque seraient, quant à eux, responsables de plein droit du simple fait causal du système : le fait dommageable, même normal, d'un système d'IA à haut risque suffirait, en effet, à engager la responsabilité de l'opérateur. Celui-ci ne pourrait s'exonérer totalement qu'en rapportant la preuve d'une force majeure. Mais, pour contrebalancer cette automatisme de la responsabilité de l'opérateur, la victime n'aurait droit qu'à une réparation partielle de ses préjudices : d'une part, tous les préjudices ne seraient pas pris en charge ; d'autre part, l'indemnisation serait plafonnée.

En second lieu, M. Méribet, dans sa thèse de doctorat précitée consacrée à l'intelligence artificielle, suggère, lui aussi, la création d'une responsabilité de plein droit du fait, même normal, de l'IA (p. 486 et s., n° 524 et s.), dont la totale automatisme aurait pour contrepois une prise en charge des seules atteintes à l'intégrité physique ou psychologique de la victime, à l'exclusion des autres chefs de préjudice (p. 499, n° 537).

On le voit, responsabilité de plein droit du fait de l'IA et réparation intégrale de tous les préjudices causés par l'IA paraissent bien difficiles à concilier.

3. Causalité

a) Quel test de causalité est principalement utilisé dans votre juridiction pour établir le lien de causalité en matière de responsabilité civile (par exemple, causalité adéquate, équivalence des conditions, causalité proximale) ?

Aucun test de causalité n'a réellement été consacré en France. La Cour de cassation s'est, en effet, abstenue de poser des règles à cet égard. Les juges font preuve d'un très grand empirisme, au point qu'on considère volontiers que la causalité est, en définitive, une affaire de sentiment. Il s'ensuit que la jurisprudence sur la question de la causalité est impossible à systématiser. La preuve en est que les analyses doctrinales cherchant à rendre compte de la jurisprudence relative à la causalité se contredisent entre elles : selon certains auteurs, c'est la théorie de l'équivalence des conditions qui prédominerait ; selon d'autres, ce serait plutôt la théorie de la causalité adéquate, tandis que d'autres encore suggèrent qu'il faudrait distinguer selon la nature de la responsabilité : prédominance de la théorie de l'équivalence des conditions en matière de responsabilité pour faute ; prédominance de la théorie de la causalité adéquate dans les responsabilités de plein droit.

b) Comment ce test de causalité pourrait-il s'appliquer ou devrait-il être adapté dans les cas impliquant des systèmes d'IA, en considérant particulièrement la complexité et l'opacité de certains systèmes d'IA (effet "boîte noire") ?

La charge de la preuve du lien de causalité entre le fait générateur de responsabilité et le dommage incombe à la victime. Or, l'obstacle que constitue l'opacité de fonctionnement des systèmes d'IA dans l'administration de la preuve de la causalité est tel que le remède réside moins dans une adaptation du test de causalité (dont on n'a vu qu'aucun ne s'impose en France, le juge se montrant toujours très pragmatique dans son appréciation de la causalité dans chaque cas d'espèce) que dans le recours à la technique présomptive : seule l'édiction de présomptions de causalité pourrait s'avérer utile aux victimes de l'IA. Quelques avancées (timides) en ce sens peuvent être signalées.

En matière de responsabilité du fait des produits défectueux, la directive européenne n° 2024/2853 du 23 octobre 2024 instaure des présomptions de relation causale entre le défaut de sécurité du système d'IA et le dommage.

D'abord, le lien de causalité entre le défaut de sécurité et le dommage sera présumé lorsque celui-ci est d'une nature généralement compatible avec le défaut en question (article 10.3).

Ensuite, lorsqu'en raison de la complexité technique ou scientifique du produit (ce qui est, à l'évidence, le cas des systèmes d'IA), la preuve du rôle causal du défaut de sécurité du produit dans la réalisation du dommage s'avère excessivement difficile,

le lien de causalité pourra être présumé (article 10.4). Cette présomption de causalité est cependant subordonnée à la condition que la victime ait préalablement établi « *qu'il est probable (...) qu'il existe un lien de causalité entre la défektivité du produit et le dommage* » (article 10.4 b). Compte-tenu de l'asymétrie d'information existant, en matière d'IA, entre la victime et le défendeur, cette condition préalable pourrait représenter une sérieuse limite au jeu de la présomption. L'intérêt pratique de celle-ci pour les victimes de l'IA dépendra donc, dans une large mesure, de l'appréciation plus ou moins souple que les juges feront de cette condition d'application de la présomption.

En matière de responsabilité extracontractuelle pour faute, une proposition de directive européenne du 28 septembre 2022 avait envisagé l'instauration, au profit des victimes de l'IA, d'une présomption simple de lien causal entre la faute et le dommage mais cette proposition a finalement été retirée (*supra* II.2 d).

c) Votre système juridique reconnaît-il la notion de causalité partielle ou proportionnelle ? Si oui, comment cette notion pourrait-elle être appliquée dans les cas où un système d'IA est l'un des multiples facteurs contribuant au dommage ?

La réponse est négative. Le droit français adopte le système de la causalité « intégrale » : dès lors que le fait générateur de responsabilité imputé au défendeur est l'une des causes ayant contribué à la production du dommage, le défendeur est réputé l'avoir causé entièrement (indivisibilité de la causalité) et doit en conséquence le réparer intégralement. Ainsi, les coresponsables d'un même dommage sont, par principe, condamnés in solidum envers la victime.

4. Faute de la victime / Minimisation du dommage

- a) **Comment la notion de faute de la victime pourrait-elle s'appliquer différemment dans les cas impliquant des systèmes d'IA ?**
- b) **Dans votre pays, la faute de la victime constitue-t-elle une défense totale ou une défense partielle en matière de responsabilité ?**

Réponse couplée aux a) et b)

En droit commun, la faute de la victime exonère totalement le défendeur lorsqu'elle présente les caractères de la force majeure (irrésistibilité et imprévisibilité). Dans le cas contraire, elle ne produit qu'un effet partiellement exonératoire.

Dans les régimes spéciaux, l'effet exonératoire de la faute de la victime peut obéir à un régime dérogatoire, susceptible de varier d'un régime à l'autre.

S'agissant, plus spécifiquement de la faute de la victime d'un dommage causé par un système d'IA, on se trouve, en l'absence de contentieux à cet égard, réduit aux conjectures. Avec toute la prudence qui s'impose, on serait tenté d'avancer l'hypothèse qu'en ce domaine, la faute de la victime ne devrait trouver à s'appliquer qu'assez rarement.

Pour s'en convaincre, on se concentrera sur les deux régimes de réparation appelés à jouer un rôle prépondérant dans l'indemnisation des dommages causés par l'IA : la responsabilité du fait des produits défectueux et le régime d'indemnisation des victimes d'accidents de la circulation (*supra* II.1 a)

Considérons, en premier lieu, la responsabilité du fait des produits défectueux. Le responsable est exonéré totalement ou partiellement « *lorsque le dommage est causé conjointement par un défaut du produit et par la faute de la victime (...)* » (article 1245-12 du Code civil et article 13 de la directive du 23 octobre 2024). Deux considérations mènent néanmoins à penser que la faute de la victime devrait jouer assez rarement, s'agissant des dommages causés par un système d'IA défectueux.

La première a un caractère général. La Cour de cassation adopte globalement une attitude assez restrictive à l'égard de la faute de la victime d'un produit défectueux, en ne lui reconnaissant un effet exonératoire (total ou partiel) que lorsqu'elle a contribué à l'apparition même du dommage et non lorsqu'elle l'a simplement aggravé (Cass. 1^{re} civ., 2 juin 2021, n° 19-19349).

La seconde considération est spécifique aux dommages causés par l'IA. Lorsqu'on cherche à imaginer des fautes que pourrait commettre la victime d'un système d'IA, les exemples ne se bousculent pas. On songe à une utilisation du système ne respectant pas la notice d'utilisation ou encore à une surveillance insuffisante de son fonctionnement. Mais, au-delà, on peine à trouver d'autres illustrations. La raison pourrait en être que, s'agissant des systèmes d'IA, les occurrences de faute de la victime sont moins nombreuses que celles qu'on peut rencontrer en présence de produits ordinaires (non-intelligents). En effet, si l'utilisateur de ces derniers sera bien souvent amené à les manipuler matériellement, avec toutes les occasions de maladresses ou négligences que cela peut susciter, il n'en va pas de même de l'utilisateur d'un système d'IA : c'est une évidence en ce qui concerne l'utilisateur d'un système d'IA désincarné mais c'est vrai également pour celui d'un système d'IA incarné, qui se bornera à l'activer puis le laissera agir de manière autonome.

Envisageons, en second lieu, le régime d'indemnisation des victimes d'accidents de la circulation, issu de la loi du 5 juillet 1985, qui a vocation à s'appliquer aux véhicules autonomes (*infra* scénario E). La seule cause d'exonération que le responsable (le conducteur ou le gardien du véhicule impliqué) peut invoquer est précisément la faute de la victime. L'incidence de la faute de la victime varie selon la nature du dommage subi par la victime (corporel ou non) et selon la qualité de la victime (conductrice ou non-conductrice) :

- s'agissant d'une victime conductrice, toute faute, même légère, commise par elle en relation avec son préjudice est susceptible de réduire ou supprimer son droit à indemnisation (selon le degré de gravité de la faute), quelle que soit la nature du dommage subi ;
- s'agissant d'une victime non-conductrice, il faut distinguer selon la nature du dommage subi : l'indemnisation de son dommage matériel sera réduite voire exclue à la moindre faute de sa part ; en revanche, l'indemnisation intégrale de son dommage corporel est de principe, sauf faute intentionnelle (recherche volontaire du dommage) ou, si la victime a entre 16 et 70 ans et ne présente pas un taux d'invalidité dépassant 80%, faute inexcusable cause exclusive de

l'accident.

Supposons que la victime d'un accident de la circulation dans lequel est impliqué un véhicule autonome subisse un dommage corporel. Si elle n'est pas conductrice, la faute qu'elle a pu commettre ne lui sera presque jamais opposable (la recherche volontaire du dommage, qui correspond concrètement au suicide de la victime, est rarissime et la faute inexcusable cause exclusive de l'accident est conçue de manière très étroite par la jurisprudence). Si la victime a la qualité de conductrice du véhicule autonome, il n'y aura guère de faute susceptible de lui être reprochée dès lors que le système de délégation de conduite était activé au moment de l'accident, hormis celle de n'avoir pas réagi assez promptement à une éventuelle demande de reprise en main du contrôle du véhicule formulée par le système ou encore celle de n'avoir pas bouclé sa ceinture de sécurité.

On le voit, la faute de la victime d'un dommage dans lequel est impliqué un système d'IA pourrait n'avoir, somme toute, qu'une incidence mineure sur son droit à indemnisation.

c) Quelles mesures de minimisation du dommage pourrait-on attendre des victimes des systèmes d'IA ?

La jurisprudence française refuse de mettre à la charge de la victime une quelconque obligation de minimiser son propre dommage, quelle que soit la nature de celui-ci, corporelle ou matérielle (Cass. 1^{re} civ., 15 janvier 2015, n° 13-21180 ; Cass. 2^e civ., 26 mars 2015, n° 14-16011 ; Cass. 2^e civ., 15 décembre 2022, n° 21-16712 ; Cass. 1^{re} civ., 7 janvier 2022, n° 20-23440).

5. Préjudice / Dommage

a) Quels types de préjudices ou de dommages sont généralement protégés par le droit de la responsabilité dans votre juridiction ? Cette protection diffère-t-elle entre les contextes contractuel et extra-contractuel ?

Le système français de réparation des dommages ne repose pas sur la hiérarchisation des intérêts protégés. En ce qui concerne la réparation des dommages, le droit français repose sur un double principe.

Un *principe de généralité* d'abord : tous les préjudices, sans distinction quant à leur nature, sont, sauf disposition légale spéciale contraire, juridiquement réparables : préjudices corporels (résultant d'une atteinte à l'intégrité physique de la personne), préjudices matériels (résultant de la destruction, dégradation ou vol d'un bien), préjudices économiques purs (pertes économiques non consécutives à l'atteinte à l'intégrité physique d'une personne ou à l'intégrité matérielle d'un bien), préjudices moraux purs (souffrance morale non consécutive à une atteinte primaire à une personne ou à un bien).

Un *principe d'équivalence*, ensuite : tous les préjudices ont *a priori* une égale vocation à être intégralement réparés.

Ce double principe, de généralité et d'équivalence, s'applique tant en matière contractuelle qu'en matière extracontractuelle.

b) Existe-t-il des types de dommages spécifiques qui pourraient émerger ou devenir plus prévalents avec l'utilisation croissante des systèmes d'IA (par exemple, violation de la vie privée, discrimination algorithmique, perte d'autonomie) ? Comment votre système juridique est-il équipé pour traiter ces types de dommages ?

Il n'est pas aisé de répondre à cette question car, d'un côté, les systèmes d'IA sont potentiellement susceptibles de causer des dommages de toutes natures mais, d'un autre côté, ainsi que l'a relevé un auteur, il n'existe, pour l'heure, « *tout simplement pas encore de cas de dommages causés par l'intelligence artificielle, ou en tout cas très peu* » (J. -S. Borghetti, Faut-il consacrer un régime de responsabilité du fait de l'intelligence artificielle dans le Code civil ? *in* Le renouvellement du droit civil sous l'influence du numérique, Lefebvre Dalloz, 2024, p. 156). Sans doute y-a-t-il eu, à l'étranger, quelques défaillances dramatiques de systèmes d'IA qui ont marqué les esprits (crashes de deux Boeings survenus en 2018 et 2019, en raison de la défektivité du système d'IA destiné à empêcher le décrochage de l'appareil, ayant coûté la vie à plus de 300 passagers ; piéton tué en 2018 par une voiture autonome dans l'Arizona, à la suite d'une défaillance du système de délégation de conduite) mais, en France, faute de pouvoir s'appuyer sur l'analyse objective d'un contentieux suffisamment nourri, on se trouve, à cet égard, réduit aux conjectures.

Pour proposer tout de même quelques éléments de réponse, on peut partir de la distinction évoquée plus haut (*supra* I.1 c) entre les systèmes d'IA « incarnés », qui s'incorporent dans un objet matériel (cf. véhicules autonomes, robots chirurgicaux, robots sociaux...), et les systèmes d'IA « désincarnés », qui sont, quant à eux, dépourvus de tout support physique (cf. agents conversationnels, systèmes d'IA utilisés dans le secteur bancaire et financier pour la réalisation d'investissements ou l'évaluation de la capacité d'emprunt des clients...).

Les systèmes d'IA incarnés causeront plus volontiers des dommages corporels (cf. voiture autonome provoquant un dommage corporel à la suite d'une défaillance du système de délégation de conduite) ou matériels (cf. robot d'assistance aux personnes âgées ou malades qui, par un mouvement inapproprié dû à une défaillance de son logiciel, casse un objet de valeur appartenant à son utilisateur).

En revanche, les systèmes d'IA désincarnés devraient plus volontiers causer des préjudices économiques purs (cf. privation du prêt escompté, en raison d'une erreur du système d'IA utilisé par la banque pour l'octroi de crédits) ou des préjudices moraux purs (cf. atteinte à l'honneur subie par un élu local australien dont Chat GPT indiquait qu'il avait été condamné pour corruption alors qu'en réalité, il était le lanceur d'alerte ayant permis que les coupables soient condamnés ; discrimination à l'embauche résultant d'une défaillance du système d'IA utilisé par l'employeur pour

le recrutement ; atteinte à la vie privée...). On pourrait donc assister, avec le développement des systèmes d'IA désincarnés, à une augmentation des préjudices économiques purs et des préjudices moraux purs.

A hauteur de principe, ceux-ci sont réparables en droit français, au même titre que tout autre préjudice (*supra* II.5 a). Cependant, la responsabilité du fait des produits défectueux, qui constitue pourtant un pilier essentiel de l'indemnisation des dommages causés par l'IA, se montre réfractaire à ce type de préjudices (*infra* II.7 c). Cela signifie que les victimes de préjudices économiques purs ou de préjudices moraux purs causés par des systèmes d'IA, le plus souvent désincarnés, n'auraient d'autres ressources que de se tourner vers le droit commun de la responsabilité civile qui constitue la voie de réparation subsidiaire quand les autres se ferment. Or, le droit commun de la responsabilité civile n'offre à la victime que des chances d'indemnisation réduites.

Sans doute, lui sera-t-il, en tout état de cause, possible d'agir sur le fondement de la responsabilité pour faute mais on connaît les difficultés probatoires considérables que celle-ci soulève (*supra* II.2 b).

La victime pourrait donc être tentée de solliciter un autre régime de droit commun : la responsabilité du fait des choses qui, elle, est une responsabilité de plein droit. L'article 1242, alinéa 1^{er}, du Code civil, pose, en effet, un principe général de responsabilité extracontractuelle du fait des choses : « *On est responsable non seulement du dommage que l'on cause par son propre fait, mais encore de celui qui est causé par le fait (...) des choses que l'on a sous sa garde* ». Le gardien d'une chose (c'est-à-dire celui qui en a l'usage, la direction et le contrôle) est, de plein droit, responsable des dommages causés par un fait anormal de celle-ci et ne peut s'exonérer que par la preuve d'une cause étrangère présentant les caractères de la force majeure (exonération totale) ou d'une faute de la victime ayant contribué à la réalisation du dommage (exonération partielle). Mais pour que les systèmes d'IA désincarnés entrent dans le domaine d'application de la responsabilité du fait des choses, encore faut-il qu'ils puissent être qualifiés de choses, au sens de l'article 1242, alinéa 1^{er}, du Code civil. Or, rien n'est moins sûr. Quand l'IA est incorporée dans un support physique (cf. robot intelligent), on se trouve incontestablement en présence d'une chose, au sens de ce texte. Mais lorsque le système d'IA est désincarné, il constitue une chose immatérielle. Or, la question de savoir si la responsabilité générale du fait des choses peut s'appliquer aux choses immatérielles fait débat en France. La doctrine se montre partagée à cet égard. En faveur d'une application de la responsabilité du fait des choses aux choses immatérielles, certains auteurs font valoir que la jurisprudence a, jusqu'à présent, écarté toutes les tentatives de restriction des choses entrant dans le champ d'application de l'article 1242, alinéa 1^{er}, du Code civil, lequel joue, que la chose soit mobilière ou immobilière, qu'elle soit dangereuse ou non, qu'elle soit actionnée par la main de l'homme ou non. D'autres auteurs, en revanche, contestent l'application de la responsabilité du fait des choses aux choses immatérielles, au motif que ce régime n'a été conçu que pour les choses corporelles, ainsi qu'en témoigne la manière dont la jurisprudence appréhende la notion de fait de la chose et celle de garde de la chose. Il existe donc un doute sérieux quant à l'applicabilité de la responsabilité générale du fait des choses aux systèmes d'IA désincarnés, que le droit prospectif de la responsabilité civile ne fait qu'aviver puisque tant le projet de réforme de la responsabilité civile du 13 mars 2017,

émanant du ministère de la justice (article 1243) que la proposition sénatoriale de réforme de la responsabilité civile du 20 juillet 2020 (article 1242) cantonnent *expressis verbis* la responsabilité du fait des choses aux seules choses corporelles.

En définitive, il se pourrait donc que les victimes de systèmes d'IA désincarnés, qui causent plutôt des préjudices économiques et moraux purs, soient moins bien traitées. Il y a là un véritable problème.

6. Responsabilité entre multiples acteurs

a) Comment votre système juridique traite-t-il la responsabilité plurale ou multiple dans les cas de dommages causés par plusieurs acteurs ?

Tant en matière contractuelle qu'en matière extracontractuelle, les coresponsables d'un même dommage sont, par principe, condamnés *in solidum* à le réparer. La victime est donc en droit de réclamer l'intégralité de l'indemnisation à laquelle elle a droit à l'un quelconque des coresponsables. En d'autres termes, chaque coresponsable est obligé au tout.

b) Dans le contexte des systèmes d'IA, comment la responsabilité (solidaire, in solidum, conjointe, etc.) pourrait-elle s'appliquer entre les différents acteurs de la chaîne de valeur (par exemple, développeurs, fabricants, opérateurs, utilisateurs) ? Quels critères devraient être utilisés pour déterminer l'application de la responsabilité entre multiples acteurs ?

Dès lors que plusieurs personnes différentes, quelles qu'elles soient (participant à l'élaboration du système d'IA – fabricant, développeur, opérateur – ; propriétaire du système d'IA ; utilisateur du système d'IA) peuvent se voir imputer un fait générateur de responsabilité ayant, au moins pour partie, contribué à la réalisation du dommage causé par un système d'IA, toutes seront condamnées *in solidum* à réparer le dommage subi par la victime. Cette responsabilité *in solidum* s'applique quelle que soit la nature de la responsabilité encourue par les différents protagonistes (contractuelle ou extracontractuelle, pour faute ou sans faute)

On peut, pour illustration, raisonner sur le cas d'un système d'IA incarné. On se trouve alors en présence d'un produit composite (ensemble formé par l'objet corporel et le système d'IA qui y est intégré). Supposons que la victime soit parvenue à prouver que c'est un défaut de sécurité de ce produit composite qui est à l'origine de son dommage. Elle peut donc fonder son action en réparation sur la responsabilité du fait des produits défectueux. Le juge condamnera solidairement le producteur de l'objet matériel incorporant l'IA et le producteur du système d'IA intégré (article 1245-7 du Code civil et article 12 de la directive du 23 octobre 2024). La victime pourra ainsi réclamer l'intégralité de sa créance de dommages et intérêts à l'un ou à l'autre des coresponsables, sans avoir à déterminer si la défectuosité à l'origine de son dommage provient de la fabrication de l'objet corporel ou de la conception du système d'IA intégré.

c) Comment votre système juridique traite-t-il les cas où certains acteurs potentiellement responsables ne peuvent pas être identifiés ou sont insolvables ? Cette approche devrait-elle être modifiée dans le contexte des systèmes d'IA ?

La question aborde deux problèmes distincts : celui de l'ignorance quant à l'identité du responsable et celui de l'insolvabilité de celui-ci. Ces deux problèmes seront successivement examinés au regard du droit positif, d'abord (1) puis au regard du droit prospectif, ensuite (2).

1/ En droit positif

1.1) Le responsable n'est pas identifié

Trois situations différentes peuvent se présenter en pratique.

- ◆ L'ignorance quant au(x) responsable(s) du dommage causé par l'IA est totale. En pareil cas, il n'est, à l'évidence, pas possible de faire jouer la responsabilité civile.
Parfois néanmoins, une victime de l'IA pourra être indemnisée au titre de la solidarité nationale lorsque le responsable demeure totalement inconnu. Supposons, par exemple, qu'un véhicule autonome percute un piéton. Le véhicule, qui ne s'est pas arrêté, n'a pas pu être identifié. La voie de la responsabilité civile est, de ce fait, impraticable. La victime pourra toutefois réclamer réparation intégrale à un fonds de garantie, le fonds de garantie des assurances obligatoires de dommages (FGAO) dont l'une des missions est précisément d'indemniser les victimes d'accidents de la circulation impliquant un véhicule dont le conducteur est inconnu.
- ◆ Le dommage causé par un système d'IA est imputable à plusieurs coresponsables mais un ou plusieurs d'entre eux n'ont pas pu être identifiés. Cette situation d'ignorance partielle quant à l'identité des coresponsables sera, à vrai dire, indolore pour la victime. Adeptes de la causalité intégrale (*supra* II.3 c), le droit français considère que celui qui, par son fait, n'a contribué qu'en partie à la réalisation du dommage n'en est pas moins juridiquement réputé l'avoir causé entièrement et doit, en conséquence, indemniser intégralement la victime qui s'adresse à lui. Celle-ci pourra donc obtenir réparation intégrale auprès du ou des coresponsables identifiés.
- ◆ Le responsable du dommage causé par l'IA n'est pas précisément identifié mais on sait tout de même qu'il se trouve nécessairement au sein d'un ensemble de personnes déterminées. On ne se trouve donc pas dans une situation d'ignorance absolue quant à la désignation du responsable mais plutôt dans une situation de doute quant à son identification précise. Imaginons, par exemple, qu'afin de permettre son maintien à domicile, une personne âgée ait fait l'acquisition de deux robots d'assistance à la personne. Il est établi que les deux robots présentent un défaut de sécurité. Le propriétaire est victime d'un dommage dont on sait qu'il a été causé par le défaut de sécurité de l'un des deux robots, sans qu'on puisse déterminer

lequel précisément. Est-ce le défaut de sécurité du robot fabriqué par X qui a causé le dommage ou est-ce le défaut de sécurité du robot fabriqué par Y ? On l'ignore. Le responsable n'est certes pas nommément identifié mais il n'y a qu'une alternative : soit c'est X soit c'est Y. En pareil cas, la jurisprudence accepte d'appliquer ce qu'on nomme la théorie de la causalité alternative : l'ensemble des coresponsables potentiels encourant une responsabilité de plein droit seront déclarés collectivement responsables (en l'occurrence, au titre de la responsabilité du fait des produits défectueux), sur la base d'une présomption de causalité (alternative). La victime obtiendra ainsi la condamnation *in solidum* de l'ensemble des responsables potentiels.

1.2) Le responsable est insolvable

En règle générale, l'insolvabilité du responsable n'est pas un obstacle à l'indemnisation de la victime. Il y a plusieurs raisons à cela.

D'abord, l'assurance de responsabilité est très largement répandue en France et, en pratique, les acteurs de l'IA sont couverts par une assurance. La victime d'un système d'IA pourra ainsi être indemnisée par l'assureur du responsable insolvable (sauf résiliation du contrat d'assurance par l'assurance en raison d'un défaut de paiement des primes).

Ensuite, lorsque le responsable est insolvable et non assuré, la victime pourra parfois être indemnisée par un fonds de garantie. Il en est ainsi, par exemple, en matière d'accident de la circulation : lorsque le conducteur du véhicule impliqué dans l'accident est insolvable et non assuré, la victime peut s'adresser au fonds de garantie des assurances obligatoires (FGAO), dont l'une des missions est d'indemniser les victimes d'accidents de la circulation impliquant un véhicule, autonome ou non, dont le conducteur est insolvable et non assuré.

Enfin, lorsqu'il y a plusieurs coresponsables, ceux-ci sont, par principe, condamnés *in solidum*, de sorte que l'insolvabilité de l'un d'entre eux sera supportée, non par la victime, mais par les autres coresponsables solvables.

2/ En droit prospectif

Les remèdes à l'indétermination ou à l'insolvabilité du ou des responsables sont, en droit positif, assez ponctuels et disparates. Toutes les victimes de dommages causés par l'IA n'en profiteront pas. Dès lors, dans un souci de protection plus systématique de celles-ci, ne serait-il pas judicieux d'envisager prospectivement la création d'un fonds de garantie des victimes de l'IA qui interviendrait subsidiairement dans tous les cas où le responsable d'un dommage causé par l'IA ne parvient pas à obtenir réparation par la voie de la responsabilité civile ? La mise en place d'un tel fonds de garantie des victimes de l'IA a, au demeurant, été explicitement évoquée par le Parlement européen, dans une résolution du 16 février 2017 (résolution du Parlement européen du 16 février 2017 contenant des recommandations à la

Commission concernant des règles de droit civil sur la robotique, point 58).

L'instauration d'un fonds de garantie chargé d'indemniser à titre subsidiaire les victimes de l'IA suscite des réserves. Les dommages les plus graves que peut causer l'IA, à savoir les dommages corporels, font d'ores et déjà l'objet, dans une large mesure, d'une prise en charge subsidiaire par la solidarité nationale, lorsque la responsabilité civile ne parvient pas à les indemniser. Les deux sources principales de dommages corporels causés par l'IA sont très certainement d'une part, les accidents de la circulation dans lesquels est impliqué un véhicule autonome ; d'autre part, les accidents médicaux survenus à l'occasion d'un acte de soins pour la réalisation duquel a été utilisé un système d'IA. Or, dans ces deux domaines, il existe déjà, en droit positif français, des fonds d'indemnisation intervenant à titre subsidiaire : l'office national d'indemnisation des accidents médicaux (ONIAM), pour les accidents médicaux échappant à la responsabilité civile, qu'ils impliquent ou non un système d'IA ; le fonds de garantie des assurances obligatoires de dommages (FGAO), pour les accidents de la circulation dont le responsable est inconnu ou non-assuré, qu'ils impliquent ou non un véhicule autonome.

A partir de là, le législateur désireux de créer un fonds de garantie des victimes de l'IA d'intervention subsidiaire se trouverait confronté à un véritable dilemme puisqu'il devrait choisir entre deux options contraires, toutes deux insatisfaisantes.

Première option : décider que toute victime d'un système d'IA qui ne parvient pas à obtenir réparation par la voie de la responsabilité civile ne peut subsidiairement s'adresser qu'au fonds de garantie des victimes de l'IA, à l'exclusion de l'ONIAM ou du FGAO qui pouvaient, auparavant, être le cas échéant sollicités. Mais alors, il faudrait distinguer : selon qu'un système d'IA médicale a été utilisé ou non, la victime d'un accident médical devrait subsidiairement réclamer réparation au « fonds IA » ou à l'ONIAM ; selon qu'un véhicule autonome est ou non impliqué dans l'accident de la circulation, la victime devrait subsidiairement s'adresser au « fonds IA » ou au FGAO. Or, les règles de procédure et les méthodes d'évaluation de l'indemnisation varient d'un fonds à l'autre. Donc, une inégalité serait instaurée entre les victimes d'un même type de situation dommageable (accident médical, dans un cas ; accident de la circulation, dans l'autre).

Seconde option : maintenir inchangée la compétence de l'ONIAM et du FGAO, qui continueraient d'avoir vocation à indemniser subsidiairement les dommages corporels résultant d'un accident médical impliquant une IA ou d'un accident de la circulation impliquant un véhicule autonome. En pratique, le nouveau fonds IA servirait alors surtout à indemniser les préjudices économiques purs ou moraux purs que peuvent principalement causer les systèmes d'IA désincarnés. L'indemnisation des dommages causés par l'IA que la responsabilité civile ne parvient pas à réparer relèverait alors de trois fonds de garantie différents (l'ONIAM, le FGAO et le nouveau fonds IA). Ce serait aggraver l'émiettement, souvent déploré, du droit français de la réparation des dommages.

On comprend, dans ces conditions, qu'il soit permis de s'interroger, du point de vue

de la politique juridique, sur la pertinence de la création d'un nouveau « fonds IA » d'intervention subsidiaire.

d) Existe-t-il des mécanismes juridiques dans votre juridiction pour répartir équitablement la responsabilité entre les acteurs de la chaîne de valeur de l'IA ?

Deux mécanismes peuvent être mentionnés.

Le premier, qui est de portée limitée puisqu'il ne concerne que la responsabilité du fait des produits défectueux, réside dans l'instauration d'une hiérarchie entre les opérateurs économiques intervenant dans la chaîne de production et de distribution. La responsabilité des distributeurs professionnels d'un système d'IA défectueux ne peut, en effet, être recherchée par la victime que si le producteur de celui-ci est inconnu (Code civil, article 1245-6 et directive du 23 octobre 2024, article 8).

Le second mécanisme a, quant à lui, une portée générale. Il s'agit du mécanisme de répartition finale de la dette de réparation entre les coresponsables d'un même dommage, après que l'un d'entre eux a indemnisé la victime.

En tous domaines, le principe est que lorsque plusieurs personnes peuvent être tenues pour coresponsables d'un même dommage, sur quelque fondement que ce soit, toutes sont tenues *in solidum* de la dette de réparation envers la victime. Celle-ci peut donc réclamer l'intégralité des dommages et intérêts auxquels elle a droit à l'un quelconque des coresponsables (*supra* II.6 a et b). Mais après que l'un d'entre eux aura intégralement indemnisé la victime, va se poser la question de la répartition finale de la dette de réparation entre les coresponsables. C'est la question de la contribution à la dette de réparation.

Dans le silence du législateur, c'est la jurisprudence qui a élaboré le mécanisme de contribution des coresponsables à la dette de réparation. Pour l'exposer aussi clairement que possible, on raisonnera sur l'hypothèse où il y a deux coresponsables du dommage. Quatre situations doivent être distinguées :

- 1^{re} situation : le coresponsable qui a entièrement indemnisé la victime ainsi que l'autre coresponsable ont tous les deux commis une faute. Dans cette situation, la répartition de la dette de réparation entre les deux coresponsables fautifs se fait en proportion de leurs fautes respectives. C'est donc le critère de la gravité des fautes respectives qui préside, en pareil cas, à la contribution à la dette.
- 2^e situation : aucun des coresponsables n'a commis de faute. Dans cette situation, la répartition de la dette de réparation entre les coresponsables se fait par parts égales.
- 3^e situation : le coresponsable qui a entièrement indemnisé la victime n'a pas commis de faute tandis que l'autre coresponsable est, quant à lui, fautif. Dans cette situation, le premier peut récupérer auprès du second la totalité de ce qu'il a versé à la victime.
- 4^e situation : le coresponsable qui a entièrement indemnisé la victime est fautif tandis que l'autre coresponsable n'a, quant à lui, pas commis de faute. Dans cette situation, le premier ne dispose d'aucun recours en contribution contre le second. Doté d'une portée tout-à-fait générale, ce mécanisme de répartition finale de dette

de réparation entre les coresponsables a vocation à s'appliquer aux coresponsables d'un dommage causé par un système d'IA.

On peut ajouter, pour finir, que ce mécanisme de répartition finale de la dette de réparation entre les différents acteurs de l'IA ne s'applique que sous réserve de correctifs ponctuels prévus par une disposition légale spéciale contraire (cf. directive européenne du 23 octobre 2024 relative à la responsabilité du fait des produits défectueux, article 12.2 a, excluant le recours du fabricant ayant intégré un logiciel défectueux dans un produit contre la start-up ou la petite ou moyenne entreprise ayant créé ce logiciel).

7. Responsabilité du fait des produits

a) Existe-t-il un régime spécifique de responsabilité du fait des produits dans votre juridiction ?

Il existe, en droit positif français, un régime de responsabilité de plein droit du fait des produits défectueux. Ce régime, issu de la directive européenne n° 85/374 du 25 juillet 1985, transposée en droit interne par une loi du 19 mai 1998, est régi par les articles 1245 et suivants du Code civil.

Pour mettre en œuvre la responsabilité du fait des produits défectueux, la victime doit prouver que les trois éléments suivants sont réunis :

- un défaut de sécurité du produit : un produit (tout bien meuble, même s'il est incorporé dans un autre bien) est défectueux lorsqu'il n'offre pas la sécurité à laquelle on peut légitimement s'attendre ;
- un dommage : sont réparables, au titre de la responsabilité du fait des produits défectueux, les atteintes à la personne et les atteintes aux biens autres que le produit défectueux lui-même ;
- un lien de causalité entre le défaut du produit et le dommage.

Lorsque cette triple preuve est rapportée, la responsabilité du fait des produits défectueux peut s'appliquer. Elle incombe, à titre principal, au producteur (personne agissant à titre professionnel en qualité de fabricant d'un produit fini ou d'une partie composante) ainsi qu'à des personnes que la loi assimile au producteur (importateur du produit dans l'Union européenne ; personne qui, agissant à titre professionnel, se présente comme producteur en apposant sur le produit son nom, sa marque ou un autre signe distinctif) et, à titre subsidiaire, dans le cas où le producteur ne peut pas être identifié, aux revendeurs professionnels du produit.

L'action fondée sur la responsabilité du fait des produits défectueux doit être exercée dans un délai de 3 ans, à compter du moment où la victime a connaissance ou aurait dû avoir connaissance du dommage, du défaut et de l'identité du producteur (délai de prescription), sans pouvoir être introduite au-delà de 10 ans après la mise en circulation du produit (délai de forclusion).

Le producteur, responsable de plein droit des dommages causés par le défaut de

sécurité du produit, ne peut s'exonérer qu'en invoquant l'une des causes d'exonération prévues par la loi (Code civil, articles 1245-10 et 1245-12) :

- le produit litigieux n'a pas été mis en circulation par le producteur ou n'a pas été fabriqué pour être mis sur le marché ;
- le défaut de sécurité n'existait pas au moment où le produit a été mis en circulation ;
- l'état des connaissances scientifiques et techniques le plus avancé au niveau mondial, au moment de la mise en circulation du produit, ne permettait pas de déceler le défaut (exonération pour « risque de développement ») ;
- le défaut de sécurité procède du respect par le producteur d'une règle impérative (ordre de la loi) ;
- la victime a commis une faute ayant contribué à la réalisation du dommage (exonération partielle).

Ce régime de responsabilité du fait des produits défectueux vient de faire l'objet d'une réforme. La directive européenne n° 2024/2853 du 23 octobre 2024 ambitionne de mieux adapter la responsabilité du fait des produits défectueux à l'intelligence artificielle. La transposition en droit interne de cette nouvelle directive doit intervenir au plus tard le 9 décembre 2026, le régime antérieur continuant de s'appliquer transitoirement aux produits mis sur le marché jusqu'à cette date.

Les apports de cette nouvelle législation seront exposés dans les réponses aux questions suivantes.

b) Comment les principes existants de la responsabilité du fait des produits pourraient-ils s'appliquer aux systèmes d'IA ? Devrait-on distinguer entre les systèmes d'IA et les produits intégrant l'IA ?

Lorsque l'IA est intégrée dans un support physique (robot, véhicule autonome), on admet aisément que l'objet complexe composé du support matériel et de l'IA incorporée constitue un produit, au sens de l'article 1245-2 du Code civil. La responsabilité du fait des produits défectueux trouve donc incontestablement à s'appliquer en pareil cas. Le producteur du système d'IA et celui de l'objet matériel incorporant l'IA seront même solidairement responsables d'un éventuel défaut de sécurité (Code civil, article 1245-7 et directive du 23 octobre 2024, article 8.1).

Mais un système d'IA désincarné peut-il, en l'absence de tout substrat tangible, être qualifié de produit ? Une réponse négative signifierait que la responsabilité du fait des produits défectueux ne s'applique pas aux systèmes d'IA désincarnés.

Des hésitations se sont manifestées à cet égard en doctrine, certains auteurs doutant que la notion de produit, au sens de l'article 1245-2 du Code civil, englobe les choses immatérielles.

Plusieurs éléments pouvaient cependant contribuer à réduire ce doute quant à la possibilité de considérer un système d'IA désincarné comme un produit, au regard de la législation sur la responsabilité du fait des produits défectueux :

- dès 1988, la Commission européenne indiquait que « *la directive s'applique aux logiciels* » (JOCE, 8 mai 1989, n° C 114/42) ;
- le règlement européen n° 2023/988 du 10 mai 2023 relatif à la sécurité générale des produits énonce que « *tout élément, qu'il soit ou non interconnecté avec d'autres éléments (...)* » constitue un produit ;
- en France, une réponse ministérielle expliquait que la responsabilité du fait des produits défectueux « *a vocation à englober l'intégralité de la catégorie juridique des meubles, à laquelle appartiennent les logiciels* » (réponse ministérielle n° 15677, JOAN Q, 24 août 1998, p. 4728) et le Code de la santé publique admet qu'un système d'IA conçu à des fins diagnostiques ou thérapeutiques entre dans la catégorie des produits de santé (articles L. 5211-1 et L. 5311-1).

Le doute (léger) qui pouvait subsister quant à la possibilité de considérer un système d'IA désincarné comme un produit, au regard de la responsabilité du fait des produits défectueux, est désormais totalement levé par la nouvelle directive du 23 octobre 2024 qui proclame *expressis verbis* que le terme « produit » comprend les logiciels (article 4.1).

La responsabilité du fait des produits défectueux a ainsi vocation à s'appliquer à tous les systèmes d'IA.

c) Comment définiriez-vous un "défaut" dans le contexte d'un système d'IA, en particulier dans les cas où le préjudice est causé par une décision prise par un système d'IA, plutôt que par un défaut traditionnel du produit ?

Aux termes de l'article 1245-3 du Code civil, un produit est défectueux « *lorsqu'il n'offre pas la sécurité à laquelle on peut légitimement s'attendre* ». En d'autres termes, le défaut de sécurité est un défaut empêchant le produit d'atteindre le niveau de sécurité auquel le grand public est en droit de s'attendre. Le défaut de sécurité peut être intrinsèque (défaut de conception ou de fabrication) ou extrinsèque (défaut d'information ou de mise en garde).

Appliquée à l'IA, cette définition générale du défaut de sécurité soulève des difficultés qui ont trait à sa portée, d'une part, et à sa mise en œuvre, d'autre part.

1) Difficulté relative à la portée de la notion de défaut de sécurité en matière d'IA

De la combinaison de l'article 1245-3 du Code civil énonçant, on l'a vu, qu'un produit est défectueux lorsqu'il n'offre pas la sécurité à laquelle le grand public peut légitimement s'attendre, et de l'article 1245-1 prévoyant que sont réparables, au titre de la responsabilité du fait des produits défectueux, d'une part, les atteintes à la personne, d'autre part les atteintes aux biens autres que le produit défectueux lui-même, il s'évince que le défaut de sécurité consiste fondamentalement en une dangerosité anormale du produit pour la santé des personnes ou l'intégrité matérielle des biens.

Si cette définition générale du défaut de sécurité s'ajuste assez aisément aux systèmes d'IA incarnés (cf. véhicule autonome, robot intelligent), dans la mesure où ce sont, pour l'essentiel, des dommages corporels ou matériels que ceux-ci sont susceptibles de provoquer, il n'en va pas de même s'agissant des systèmes d'IA désincarnés. Sans doute, ceux-ci peuvent-ils, à l'occasion, causer un dommage corporel (cf. système d'IA médicale fournissant un diagnostic erroné, suivi par le praticien, à l'origine du dommage corporel que le traitement inapproprié a pu engendrer). Mais, le plus souvent, les systèmes d'IA désincarnés occasionneront bien plutôt des préjudices économiques ou moraux qualifiés de « purs » parce qu'ils ne dérivent pas d'une atteinte primaire à l'intégrité physique d'une personne ou à l'intégrité matérielle d'un bien. Les exemples de préjudices économiques ou moraux purs provoqués par des systèmes d'IA désincarnés abondent : refus de prêt bancaire résultant d'une décision erronée du système d'IA utilisé par un établissement financier ; discrimination à l'embauche en raison d'un biais dans le fonctionnement du système d'IA utilisé par l'employeur pour procéder au recrutement ; atteinte à l'honneur ou à la réputation résultant de la diffusion d'une vidéo hypertruquée (deepfake) réalisée au moyen d'un système d'IA, pertes financières provoquées par une interprétation erronée des tendances du marché par le système d'IA...

Dès lors qu'ils ne sont pas, à titre principal, générateurs d'un danger anormal pour la santé des personnes ou l'intégrité des biens, les systèmes d'IA désincarnés s'avèrent par là-même assez largement réfractaires à la notion de défaut de sécurité, au sens de l'article 1245-3 du Code civil.

La directive du 23 octobre 2024 réformant la responsabilité du fait des produits défectueux n'apporte, sur ce point, aucune évolution. En excluant la réparation des préjudices économiques purs et des préjudices moraux purs, implicitement dans son article 6 et plus explicitement dans son considérant 24 (« *Les types de dommages autres que ceux prévus par la présente directive, tels que les pertes purement économiques, les atteintes à la vie privée ou la discrimination, ne devraient pas, par eux-mêmes, déclencher une responsabilité au titre de la présente directive* »), la nouvelle directive semble bien s'en tenir à une conception stricte du défaut de sécurité entendu comme un danger anormal pour la santé des personnes ou l'intégrité des biens, de sorte que les préjudices occasionnés par les systèmes d'IA désincarnés continueraient d'échapper, pour une large part, à la responsabilité du fait des produits défectueux.

Reste qu'il y a tout de même quelque paradoxe à reconnaître, à hauteur de principe, que les systèmes d'IA désincarnés entrent dans le champ d'application de la responsabilité du fait des produits défectueux (*supra* II.7 b), pour, en pratique, les en soustraire aussitôt, dans une large mesure, en excluant de ce régime les préjudices économiques et moraux purs qui sont précisément ceux que les systèmes d'IA désincarnés causent le plus volontiers. Cela incitera-t-il la loi de transposition ou la jurisprudence à promouvoir une conception plus large du défaut de sécurité et des préjudices réparables ? L'avenir le dira.

2) *Difficulté relative à la mise en œuvre de la notion de défaut de sécurité en matière*

d'IA

2.1) Constat

C'est à la victime qu'incombe la charge de la preuve du défaut de sécurité du produit. Même si la jurisprudence se contente, à cet égard, d'un faisceau d'indices graves, précis et concordants, ce fardeau probatoire risque de s'avérer particulièrement lourd pour les victimes de l'IA.

Hors quelques cas particuliers dans lesquels le défaut de sécurité d'un système d'IA s'évincerait de lui-même des circonstances de l'espèce (non-respect évident des standards de sécurité imposés notamment par le règlement européen sur l'IA ; insuffisance d'information sur les risques liés à l'utilisation du système), la preuve du défaut de sécurité d'un système d'IA sera, pour la victime, une tâche des plus épineuse en raison de l'autonomie et de l'opacité de fonctionnement des systèmes d'IA connexionnistes.

Parce qu'ils mettent en œuvre des algorithmes non-déterministes dont on ignore comment ils traitent la masse de données qui les alimente, parce qu'ils sont auto-apprenants et évolutifs, parce que les résultats auxquels ils parviennent s'expriment sous forme de prédictions probabilistes, les systèmes d'IA connexionnistes commettent inévitablement des erreurs. Cela est, en quelque sorte, inhérent à leur fonctionnement normal. Qu'un système d'IA ait fourni un résultat erroné à l'origine du dommage subi par la victime ne trahit donc pas, à soi seul, sa défaillance objective, ne révèle pas automatiquement un défaut de sécurité. Comment, dans ces conditions, la victime peut-elle rapporter la preuve d'un défaut de sécurité du système d'IA ?

Deux méthodes paraissent, de prime abord, envisageables.

La première peut être qualifiée de « quantitative ». Pour rapporter la preuve du défaut de sécurité, la victime devrait démontrer que le système d'IA litigieux atteint, de manière générale, un taux d'erreur anormal. Ce qui implique *primo*, de déterminer le taux d'erreur du système d'IA en cause ; *secundo*, d'établir qu'il est anormal. C'est dire que, pour la victime, la difficulté se dédouble. D'abord, la détermination du taux effectif d'erreur du système d'IA litigieux risque de s'avérer malcommode. Ensuite, pour caractériser l'anormalité du taux d'erreur du système d'IA, il faut nécessairement comparer celui-ci à un étalon de référence. Mais lequel ? Les déclarations du producteur à cet égard ? Mais *quid* si celui-ci s'est abstenu d'annoncer lui-même un taux d'erreur (par hypothèse bas) ? Le comportement d'un humain normalement compétent dans l'activité concernée ? Mais *quid* si le système d'IA accomplit une tâche que l'humain n'est pas en mesure d'exécuter ? Les performances d'un modèle-type de système d'IA comparable ? Mais sur quelles bases tangibles déterminer celles-ci ?

La seconde méthode, qu'on qualifiera, par contraste, de « qualitative » consiste à exiger de la victime qu'elle rapporte spécifiquement la preuve d'un défaut de conception de l'algorithme ou d'une faille dans la constitution de la base de données alimentant le système d'IA ou encore d'un mauvais apprentissage du système... Mais

cette tâche risque fort de s'avérer diabolique – même avec l'aide d'experts en informatique – dès lors que les ressorts profonds du fonctionnement des systèmes d'IA demeurent hermétiques (l'effet « boîte noire »).

Au fond, « *le critère de défektivité sied en définitive mieux au dommage causé par des produits de série qui sont appelés à tous faire l'objet d'un usage similaire qu'à des objets intelligents appelés à interagir avec leur environnement, voire à apprendre seuls et à s'améliorer au contact de l'utilisateur* » (A. Bonnet, La responsabilité du fait de l'intelligence artificielle, mémoire Paris 2, 2015, p. 25).

2.2) Remèdes

Afin de faciliter la mise en œuvre de la notion de défaut de sécurité en matière d'IA, la directive européenne du 23 octobre 2024 met en place trois mesures (qui ne s'appliqueront toutefois qu'aux systèmes d'IA mis sur le marché après le 9 mars 2026).

Primo, le nouveau texte prévoit que, pour apprécier le défaut de sécurité d'un système d'IA, il faut tenir compte de « *l'effet sur le produit de sa capacité à poursuivre son apprentissage (...) après sa mise sur le marché* » (article 7.2 c). Cela signifie que l'apprentissage en continu du système d'IA postérieurement à sa mise sur le marché est une donnée que le juge devra prendre en considération pour admettre plus aisément le défaut de sécurité.

Secundo, la nouvelle directive met en place une procédure d'injonction de divulgation d'éléments de preuve (article 9). Sur sollicitation de la victime, dès lors que celle-ci présente des éléments rendant plausible sa demande en réparation, le juge devra ordonner au défendeur de divulguer les éléments de preuve pertinents dont il dispose (dans la mesure de ce qui est nécessaire et proportionné, au regard notamment des informations confidentielles et des secrets d'affaires). A vrai dire, cette mesure ne présente, au regard du droit français, qu'un intérêt limité dans la mesure où l'article 11, alinéa 2, du Code de procédure civile permet déjà au juge d'enjoindre au défendeur de divulguer les éléments de preuve dont il dispose.

Tertio, le défaut de sécurité sera, dans certains cas, présumé.

D'abord, le défaut de sécurité sera présumé si le défendeur ne respecte pas l'injonction de divulgation de preuves que le juge lui a adressée (article 10.2 a).

Ensuite, le dysfonctionnement manifeste d'un produit utilisé dans des conditions normales permet de présumer le défaut de sécurité (article 10.4 c). Ainsi, par exemple, le dysfonctionnement manifeste d'un système d'IA gérant la distribution de l'eau dans une grande exploitation agricole entraînant l'inondation de plusieurs champs et la destruction des cultures pourrait faire présumer le défaut de sécurité.

Enfin, lorsqu'en raison de la complexité technique ou scientifique du produit – ce qui est le cas des systèmes d'IA –, la preuve du défaut de sécurité s'avère excessivement difficile, celui-ci pourra être présumé (article 10.4). Cette présomption de défaut de sécurité est cependant subordonnée à la condition que la victime ait préalablement établi « *qu'il est probable que le produit est défectueux* » (article 10.4 b). Compte-

tenu de l'asymétrie d'information existant, en matière d'IA, entre la victime et le défendeur, cette condition préalable pourrait représenter un obstacle notable au déclenchement de la présomption. L'intérêt pratique de celle-ci pour les victimes de l'IA dépendra donc, dans une large mesure, de l'appréciation plus ou moins souple que les juges feront de cette condition d'application de la présomption.

d) Dans le cadre de la responsabilité du fait des produits, comment devrait-on traiter les mises à jour logicielles ou les changements dans les données d'apprentissage qui modifient le comportement d'un système d'IA après sa mise sur le marché ?

Les articles 1245 et suivants du Code civil, qui régissent actuellement la responsabilité du fait des produits défectueux, n'envisagent pas l'IA de manière spécifique. On ne s'étonnera donc pas qu'ils soient totalement muets sur la façon dont il convient de traiter les mises à jour logicielles ou les changements dans les données d'apprentissage entraînant une modification du comportement d'un système d'IA postérieure à sa mise sur le marché. En revanche, la directive n° 2024/2853 du 23 octobre 2024 réformant la responsabilité du fait des produits défectueux dans le but, notamment, de l'adapter à l'IA traite de cette question avec précision.

Le nouveau texte (qui, rappelons-le ne s'appliquera qu'aux systèmes d'IA mis sur le marché après le 9 mars 2026) invite à distinguer selon que la modification du système d'IA qu'entraînent les mises à jour logicielles ou les changements dans les données d'apprentissage a un caractère substantiel ou non, étant précisé qu'une modification substantielle est celle qui « *modifie la performance d'origine, la destination d'origine ou le type d'origine du produit, sans que cette modification ait été prévue dans l'évaluation initiale des risques effectuée par le fabricant* » (article 4.18 b).

1) Première hypothèse : les mises à jour logicielles ou les changements dans les données d'apprentissage entraînent une modification substantielle du système d'IA postérieure à sa mise sur le marché.

Un système d'IA faisant l'objet d'une modification substantielle doit être considéré comme un nouveau produit mis sur le marché (cf. considérant 58 : « *les produits substantiellement modifiés sont essentiellement des produits nouveaux* »). La directive du 23 octobre 2024 en tire une double conséquence.

La première a trait au délai pour agir sur le fondement de la responsabilité du fait des produits défectueux. La nouvelle directive européenne conserve le système du double délai actuellement en vigueur, en distinguant un délai de prescription (article 16) et un délai de forclusion (article 17) :

- délai de prescription : l'action fondée sur la responsabilité du fait des produits défectueux doit être exercée dans un délai de 3 ans, à compter du moment où la victime a connaissance ou aurait dû avoir connaissance du dommage, du défaut et de l'identité du producteur ;
- délai de forclusion : l'action n'est, en tout état de cause, plus recevable au-delà de 10 ans après que le produit a été mis sur le marché (le délai de forclusion est porté à 25

ans lorsque la défektivité du produit est à l'origine d'une lésion corporelle dont les symptômes sont d'apparition lente, ce qui pourra s'avérer utile notamment pour les victimes de systèmes d'IA médicaux).

C'est le délai de forclusion qui nous intéresse ici. Son point de départ est normalement fixé à la date à laquelle le produit a été mis sur le marché (la mise sur le marché étant définie par l'article 4.8 comme « *la première mise à disposition d'un produit sur le marché de l'Union* »). Mais « *lorsqu'une modification substantielle est apportée au moyen d'une mise à jour ou d'une mise à niveau logicielle, ou en raison de l'apprentissage continu d'un système d'IA, le produit substantiellement modifié devrait être considéré comme étant mis à disposition sur le marché ou mis en service au moment où la modification est effectivement effectuée* » (considérant 40). La modification substantielle d'un produit préexistant étant considérée comme la mise sur le marché d'un nouveau produit, elle constitue fort logiquement le point de départ d'un nouveau délai de forclusion de 10 ans (ou de 25 ans, en cas de lésion corporelle d'apparition lente).

La seconde conséquence concerne l'exonération du producteur. Dans la liste légale des causes d'exonération ouvertes au producteur, l'une de celles que les producteurs de systèmes d'IA pourraient être plus particulièrement enclins à invoquer est celle qui consiste à soutenir que le défaut de sécurité n'existait pas au moment où le produit a été mis sur le marché (Code civil, article 1245-10.2° et directive du 23 octobre 2024, article 11.1 c). Mais lorsque le défaut de sécurité est dû à une modification substantielle du système d'IA (laquelle peut résulter d'une mise à jour logicielle ou d'un changement des données d'apprentissage), le producteur ne pourra pas s'exonérer en alléguant la postériorité du défaut par rapport à la date initiale de mise sur le marché (directive du 23 octobre 2024, article 11.2 d) : la modification substantielle correspondant à la mise sur le marché d'un nouveau produit, on se trouve, en effet, en présence d'un nouveau système d'IA qui était déjà défectueux au moment où il a été mis sur le marché.

2) *Seconde hypothèse : les mises à jour logicielles ou les changements dans les données d'apprentissage entraînent une modification non-substantielle du système d'IA postérieure à sa mise sur le marché.*

La directive du 23 octobre 2024 interdit au producteur de se prévaloir de l'exonération résultant de la postériorité du défaut de sécurité par rapport à la date de mise sur le marché lorsque le défaut est dû à un logiciel qui est resté sous son contrôle après la mise sur le marché (article 11.2 b) et, précise le considérant 50 de la directive, « (...) *étant donné que les technologies numériques permettent aux fabricants d'exercer un contrôle après le moment de la mise sur le marché (...), il convient que les fabricants restent responsables de toute défektivité apparue après ce moment du fait des logiciels (...)* qu'ils contrôlent, que ce soit sous la forme de mises à niveau ou de mises à jour ou bien d'algorithmes d'apprentissage automatique. Ces logiciels (...) devraient être considérés comme étant sous le contrôle du fabricant (...) ». La faculté d'exonération tirée de la postériorité du défaut de sécurité par rapport à la date de mise sur le marché du produit se trouvera donc neutralisée en matière d'IA tant que le producteur du système conserve un contrôle sur celui-ci après sa mise sur

le marché (que ce soit par la fourniture de mises à jour logicielles ou par la mise en place d'un algorithme d'apprentissage en continu).

e) Comment les concepts de 'état des connaissances scientifiques' et du 'risque de développement' devraient-ils être appliqués aux systèmes d'IA dans le contexte de la responsabilité du fait des produits ?

L'exonération pour « risque de développement » permet au producteur d'échapper à sa responsabilité en rapportant la preuve que l'état des connaissances scientifiques et techniques le plus avancé au niveau mondial, au moment où il a mis le produit sur le marché, ne permettait pas de déceler le défaut de sécurité (Code civil, article 1245-10. 4° ; directive du 23 octobre 2024, article 11.1 e).

L'opacité de fonctionnement des systèmes d'IA ne peut qu'inciter les producteurs à invoquer cette cause d'exonération, en soutenant que le défaut était absolument indétectable au moment de la mise sur le marché.

La directive du 13 octobre 2024 réformant la responsabilité du fait des produits défectueux devrait cependant entraîner, à l'avenir, un recul de l'exonération pour risque de développement s'agissant des dommages causés par des systèmes d'IA défectueux.

En premier lieu, si la nouvelle directive maintient l'exonération pour risque de développement, elle permet à chaque État-membre de l'UE de l'écarter, dans la loi de transposition, pour certaines catégories spécifiques de produits (article 18). Le législateur français pourrait donc fort bien interdire l'exonération pour risque de développement en ce qui concerne les systèmes d'IA, conformément au souhait formulé par de nombreux auteurs. Il apparaît néanmoins peu probable que la loi française de transposition de la directive supprime purement et simplement l'exonération pour risque de développement en matière d'IA : la justification de celle-ci résidant dans la protection de l'innovation, le législateur français renoncera vraisemblablement à la faire disparaître, de crainte de nuire à la compétitivité des acteurs français de l'IA face à leurs concurrents étrangers.

En second lieu, si le législateur français fait le choix de conserver l'exonération pour risque de développement en matière d'IA, il est à noter qu'une restriction sera tout de même apportée à celle-ci. Jusqu'à présent, le fabricant était exonéré dès lors que le défaut de sécurité était absolument indécélable au moment de la mise en circulation. Désormais, pour que le fabricant s'exonère, il faudra non seulement que le défaut de sécurité soit absolument indécélable au moment de la mise sur le marché mais, de surcroît, qu'il le demeure après, tant que le produit reste sous son contrôle (article 11.1 e), c'est-à-dire concrètement tant que le fabricant du système d'IA fournit des mises à jour ou, si le système est doté d'une faculté d'apprentissage en continu, tant que cet auto-apprentissage perdure.

III. RESOLUTION DES SCENARIOS HYPOTHETIQUES

Pour conclure ce questionnaire, veuillez expliquer comment chacun des scénarios hypothétiques A~E présentés précédemment serait probablement résolu dans votre juridiction. Si la résolution était similaire pour plusieurs scénarios, vous pouvez les regrouper dans votre réponse.

Scénario A : Erreur de diagnostic médical par IA (Préjudice corporel)

Un système de diagnostic alimenté par l'IA ne détecte pas un cancer traitable à un stade précoce sur le scan d'un patient, entraînant un retard de traitement et des complications de santé importantes.

La victime peut envisager de demander réparation de son préjudice soit à l'établissement ou au professionnel de santé qui s'est servi du système d'IA d'aide au diagnostic médical (1) soit au producteur de celui-ci (2).

1) Action en responsabilité contre l'établissement ou le professionnel de santé utilisateur du système d'IA

Deux hypothèses doivent être distinguées. Si la victime a été soignée dans le secteur privé (clinique privée, professionnel de santé privé), son action en réparation relève de la compétence des juridictions judiciaires (avec à leur sommet la Cour de cassation). En revanche, si la victime a été soignée dans un hôpital public, son action en réparation relève de la compétence des juridictions administratives (avec à leur sommet le Conseil d'État). Or, les deux ordres de juridictions adoptent des positions divergentes au sujet de la responsabilité médicale à raison des produits de santé utilisés à des fins diagnostiques ou thérapeutiques.

1.1) Première hypothèse : le système d'IA est utilisé par une clinique privée ou un professionnel de santé privé

Selon la Cour de cassation, fidèle au principe selon lequel la responsabilité médicale repose sur la faute (Code de la santé publique, article L. 1142-1. I), la responsabilité des cliniques privées ou des professionnels de santé privés, à raison des produits et dispositifs de santé qu'ils utilisent, ne peut être engagée que sur la base d'une faute prouvée. La victime dont la tumeur n'a pas été détectée par l'IA devra donc rapporter la preuve d'une faute commise par le défendeur. En quoi peut consister cette faute ? Elle ne saurait résider dans le fait d'avoir eu recours à un système d'IA d'aide au diagnostic médical, le médecin devant « *toujours élaborer son diagnostic avec le plus grand soin (...) en s'aidant dans toute la mesure du possible des méthodes scientifiques les mieux adaptées* » (Code de la santé publique, article R. 4127-33).

Aux termes de l'article L. 4001-3. I du Code de la santé publique, « *[l]e professionnel de santé qui décide d'utiliser, pour un acte de prévention, de diagnostic ou de soin, un dispositif médical comportant un traitement de données algorithmique dont l'apprentissage a été réalisé à partir de données massives s'assure que la personne concernée en a été informée et qu'elle est, le cas échéant, avertie de l'interprétation*

qui en résulte ». Le non-respect de cette obligation spécifique d'information constituerait une faute.

Si la victime a été dûment informée du recours à l'IA pour le dépistage de son cancer, la faute médicale ne pourrait plus guère consister qu'en un contrôle insuffisant du résultat négatif auquel est parvenu, en l'espèce, le système d'IA. Or, la preuve d'une telle faute technique ne sera pas aisée car plus la fiabilité générale du système d'IA utilisé sera élevée, moins le juge sera enclin à retenir une faute à l'encontre du praticien qui a suivi le résultat proposé par le système.

1.2) Seconde hypothèse : le système d'IA est utilisé par un hôpital public

A l'opposé de la Cour de cassation, le Conseil d'État met à la charge des hôpitaux publics une responsabilité de plein droit du fait des produits de santé qu'ils utilisent lorsque la défaillance de ceux-ci occasionne un dommage au patient : « *le service public hospitalier est responsable, même en l'absence de faute de sa part, des conséquences dommageables pour les usagers de la défaillance des produits et appareils de santé qu'il utilise* » (voir notamment, CE, 10 juillet 2024, n° 479613). Or, l'article L. 5211-1 du Code de la santé publique range les systèmes d'IA conçus à des fins médicales dans la catégorie des produits de santé. Il s'ensuit que l'hôpital public est responsable de plein droit des dommages causés au patient par la défaillance d'un système d'IA médical qu'il utilise. En l'espèce, si la victime est soignée dans un hôpital public, elle pourra donc assigner celui-ci devant le juge administratif afin d'obtenir réparation sur le fondement de ce régime de responsabilité objective. Mais encore faudra-t-il que la victime parvienne à établir que l'erreur commise par le système d'IA utilisé par l'hôpital public constitue une « défaillance », au sens de la jurisprudence administrative.

2) Action en responsabilité contre le producteur du système d'IA

Si la victime choisit de demander réparation au producteur du système d'IA, c'est préférentiellement sur le fondement de la responsabilité du fait des produits défectueux qu'elle agira (2.1). Une action fondée sur la responsabilité civile de droit commun pourrait également être envisagée (2.2).

2.1) Responsabilité du fait des produits défectueux

Une action intentée contre le producteur du système d'IA d'aide à la décision médicale, sur le fondement de la responsabilité du fait des produits défectueux, risque de se heurter à un obstacle probatoire. Celui-ci a trait au défaut de sécurité du système d'IA litigieux, dont il appartient à la victime de rapporter la preuve.

Un système d'IA d'aide au diagnostic médical ne se contente pas d'exécuter automatiquement des instructions préprogrammées ; il apprend, grâce à un algorithme basé sur un réseau de neurones artificiels, à traiter une masse importante de données en cherchant des similitudes, des corrélations, des récurrences afin de dégager un résultat par induction probabiliste. Or, ce mode de fonctionnement génère inévitablement des erreurs (cf. *supra* II.7 c). Le simple constat que le système

d'IA litigieux a, en l'espèce, commis une erreur en ne détectant pas un cancer précoce ne saurait donc suffire à établir le défaut de sécurité. Il faut qu'au-delà, la victime établisse que le système d'IA litigieux ne présente pas la sécurité à laquelle on peut légitimement s'attendre, autrement dit, qu'il présente un danger anormal pour la santé des patients. Mais comment rapporter cette preuve ?

Suffit-il d'établir qu'un radiologue normalement compétent aurait détecté la tumeur dont la victime est, en l'espèce, atteinte ? Il est permis de douter que cela suffise à caractériser le défaut de sécurité car, dans le même temps, le système d'IA peut être capable de détecter des tumeurs qu'un humain n'aurait pas décelées.

Puisqu'en l'espèce, le système d'IA litigieux constitue un produit de santé, on peut songer à solliciter la technique du rapport bénéfices/risques, volontiers utilisée en ce domaine. Dans cette perspective, il appartiendrait à la victime d'établir que le système d'IA d'aide au diagnostic médical utilisé en l'espèce présente un rapport bénéfices/risques défavorable. Mais pour ce faire, faut-il mettre en balance le nombre de diagnostics erronés réalisés par le système d'IA par rapport au nombre de diagnostics exacts qu'il produit, ou bien, de manière plus pointue, le nombre de tumeurs non détectées par le système d'IA qu'un humain aurait décelées par rapport au nombre de tumeurs qu'il a détectées et qu'un humain n'aurait pas vues ? Et, en tout état de cause, il faudrait ensuite établir que le ratio auquel on parvient est suffisamment défavorable au système d'IA pour que l'on puisse considérer que celui-ci présente un danger anormal pour la santé des patients. On mesure aisément la complexité de la tâche de la victime.

Faut-il plutôt que la victime rapporte précisément la preuve d'un défaut de conception de l'algorithme ou d'une faille dans la constitution de la base de données alimentant le système d'IA ou d'un mauvais apprentissage du système... ? Mais elle risque alors de se heurter à l'opacité de fonctionnement des systèmes d'IA contemporains, qui est telle que même leurs concepteurs ne sont pas en mesure de comprendre le processus suivi par l'IA jusqu'au résultat proposé.

Sans doute, la directive européenne du 23 octobre 2024 réformant la responsabilité du fait des produits défectueux met-elle en place différents procédés destinés à alléger la tâche probatoire des victimes de l'IA, en instaurant notamment, à certaines conditions, des présomptions de défaut de sécurité (*supra* II.7 c) mais elle ne s'appliquera qu'aux produits mis sur le marché après le 9 décembre 2026.

2.2) Responsabilité civile de droit commun

La victime dont la tumeur n'a pas été détectée précocement à la suite de l'erreur du système d'IA n'augmentera pas ses chances d'indemnisation si elle agit contre le producteur de celui-ci sur le fondement de la responsabilité civile de droit commun.

Une action fondée sur la responsabilité pour faute serait certes recevable mais elle se heurterait à un obstacle probatoire encore supérieur à celui que soulève le défaut de sécurité, en ce qu'il incomberait à la victime non seulement de rapporter la preuve d'une faute commise par le producteur mais également d'établir que cette faute est distincte d'un défaut de sécurité du système d'IA (*supra* II.2 b).

Quant à une action fondée sur la responsabilité générale du fait des choses, elle est exclue car la jurisprudence interdit formellement aux victimes d'agir contre les producteurs sur le fondement de la responsabilité générale du fait des choses (voir Cass. 1^{re} civ., 26 oct. 2022, n° 20-23425).

Scénario B - Dysfonctionnement d'un système d'irrigation contrôlé par IA (Préjudice matériel)

Un système d'IA gérant la distribution d'eau dans une grande exploitation agricole dysfonctionne, inondant plusieurs champs et détruisant les cultures. Le dysfonctionnement est attribué à une erreur dans l'interprétation par l'IA des données des capteurs d'humidité du sol.

L'agriculteur victime de la destruction des cultures provoquée par le système d'IA gérant la distribution d'eau dans son exploitation peut envisager d'agir en réparation contre le producteur du système d'IA sur le fondement de la responsabilité du fait des produits défectueux (1) ou contre le revendeur professionnel du système d'IA sur le fondement de la garantie des vices cachés (2).

1) Action en réparation dirigée contre le producteur du système d'IA sur le fondement de la responsabilité du fait des produits défectueux

Pour obtenir réparation de son préjudice matériel (destruction de cultures), la victime peut agir contre le producteur du système d'IA, sur le fondement de la responsabilité du fait des produits défectueux. Celle-ci requiert la réunion de trois conditions : un défaut de sécurité du système d'IA, un dommage réparable au titre de la responsabilité du fait des produits défectueux et un lien de causalité entre les deux. Reprenons ces trois conditions.

En premier lieu, la victime doit rapporter la preuve que le système d'IA litigieux présente un défaut de sécurité, autrement dit qu'il n'offre pas la sécurité à laquelle on peut légitimement s'attendre. Plus précisément, il s'agira, en l'espèce, d'établir que le système d'IA gérant l'irrigation des terres créait un danger anormal pour les cultures agricoles.

Dans la réponse au scénario A, on a souligné combien la preuve du défaut de sécurité d'un système d'IA était difficile pour la victime. Mais, en l'espèce, celle-ci peut nourrir l'espoir de vaincre cet obstacle probatoire. L'exactitude des données transmises par les capteurs d'humidité du sol n'est, en l'occurrence, pas contestée. C'est donc bel et bien l'erreur d'interprétation de ces données par le système d'IA qui est à l'origine du dommage. Cette erreur, par l'ampleur de ses conséquences (inondation de plusieurs champs et destruction des cultures) constitue une erreur qu'on peut qualifier de grossière. Or, cette erreur grossière peut être considérée comme révélatrice d'un dysfonctionnement manifeste du système d'IA, d'où peut être raisonnablement inféré le défaut de sécurité (en ce sens, on relèvera que l'article 10 de la directive européenne du 23 octobre 2024, admet que le dysfonctionnement manifeste du produit fait présumer sa défectuosité). Il serait également possible de voir dans cette erreur grossière le signe d'un mauvais apprentissage de l'IA dans l'interprétation des données qu'elle doit traiter, lequel pourrait pareillement dénoter le défaut de

sécurité.

En deuxième lieu, en supposant que la preuve du défaut de sécurité soit rapportée, la victime doit justifier d'un dommage réparable au titre de la responsabilité du fait des produits défectueux. En l'espèce, le dommage allégué consiste en la destruction des cultures. Il s'agit donc d'un dommage matériel. Aux termes de l'article 1245-1, alinéa 2, du Code civil, la responsabilité du fait des produits défectueux permet à la victime d'obtenir réparation « d'une atteinte à un bien autre que le produit défectueux lui-même ». Tel est le cas, en l'occurrence. Cependant, les textes européens d'où procède la responsabilité du fait des produits défectueux ne visent que les dommages causés à des biens destinés à la consommation privée (directive n° 85/374 du 25 juillet 1985, article 9.b ; directive n° 2024/2853 du 23 octobre 2024, article 6.b). Or, en l'espèce, les cultures agricoles détruites sont des biens à destination professionnelle (elles avaient vocation à être vendues par la victime dans le cadre de son activité professionnelle). Il ne faudrait pourtant pas en déduire que la destruction des récoltes ne pourra pas être indemnisée sur le fondement de la responsabilité du fait des produits défectueux. Le législateur français, dans la loi transposant en droit interne la directive européenne, n'a, en effet, pas repris cette exclusion des dommages causés aux biens à usage professionnel, l'article 1245-1, alinéa 2, du Code civil énonçant que sont réparables les atteintes aux biens (autre que le produit défectueux lui-même), sans distinguer aucunement selon leur destination. Cet élargissement des dommages matériels réparables a au demeurant reçu l'aval de la Cour de justice de l'Union européenne, qui a admis que les lois de transposition nationale peuvent inclure les dommages causés aux biens à usage professionnel dans le champ de la responsabilité du fait des produits défectueux (CJUE 4 juin 2009, n° C-285/08).

Le dommage matériel allégué par la victime est donc indemnisable au titre de la responsabilité du fait des produits défectueux, sous déduction cependant de la franchise de 500 € prévue par l'article 1245-1, alinéa 2, du Code civil (laquelle disparaîtra lorsque la nouvelle directive du 23 octobre 2024, qui l'abolit, aura été transposée en droit interne).

En troisième lieu, il appartient à la victime de rapporter la preuve d'un lien de causalité entre le défaut de sécurité et le dommage. A cet égard, le producteur du système d'IA pourrait, en l'espèce, être tenté d'invoquer une faute de la victime ayant pour effet, si ce n'est de rompre, du moins de réduire le lien causal entre le défaut de sécurité et le dommage. En effet, aux termes de l'article 1245-12 du Code civil, « La responsabilité du producteur peut être réduite ou supprimée, compte tenu de toutes les circonstances, lorsque le dommage est causé conjointement par un défaut du produit et par la faute de la victime ou d'une personne dont la victime est responsable ». La faute de la victime pourrait, en l'occurrence, consister à n'avoir pas suffisamment surveillé le fonctionnement du système d'IA et à ne l'avoir pas stoppé à temps, avant que plusieurs champs ne soient inondés et les cultures détruites. Le dommage final (destruction des récoltes) résulterait ainsi conjointement du défaut de sécurité du système d'IA (qu'on suppose avéré) et de la faute de la victime. Il n'est pas sûr cependant que le juge retienne ici une faute à l'encontre de la victime de nature à réduire son indemnisation. La Cour de cassation se montre, en effet, très

restrictive dans l'admission de la faute de la victime d'un produit défectueux, puisqu'elle ne lui reconnaît un effet exonératoire (total ou partiel) que lorsqu'elle a contribué à l'apparition même du dommage et non lorsqu'elle l'a simplement aggravé (Cass. 1^{re} civ., 2 juin 2021, n° 19-19349). Or, en l'espèce, la faute de la victime n'a-t-elle pas contribué seulement à l'aggravation du dommage, dont l'élément déclencheur a été le dysfonctionnement du système d'IA ?

2) Action en réparation dirigée contre le revendeur professionnel du système d'IA sur le fondement de la garantie des vices cachés

En l'espèce, c'est au moyen d'un contrat que l'agriculteur a obtenu le système d'irrigation de ses terres. Mais de quel type de contrat s'agit-il ? L'énoncé suggère que ce qui a été mis à la disposition de l'agriculteur n'est pas uniquement un logiciel, un programme informatique désincarné (auquel cas on serait, selon la doctrine majoritaire, en présence d'un contrat *sui generis*). C'est tout un système d'irrigation qui lui a été fourni avec des éléments matériels (notamment les capteurs d'humidité) et un élément immatériel (le logiciel pilotant le système). Il y a donc tout lieu de penser que le système d'irrigation intelligent de l'exploitation agricole a fait l'objet d'un contrat de vente. Dans l'hypothèse où l'agriculteur aurait acheté le système d'IA non pas directement au producteur mais à un revendeur professionnel, on pourrait s'interroger sur l'éventualité d'une action en réparation dirigée contre celui-ci, sur le fondement de la garantie des vices cachés.

Pour qu'une telle action soit envisageable, encore faut-il qu'un vice caché du système d'IA vendu soit caractérisé. Le vice caché s'entend d'un défaut inhérent à la chose qui la rend inapte à son usage normal et qui ne s'est révélé qu'après la conclusion du contrat de vente (Code civil, article 1641). En l'espèce, le fait que le dispositif d'irrigation acquis par l'agriculteur ait provoqué l'inondation de plusieurs champs et détruit ses cultures ne révèle-t-il l'inaptitude à son usage normal, lequel est précisément d'assurer une irrigation raisonnée des sols ?

La victime pourrait alors agir en réparation contre le revendeur professionnel du système d'IA vicié, sur le fondement de l'article 1645 du Code civil, aux termes duquel « Si le vendeur connaissait les vices de la chose, il est tenu (...) de tous les dommages et intérêts envers l'acheteur ». La Cour de cassation juge, en effet, de façon constante, qu'« il résulte de ce texte une présomption irréfragable de connaissance par le vendeur professionnel du vice de la chose vendue, qui l'oblige à réparer l'intégralité de tous les dommages qui en sont la conséquence » (Cass. com., 29 janvier 2025, n° 23-17954).

Scenario C - Erreur de jugement d'un conseiller financier IA (Préjudice économique)

Un algorithme d'investissement IA recommande une stratégie à haut risque basée sur une interprétation erronée des tendances du marché, entraînant des pertes financières substantielles pour ses clients.

En l'espèce, les pertes financières subies par la victime ne résultent ni d'une atteinte à l'intégrité physique d'une personne ni d'une atteinte à un bien matériel. Elles constituent donc ce qu'on nomme un préjudice économique pur (*supra* II.7 c).

Compte-tenu de la nature du préjudice subi, il paraît hasardeux d'agir sur le fondement de la responsabilité du fait des produits défectueux contre le producteur du système d'IA dont l'erreur dans l'interprétation des tendances du marché est à l'origine des pertes financières. Il est, en effet, douteux que la réparation d'un tel préjudice économique pur entre dans le périmètre de ce régime de responsabilité. Trois arguments peuvent être invoqués en ce sens.

Primo, l'article 1245-1 du Code civil limite les dommages réparables aux atteintes à la personne et aux atteintes aux biens (matériels) autres que le produit défectueux lui-même.

Secundo, la nouvelle directive du 23 octobre 2024 réformant la responsabilité du fait des produits défectueux exclut la réparation du préjudice économique pur, implicitement dans son article 6 (qui, outre les atteintes à l'intégrité physique des personnes ou à l'intégrité matérielle des biens, ne vise que la destruction ou corruption de données) et expressément dans son considérant 24 (« *les pertes purement économiques (...) ne devraient pas (...) déclencher une responsabilité au titre de la présente directive* »).

Tertio, la Cour de cassation ne paraît pas disposée à indemniser, au titre de la responsabilité du fait des produits défectueux, un préjudice économique qui ne dérive en aucune manière d'une atteinte primaire causée par la défectuosité du produit à l'intégrité physique d'une personne ou à l'intégrité matérielle d'un bien autre que le produit défectueux lui-même (cf. Cass. 1^{re} civ., 14 oct. 2015, n° 14-13847 ; Cass. 1^{re} civ., 9 décembre 2020, n° 19-21390 ; Cass. 1^{re} civ. 25 mai 2023, n° 21-23174).

C'est plutôt sur le terrain de la responsabilité contractuelle que la victime devrait se placer.

On ignore, en l'espèce, si le client a traité avec une personne physique ou morale (conseiller en investissements financiers, société de gestion de portefeuille) utilisant l'IA pour accomplir sa prestation ou si elle a interagi directement, en ligne, avec un « robo-advisor » (robot conseiller). Mais, dans les deux hypothèses, un contrat a bel et bien été conclu :

- dans la première, le contrat a été passé entre le client et le professionnel auquel celui-ci s'est adressé ;
- dans la seconde, le contrat a été conclu entre le client et l'exploitant du « robo-advisor » (qui a au demeurant, le plus souvent, le statut juridique de conseiller en investissements financiers ou, plus rarement, de société de gestion de portefeuille).

La réparation des pertes financières subies peut donc peut, en tout état de cause, être recherchée sur le terrain de la responsabilité contractuelle du débiteur du service d'investissements financiers, utilisateur de l'IA.

Pour obtenir réparation, le client va devoir démontrer que son préjudice (pertes financières) résulte de la mauvaise exécution par le prestataire cocontractant de son

obligation contractuelle.

En droit français, la preuve de l'inexécution d'une obligation contractuelle s'opère différemment selon que l'obligation litigieuse est une obligation de résultat (le débiteur s'engage à fournir un résultat précis) ou une obligation de moyens (le débiteur promet simplement de faire preuve de diligence en vue d'un objectif souhaité mais non garanti). Si l'obligation contractuelle est de résultat, le simple constat objectif de la non-réalisation du résultat promis engage de plein droit la responsabilité du débiteur, qui ne peut s'exonérer totalement qu'en rapportant la preuve d'une cause étrangère présentant les caractères de la force majeure. Si l'obligation contractuelle est de moyens, le créancier doit rapporter la preuve d'une faute commise par le débiteur, en établissant que celui-ci n'a pas déployé le niveau de diligence requis.

En l'espèce, le conseiller en investissements financiers ou la société de gestion de portefeuille utilisant l'IA pour exécuter sa prestation contractuelle est débiteur d'une simple obligation de moyens (cf. Code monétaire et financier, articles L. 541-8-1, L. 533-22-2-1 et L. 533-11, aux termes desquels les professionnels fournissant un service d'investissements financiers doivent agir « *d'une manière honnête, loyale et professionnelle, servant au mieux les intérêts des clients* »). En conséquence, pour être indemnisée de ses pertes financières, la victime devra rapporter la preuve d'une faute commise par le débiteur qui n'a pas fait preuve de la diligence qu'on est en droit d'attendre d'un bon prestataire de services d'investissement. Concrètement, il lui faudra démontrer que son cocontractant n'a pas vérifié suffisamment les recommandations du système d'IA dont il se sert, alors même qu'elles étaient à haut risque et entachées d'une erreur d'interprétation des tendances du marché, ou bien encore qu'il n'a pas mis en place les « coupe-circuits » requis en matière de « trading algorithmique » en cas de chute brutale des cours...

Scénario D - Vidéo *deepfake* générée par IA (Préjudice moral)

Un système d'IA crée une vidéo très convaincante mais fautive d'une personne se livrant à un comportement scandaleux. La diffusion virale de la vidéo cause de graves dommages à la réputation et un traumatisme émotionnel à la personne représentée.

En l'espèce, une personne a réalisé un hypertrucage vidéo (*deepfake*) en utilisant un système d'IA. Ce *deepfake* a été mis en ligne sur internet. On supposera, en l'absence d'indication contraire, que c'est l'auteur même de la vidéo hypertruquée qui l'a diffusée. La diffusion virale de ce *deepfake* cause un préjudice moral à la personne représentée dans la vidéo car elle y apparaît en train de se livrer à un comportement scandaleux, ce qui est de nature à porter atteinte à sa réputation, à son honneur.

La victime pourra aisément rechercher la responsabilité de l'auteur (et diffuseur) du *deepfake* (1). Celle de l'hébergeur du *deepfake* sera, en revanche, plus difficile à engager (2).

1) Responsabilité de l'auteur (et diffuseur) du deepfake

Il n'est pas douteux que le comportement de l'auteur du deepfake soit, en l'occurrence, fautif. Il contrevient, en effet, à plusieurs dispositions légales.

En premier lieu, reproduire, comme le fait le deepfake litigieux, l'image d'une personne sans son consentement constitue une violation de l'article 9 du Code civil, aux termes duquel « *Chacun a droit au respect de sa vie privée* », étant précisé que les tribunaux sanctionnent les atteintes à l'image et à la voix sur le fondement de ce texte qui sert de matrice commune aux droits de la personnalité protégeant l'intégrité morale de la personne.

En deuxième lieu, la réalisation et la diffusion du deepfake litigieux enfreignent l'article 226-8 du Code pénal (issu d'une loi du 21 mai 2024). Ce texte réprime « *le fait de porter à la connaissance du public ou d'un tiers, par quelque voie que ce soit, un contenu visuel ou sonore généré par un traitement algorithmique et représentant l'image ou les paroles d'une personne, sans son consentement, s'il n'apparaît pas à l'évidence qu'il s'agit d'un contenu généré algorithmiquement ou s'il n'en est pas expressément fait mention (...)* ». Or, en l'espèce, la vidéo hypertruquée a été diffusée sur internet sans le consentement de la personne représentée et sans mention de sa manipulation par l'IA.

En troisième lieu, le comportement de l'auteur et diffuseur du deepfake litigieux tombe sous le coup de l'article 29 de la loi du 29 juillet 1881 sanctionnant la diffamation. Selon ce texte, constitue une diffamation toute allégation ou imputation intentionnelle d'un fait qui porte atteinte à l'honneur ou à la considération de la personne, communiquée au public. Les éléments constitutifs de la diffamation sont, en l'occurrence, réunis : d'une part, la vidéo hypertruquée a été réalisée et diffusée avec l'intention de nuire ; d'autre part, elle porte atteinte à l'honneur ou à la considération de la personne représentée puisqu'elle la représente se livrant à un comportement scandaleux ; enfin, la vidéo a fait l'objet d'une communication au public, celle-ci pouvant se faire par voie électronique, donc par internet (loi n° 82-652 du 29 juillet 1982, article 93-3). Dans ces conditions, l'auteur du deepfake s'est rendu coupable de diffamation et peut être condamné sur le fondement de l'article 29 de la loi du 29 juillet 1881 (Cass. crim., 17 mai 2022, n° 21-82949). Il est à noter que si l'auteur du deepfake l'a mis en ligne sur un site internet dont il n'est pas le directeur de publication, la victime pourrait, si elle le souhaite, choisir de demander réparation à ce dernier.

C'est ce fondement qui apparaît le plus en adéquation avec la nature du préjudice allégué (préjudice moral résultant d'une atteinte à la réputation).

En l'espèce, la victime pourra donc, de manière certaine, obtenir réparation de son préjudice moral.

2) Responsabilité de l'hébergeur du deepfake

La responsabilité des hébergeurs (personnes physiques ou morales qui stockent des

contenus aux fins de communication au public en ligne) est difficile à mettre en œuvre. Aux termes de l'article 6 du règlement européen n° 2022/2065 du 19 octobre 2022 sur les services numériques, les hébergeurs ne sont responsables d'un contenu stocké que s'ils ont eu connaissance de son caractère illicite et n'ont pas agi promptement pour retirer le contenu illicite ou rendre impossible l'accès à celui-ci.

Si la victime du deepfake choisissait d'agir en responsabilité contre l'hébergeur, il lui faudrait donc établir d'une part, que celui-ci a eu, d'une manière ou d'une autre, connaissance de la vidéo illicite ; d'autre part, qu'il n'a pas réagi avec une promptitude suffisante pour la retirer ou en bloquer l'accès.

Scenario E - Collision de véhicule autonome

Une voiture autonome interprète mal les données des capteurs dans des conditions météorologiques inhabituelles, la faisant dévier dans la circulation en sens inverse et provoquant un accident impliquant plusieurs véhicules avec des blessés.

Il existe, en droit français, un régime spécial d'indemnisation des victimes d'accidents de la circulation impliquant un ou plusieurs véhicules terrestres à moteur. Ce régime est régi par la loi n° 85-677 du 5 juillet 1985 (dite loi *Badinter*).

Les conditions d'application de ce régime spécial sont au nombre de trois :

- un accident de la circulation dans lequel sont impliqués un ou plusieurs véhicules terrestres à moteur (l'implication, qui s'entend d'un rôle quelconque du véhicule dans l'accident, ne requiert pas une anomalie du comportement du véhicule) ;
- un lien de causalité entre l'accident de la circulation et le dommage subi par la victime (ce qui est requis est un lien de causalité entre l'accident et le dommage, non entre le véhicule du défendeur et le dommage, de sorte que la responsabilité du défendeur peut être engagée quand bien même son véhicule n'aurait joué aucun rôle causal dans la réalisation du dommage subi par la victime) ;
- un responsable identifié en la personne du conducteur ou gardien d'un véhicule impliqué dans l'accident (la loi *Badinter* identifie deux responsables possibles, le conducteur ou le gardien du véhicule impliqué, mais, en pratique, ces deux qualités se trouvent le plus souvent réunies sur la tête d'une seule et même personne).

Lorsque les conditions évoquées ci-dessus sont réunies, la responsabilité qui incombe au conducteur (ou gardien) de chaque véhicule impliqué dans l'accident est extrêmement rigoureuse. C'est pourquoi la souscription d'une assurance de responsabilité automobile est obligatoire.

Contrairement au droit commun de la responsabilité, la force majeure n'a, en effet, aucun effet exonératoire. La seule cause d'exonération que le conducteur ou gardien du véhicule impliqué peut invoquer est la faute de la victime, selon des modalités qui varient en considération de la nature du dommage subi par la victime – corporel ou non – et de la qualité de la victime – conductrice ou non-conductrice.

La question qui se pose est celle de savoir si le régime d'indemnisation des victimes d'accidents de la circulation prévu par la loi du 5 juillet 1985, qui n'a été conçue qu'en considération des véhicules à conduite manuelle, peut s'appliquer à un accident impliquant une voiture autonome.

A vrai dire, les deux premières conditions d'application de la loi *Badinter* ne posent pas problème : il n'est pas douteux qu'une voiture autonome peut être qualifiée de véhicule terrestre à moteur et qu'elle peut être impliquée dans un accident de la circulation, à l'origine du dommage subi par la victime.

La troisième condition, qui a trait à la désignation du responsable, pourrait, en revanche, constituer un point de blocage. Un syllogisme simple vient, en effet, immédiatement à l'esprit : une voiture autonome, par cela même qu'elle est équipée d'un système de délégation de conduite à l'IA, n'a pas de conducteur ni, par voie de conséquence de gardien puisqu'en pratique, le plus souvent, le conducteur d'un véhicule en est en même temps gardien (la garde se définissant comme l'usage, la direction et le contrôle du véhicule) ; or, qui dit absence de conducteur ou gardien dit absence de responsable, au titre de la loi du 5 juillet 1985 ; donc, celle-ci n'est pas applicable aux accidents de la circulation impliquant une voiture autonome.

En réalité, ce syllogisme est vicié car sa majeure est inexacte. Admettre que les voitures autonomes ont un conducteur, au sens de la loi du 5 juillet 1985, s'impose, au contraire, comme une nécessité, pour une double raison.

La première est textuelle. Il ressort de plusieurs textes qu'une voiture autonome doit, en tout état de cause, avoir un conducteur.

D'abord, la Convention de Vienne du 8 novembre 1968 sur la circulation routière, après avoir énoncé que « *tout véhicule en mouvement (...) doit avoir un conducteur* » (article 8 §1), précise que « *[l'] exigence selon laquelle tout véhicule ou tout ensemble de véhicules en mouvement doit avoir un conducteur est réputée satisfaite lorsque le véhicule utilise un système de conduite automatisé qui est conforme* » à la législation nationale et internationale (article 34^{bis}, ajouté en 2022). Ainsi, selon la Convention de Vienne, toute voiture autonome a un conducteur ; or, ce texte a une valeur normative supranationale ; donc, le droit français doit considérer que tout véhicule autonome a un conducteur.

Ensuite, le Code de la route vise, dans plusieurs articles, le conducteur du véhicule autonome (articles L. 123-1 ; R. 319-1 ; R. 412-17-1).

Enfin, les textes encadrant l'expérimentation des véhicules terrestres autonomes imposent la présence d'un conducteur en capacité de reprendre le contrôle du véhicule à tout moment et notamment en cas d'urgence, qu'il soit ou non à bord du véhicule (D. n° 2018-211 du 28 mars 2018, article 12.II) et précisent que le système de délégation (partielle ou totale) de conduite doit pouvoir à tout moment être neutralisé ou désactivé par un conducteur, situé dans ou hors du véhicule (L. n° 2019-486 du 22 mai 2019, article 125).

La seconde raison d'admettre qu'une voiture autonome a, en tout état de cause, un conducteur, au sens de la loi du 5 juillet 1985, relève de la politique juridique : refuser d'appliquer la loi *Badinter* aux accidents de la circulation impliquant ce type de

véhicules, au motif qu'ils n'ont pas de conducteur, aboutirait à créer une inégalité entre les victimes très difficile à justifier : la victime d'un véhicule classique bénéficie de ce texte protecteur, tandis que celle d'une voiture autonome, qui ne pourrait pas s'en prévaloir, serait, par là-même moins bien traitée.

Reste tout de même à vérifier qu'en pratique, toute voiture autonome possède bien un conducteur identifiable susceptible d'endosser la responsabilité d'un éventuel accident.

Le législateur n'a pas défini la notion de conducteur. Selon la jurisprudence, est conducteur celui qui a « *conservé une certaine maîtrise* » dans la conduite du véhicule (voir notamment, Cass. 2^e civ., 14 janvier 1987, n° 85-14655). C'est ce critère, au demeurant assez flou, qu'il faut mettre à l'épreuve des voitures autonomes.

Pour ce faire, il convient de partir des classifications des véhicules autonomes selon leur niveau d'autonomie. Il existe des classifications, dépourvues de valeur normative, élaborées par les constructeurs automobiles. Celle émanant de l'organisation internationale des constructeurs automobiles (OICA), populaire en Europe, distingue 6 niveaux (de 0 à 5). Le Code de la route français retient, quant à lui, une classification juridique qui ne comporte que trois niveaux : véhicules partiellement automatisés, véhicules hautement automatisés et véhicules totalement automatisés (article R. 311-1). En combinant ces sources, on peut, au prisme de la question du conducteur, distinguer quatre catégories de voitures autonomes : les voitures faiblement autonomes, partiellement autonomes, hautement autonomes et totalement autonomes.

Ne disposant pas de précisions quant au degré d'autonomie de la voiture dans le cas proposé, il convient d'envisager les quatre hypothèses, étant toutefois précisé que les deux dernières (voitures hautement ou totalement autonomes) sont des plus improbables dans la mesure où les voitures autonomes qui circulent en France, à l'heure actuelle, ne dépasse pas le niveau de l'autonomie partielle.

1^{re} hypothèse : la voiture impliquée dans l'accident est faiblement autonome.

La catégorie des voitures faiblement autonomes, qui ne figure pas dans la classification tripartite du Code de la route, correspond aux niveaux 0 à 2 de la classification élaborée par l'OICA. Au niveau 2, l'assistance à la conduite porte à la fois sur les fonctions de guidage longitudinal (régulateur de vitesse adaptatif) et latéral (maintien de la trajectoire dans la voie de circulation) mais uniquement dans certaines situations bien spécifiques (circulation sur autoroute, manœuvre de stationnement, par exemple). En tout état de cause, une personne doit en permanence contrôler les fonctionnalités de conduite automatisées et surveiller l'environnement afin d'être en mesure, à tout moment, de reprendre sans délai le contrôle du véhicule.

On admettra aisément qu'une telle voiture, par cela même qu'elle requiert un contrôle humain permanent, a bien un conducteur, au sens jurisprudentiel du terme (conservation d'une certaine maîtrise).

2^e hypothèse : la voiture impliquée dans l'accident est partiellement autonome.

La voiture est équipée d'un système de conduite automatisé exécutant toutes les actions de conduite mais seulement dans certains conditions définies

(météorologiques, géographiques ou de circulation). Le véhicule requiert néanmoins un superviseur humain, installé au poste de conduite qui, s'il n'a pas à surveiller en permanence le système de délégation de conduite (là est la différence avec la voiture faiblement autonome de niveau 2), doit être, à tout instant prêt à reprendre les commandes, avec une réactivité suffisante, soit lorsque les conditions de la délégation de conduite ne sont plus réunies (ce qui sera assez fréquent car la capacité fonctionnelle du système a un domaine limité) soit en cas de demande d'intervention formulée par le système.

Cette deuxième catégorie de voitures autonomes, que le Code de la route français désigne sous l'appellation de « *véhicules partiellement automatisés* », correspond au niveau 3 de la classification émanant de l'OICA.

La nécessité d'une supervision humaine du système de délégation de conduite, fût-elle non permanente, et d'une possible reprise en main des commandes à tout moment permet encore de considérer que le véhicule a un conducteur, au sens de la loi du 5 juillet 1985, pour peu qu'on accepte d'interpréter souplesment le critère jurisprudentiel de conservation d'une certaine maîtrise.

3^e hypothèse : la voiture impliquée dans l'accident est hautement autonome.

Le système de délégation de conduite équipant la voiture exécute toutes les actions de conduite dans certaines conditions déterminées et est apte, de surcroît, à répondre à tout aléa de circulation ou défaillance, sans demande de reprise en main (là est la différence avec la catégorie des voitures partiellement autonomes). Il faut tout de même qu'une personne à bord de la voiture se tienne constamment en état et en position de reprendre en main le contrôle du véhicule (Code de la route, article R. 412-17-1.1) soit lorsque les conditions de délégation de conduite ne sont plus réunies soit dans certains cas, énumérés par l'article R. 412-17-1. Il du Code de la route, où le besoin s'en ferait sentir (nécessité d'obtempérer à une sommation de s'arrêter émanant des forces de l'ordre ou de respecter les indications données par les agents réglant la circulation ou encore de laisser le passage à un véhicule d'intérêt général prioritaire).

Cette troisième catégorie de voitures autonomes, que le Code de la route français appelle « *véhicules hautement automatisés* », correspond au niveau 4 de la classification émanant de l'OICA.

Si la supervision humaine est moindre que dans la catégorie précédente, elle ne disparaît cependant pas totalement : la nécessité qu'une personne à bord du véhicule se tienne en état de reprendre en main le contrôle du véhicule peut, derechef, permettre de considérer que le véhicule a un conducteur, au prix, il est vrai, d'un assouplissement supplémentaire du critère de conservation d'une certaine maîtrise dans la conduite.

4^e hypothèse : la voiture impliquée dans l'accident est totalement autonome.

Le système de délégation de conduite équipant la voiture exécute, en tout temps et en tout lieu, toutes les actions de conduite et peut répondre à tout aléa de circulation ou défaillance. La délégation de conduite est ainsi totale et permanente, de sorte que la voiture est dépourvue de volant et de pédales.

Cette quatrième catégorie de voitures autonomes, que le Code de la route français nomme « *véhicules totalement automatisés* », correspond au niveau 5 de la

classification émanant de l'OICA.

Force est alors d'admettre qu'il n'y a pas de conducteur à bord de la voiture. Mais on peut le trouver à l'extérieur. Selon l'article R. 311-1, 8.3, du Code de la route, les véhicules totalement automatisés ne peuvent circuler qu'à la condition nécessaire d'être intégrés à un « *système technique de transport routier automatisé* », ce qui désigne un ensemble de véhicules automatisés ayant vocation à circuler sur des parcours ou des zones de circulation prédéfinies aux fins de fournir un service de transport collectif de personnes (Code des transports, article R. 3151-1). Or, le Code des transports prévoit qu'en pareil cas, une personne, qui doit être titulaire du permis de conduire (article R. 3151-3), dénommée « *intervenant à distance* » puisqu'elle est située à l'extérieur des véhicules, est habilitée à intervenir pour activer ou désactiver le système de navigation, donner instruction à celui-ci d'effectuer, de modifier ou d'interrompre une manœuvre, de modifier la planification de l'itinéraire ou des points d'arrêts pour les usagers (article R. 3151-1). Il ne paraît pas impossible, en franchissant un ultime palier dans l'assouplissement de la notion de conducteur, de considérer que l'« *intervenant à distance* », au regard de son rôle de supervision extérieure, soit qualifié de conducteur. En cas d'accident impliquant un véhicule totalement autonome, l'intervenant à distance engagerait sa responsabilité, sur le fondement de la loi du 5 juillet 1985, en étant couvert par l'assurance automobile obligatoire souscrite par l'exploitant du système technique de transport routier automatisé.

Mais peut-être estimera-t-on qu'appliquer aux voitures hautement ou totalement autonomes, le critère de conservation d'une certaine maîtrise dans la conduite, dont la jurisprudence se sert pour identifier le conducteur au sens de la loi du 5 juillet 1985, relève de l'artifice, au motif que le critère se trouve, en pareille occurrence, vidé de sa substance concrète. Qu'à cela ne tienne, des propositions alternatives ont été formulées en doctrine pour désigner autrement le conducteur du véhicule hautement ou totalement autonome.

Certains auteurs (cf. L. Andreu (dir.), *Des voitures autonomes. Une offre de loi*, Dalloz, 2018, n° 01-72, n° 02-151 et s.) proposent de considérer que le conducteur est tout simplement celui qui active le système de délégation de conduite du véhicule hautement ou totalement autonome. Il s'agirait en quelque sorte d'un conducteur par détermination de la loi.

D'autres auteurs (cf. A. Dumery, *Pour l'application de la loi Badinter aux véhicules autonomes*, RLDC 2019, n° 174, p. 45) suggèrent d'admettre que le conducteur (et gardien) de la voiture hautement ou totalement autonome est le constructeur de celle-ci ou le concepteur du système de délégation de conduite, du moins tant que ce système est activé car, en cas de reprise en main manuelle, il serait loisible de considérer que c'est la personne physique ayant repris les commandes qui devient conductrice.

En définitive, le régime d'indemnisation des victimes d'accidents de la circulation, fondé sur la loi du 5 juillet 1985, paraît pouvoir s'appliquer aux voitures autonomes. C'est certain en ce qui concerne les véhicules faiblement ou partiellement autonomes, qui sont actuellement les seuls qu'on puisse trouver sur le marché, et ce n'est pas inconcevable pour les véhicules hautement ou totalement autonomes qui

pourront circuler à l'avenir. Dans le scénario qui nous est soumis, les victimes devraient donc pouvoir agir contre le conducteur de la voiture (faiblement ou partiellement) autonome impliquée pour être indemnisées par l'assureur automobile.

Cela étant, l'application de la loi du 5 juillet 1985 aux accidents de la circulation impliquant des voitures autonomes ne manquera pas de soulever une difficulté qu'il sera difficile d'esquiver. Celle-ci tient à la différence de traitement très marquée qu'instaure la loi *Badinter* entre les victimes non-conductrices (qui sont, dans presque tous les cas, intégralement indemnisées de leur dommage corporel) et les victimes conductrices (dont le droit à indemnisation est exclu ou réduit à la moindre faute). Semblable discrimination entre les victimes d'accidents de la circulation deviendra, à hauteur de principe, de moins en moins justifiable à mesure que le niveau d'autonomie des véhicules augmentera. Comment une différence substantielle de traitement entre les victimes conductrices et les victimes non-conductrices pourrait-elle survivre à la perte progressive de substance de la notion même de conducteur ? Comment admettre, par exemple, que le fait de ne pas avoir bouclé sa ceinture de sécurité n'aura aucune incidence sur l'indemnisation du passager victime mais réduira celle du conducteur du véhicule autonome dont le rôle va pourtant aller en déclinant ?

L'expansion des véhicules autonomes pourrait ainsi être, à terme, le facteur déclencheur d'une réforme du régime d'indemnisation des victimes d'accidents de la circulation.