INTELLIGENCE ARTIFICIELLE, RISQUES ET ASSURANCE

Quelques réflexions à l'intention des Dirigeants et Administrateurs

Avril 2025

DiotSiaci

L'hypermédiatisation de l'intelligence artificielle (IA) est finalement toute récente. Elle remonte au lancement de ChatGPT fin 2022 et a accompagné l'introduction de l'IA dite générative. Jusqu'alors, le terme d'IA s'inscrivait sur le continuum des progrès accomplis en informatique : toujours plus de puissance de traitement, de rapidité d'exécution et, avec le développement exponentiel d'internet depuis 25 ans, de données disponibles à traiter.

L'IA générative a par contre été tout de suite perçue comme transformative à trois titres au moins :

- Formé à comprendre le langage naturel, ce nouvel outil peut être utilisé par tout un chacun, sans connaissance du langage informatique et de la programmation, élargissant ainsi considérablement le champ des utilisateurs ;
- Restituant les résultats de ces traitements dans une forme également naturelle mais aussi sous forme de codes, de textes, d'images, de vidéos, de tableaux ou même de musique, elle ouvre des perspectives quasi infinies d'applications potentielles ;
- Enfin, après l'intelligence assistée, puis l'intelligence augmentée, l'IA générative dispose d'une forme d'autonomie en s'enrichissant en permanence, après son déploiement initial, de ses propres résultats ou de toute donnée nouvelle à laquelle les modèles peuvent accéder.

On pressent que l'IA générative aura un impact majeur sur le fonctionnement des entreprises dans toutes leurs dimensions : commerciale, administrative, production, recherche. Certains dirigeants réfléchissent à réorganiser leur entreprise de fond en comble autour des fonctionnalités permises par l'IA[1]. D'autres[2] donnent des instructions pour geler tout recrutement pour exécuter des tâches dont l'IA pourrait se charger.

Comme toute évolution d'importance, l'IA présente autant d'opportunités pour le développement de l'entreprise et l'optimisation des ressources mobilisées qu'elle ne génère de nouveaux risques ou n'aggrave des risques existants. Face à ces bouleversements à peu près certains, mais dont les implications et la portée restent encore à définir, le conseil d'administration, dont la mission est de définir la stratégie de l'entreprise et de déterminer son mode d'organisation, doit accompagner l'équipe dirigeante dans cette réflexion existentielle.

Il est compréhensible que les dirigeants et administrateurs s'interrogent sur les risques personnels qu'ils peuvent encourir du fait des changements qu'ils peuvent initier, ou au contraire différer, sur les mesures de prévention qu'ils peuvent adopter pour eux-mêmes et pour l'entreprise, et sur l'efficacité de la protection apportée par l'assurance en couverture de ces nouveaux risques.

Nous nous efforçons d'apporter quelques éléments de réponse inspirés par notre expérience de gestionnaire de risques au service des entreprises et des institutions financières en distinguant :

- Les risques encourus par les dirigeants et mandataires sociaux (les dirigeants sociaux) ;
- Les risques encourus par l'entreprise.

1. *The Economist weekly edition. 22/4/2023 - The generation game*
2. *Bloomberg.com 01/05/2023 – IBM to pause hiring for back-office jobs that AI could kill*

Les principales sources de mise en cause

En matière civile, toute personne, physique ou morale, qui estime avoir subi un préjudice du fait de la faute d'un dirigeant peut en chercher réparation en engageant sa responsabilité personnelle. Ainsi, les actionnaires ou associés, les créanciers, les clients, les fournisseurs, les employés, et plus largement toutes les parties prenantes à la vie de l'entreprise, peuvent agir contre les dirigeants sociaux. Mais il faut distinguer l'engagement d'une procédure de ses chances de succès. Hormis les associés et les créanciers, sur lesquelles nous revenons ci-après, les autres catégories de demandeurs doivent établir qu'une faute détachable de ses fonctions sociales, c'est-à-dire une faute grave et intentionnelle, a été commise par le dirigeant mis en cause pour qu'une action dirigée personnellement contre lui soit recevable, à défaut de quoi l'action est en quelque sorte redirigée contre l'entreprise elle-même dont le dirigeant n'agit que comme vitrine légale.

En revanche, les actionnaires ou associés, en cas de perte de valeur de leur investissement, ou les créanciers, en cas d'alourdissement du passif, disposent des outils juridiques pour engager une action contre les dirigeants sociaux personnellement s'ils estiment qu'ils ont commis une faute.

Aux actions civiles engagées par les actionnaires, dans le cas d'une société in-bonis, et par les créanciers, ou en leur nom, dans le cas d'une liquidation judiciaire, il faut ajouter les procédures initiées par une autorité administrative telle que l'AMF, la CNIL, l'Autorité de la Concurrence ou, prochainement, celles qui pourraient être engagées en cas de non-respect du Règlement Européen sur l'Intelligence Artificielle et qui peuvent se traduire par des sanctions ou des amendes administratives.

Fondements des mises en cause

Extrapolant certains griefs retenus dans des instances mettant en jeu la responsabilité personnelle des dirigeants, nous imaginons ci-après quelques exemples de fautes qui pourrait trouver leur source dans des sujets liés à l'intelligence artificielle et fonder une demande en réparation :

- Une action récursoire contre les dirigeants suite à une action en responsabilité contre l'entreprise pour un dommage subi par un tiers (divulgation de données confidentielles, violation de la propriété intellectuelle d'autrui) ou d'un dommage subi par l'entreprise elle-même (dissémination de données sensibles appartenant à l'entreprise, fraude au préjudice de l'entreprise facilitée par la technologie de l'intelligence artificielle, sanction administrative lourde infligée à l'entreprise pour non-respect des règlements relatifs à l'IA). Les dirigeants sociaux pourraient se voir reprocher de n'avoir pas mis en place les outils de prévention (gouvernance formalisée, formation du personnel, encadrement des usages, vérification humaine des productions de l'IA) ou de protection (adaptation des contrats d'assurance de l'entreprise à la couverture des risques découlant de l'IA) ;
- Des investissements disproportionnés dans la technologie de l'IA par rapport aux moyens de l'entreprise et aux résultats effectivement obtenus ;
- Une dépendance excessive aux technologies de l'IA si, à la suite d'un incident d'origine interne ou externe à l'entreprise, celle-ci devait faire face à une interruption totale ou partielle de sa production ou des services rendus ;

- Une communication inexacte ou trompeuse sur les bénéfices attendus du déploiement de solutions IA (augmentation du chiffre d'affaires, amélioration de la rentabilité par une optimisation des coûts ; efficacité augmentée des produits vendus ou des services rendus). Les anglo-saxons regroupent ces griefs sous le label « d'AI washing » par analogie au « greenwashing ». Le Président de la Securities and Exchange Commission a annoncé « une guerre sans merci » contre les entreprises et leurs dirigeants qui abuseraient les investisseurs par une communication insincère[3]. Mi-mars 2025, quatre instances, faisant suite à quinze en 2024, avaient déjà été engagées devant des tribunaux américains sous forme de « securities class actions » fondées sur des informations trompeuses de l'entreprise ou de sa direction au sujet des bénéfices que l'entreprise pourrait retirer du développement de l'IA ou de son utilisation dans ses procédures. Ces actions étaient consécutives à une baisse du cours du titre de l'entreprise, risque d'autant plus élevé que les titres de ces entreprises peuvent être considérés comme chers, les investisseurs attribuant une prime aux entreprises susceptibles de profiter du marché de l'IA.

Conséquences financières

Les conséquences financières de la mise en jeu de responsabilité personnelle des dirigeants sont de trois natures :

- Les **frais de défense** : Dans tous les cas de mise en cause, civile, pénale ou administrative, les dirigeants doivent se défendre dans des procédures complexes, d'autant plus coûteuses qu'il est souvent nécessaire, outre les avocats, de faire intervenir des experts « métiers ».
- Le cas échéant, les **dommages-intérêts** fixés au terme d'un jugement ou négocié dans le cadre d'un accord transactionnel. L'extension aux dirigeants d'une demande en réparation répond parfois à une tactique d'intimidation pour forcer un accord transactionnel avec l'entreprise ou ses assureurs ;
- Les **sanctions financières** prononcées par une autorité administrative ou une juridiction pénale.

Prévention et protection

Prévention

La seule constatation d'un préjudice (perte de valeur des titres, alourdissement du passif) ne suffit pas à obtenir réparation. Les juges intègrent le risque économique dans leur appréciation d'une situation et la possibilité qu'une décision ne produise les effets positifs escomptés. Le demandeur doit donc établir qu'une faute a été commise et démontrer un lien de causalité entre la faute et le préjudice. Pour se défendre, les dirigeants mis en cause devront de leur côté faire la démonstration qu'ils ont eu un comportement responsable.

Compte tenu du potentiel transformatif de l'IA, que nous avons rappelé en introduction et sachant que des usages clandestins peuvent en être fait ici où là dans l'entreprise, il est souhaitable :

- D'inscrire toute initiative dans une réflexion globale, pluridisciplinaire et documentée, menée par une équipe ad-hoc associant à ses travaux un ou plusieurs administrateurs rendant eux-mêmes compte aux autres membres du conseil ;

3. U.S.Sec - Speeches-Statements 13/2/2024 / Gensler and AI

- De déployer la technique par petites touches sous le contrôle de collaborateurs expérimentés, en cohérence avec la démarche globale et en ajustant celle-ci avec les résultats et observations tirés des initiatives décentralisées ;
- D'encadrer les usages (applications autorisées, données communicables) en commençant par alerter les équipes des risques que l'utilisation non autorisée peut faire courir à l'entreprise (perte ou dissémination de données sensibles et précieuses, recours de tiers) ;
- Et plus généralement, compte tenu de la nouveauté de la plupart des sujets ayant trait à l'IA, de conserver une attitude prudente et raisonnée dans le déploiement de nouvelles applications.

Protection

En matière de responsabilité des dirigeants, la forme précède et souvent préjuge du fond. Si les dirigeants peuvent démontrer (d'où l'importance de la communication écrite) qu'ils ont fait des efforts raisonnables pour limiter la survenance du risque, la démonstration de la faute, quand bien même un préjudice a été subi, devient plus ardue pour le tiers qui s'estime lésé.

Toutefois, en dépit de pratiques vertueuses évoquées ci-dessus, la citation d'un dirigeant dans une procédure reste possible. Il doit alors se défendre. C'est le moment où le contrat d'assurance dit « RCMS » témoigne de son utilité. Des précautions doivent être prises concernant notamment :

- Le champ des **exclusions**. Les contrats RCMS comportent peu d'exclusions, l'exception principale portant sur les sanctions pénales, inassurables. On sera attentif à deux exclusions qui ne sont pas systématiques mais que les assureurs ont parfois introduites : (i) l'exclusion des réclamations consécutives à l'engagement d'une procédure collective. Cette restriction dénature l'objet de la garantie et doit être proscrite ; (ii) l'exclusion des réclamations fondées sur des pertes « cyber » dont la portée pourrait s'étendre aux réclamations fondées sur un préjudice lié à l'IA ;
- L'intégration de **nouvelles filiales**. Les garanties ne s'appliquent qu'au titre de fautes commises postérieurement à l'acquisition. Pour ne pas risquer d'être tenus responsables de fautes antérieures, les dirigeants doivent faire preuve d'une vigilance particulière lors des « due diligences » pré-acquisition et dans la rédaction de la documentation contractuelle (contrat de cession et garantie de passif)
- Le **montant de garantie**. Ce montant constitue l'engagement maximum de l'assureur. En cas de pluralité de dirigeants mis en cause, ce montant est partagé, la pratique étant que chaque dirigeant retienne son propre conseil. Il faut s'interroger régulièrement sur l'adéquation entre le montant de garantie et les coûts probables de défense sans parler du montant en principal d'éventuelles condamnations.

Les risques encourus par l'entreprise

Une bonne maîtrise des risques encourus par l'entreprise contribue à la stabilité de ses résultats et à sa pérennité en même temps qu'elle constitue, si elle est effective et documentée, le premier garde-fou contre les risques de responsabilité auxquels les dirigeants sociaux sont personnellement exposés. Elle implique un travail minutieux de cartographie des risques. A quels risques l'entreprise est-elle exposée du fait de l'utilisation active ou passive de l'IA ? Ces risques sont-ils nouveaux ou existants, éventuellement aggravés ? Comment les inventorier, en mesurer l'impact et les prévenir ? Les contrats d'assurance existants suffisent-ils à en couvrir les conséquences financières ? Les questions sont nombreuses. Les réponses plus encore suivant que les modèles et applications d'IA déployés dans l'entreprise sont développés en interne (ils le sont rarement en totalité) ou achetés et « customisés », ou encore selon l'origine des données dont ils se nourrissent.

A la lumière des litiges ou incidents qui commencent à essaimer, nous nous contenterons de souligner certains risques que la nouvelle génération de l'intelligence artificielle, l'IA générative, a pour effet d'exacerber avant de développer quelques recommandations en matière d'assurance.

Risque accru de fraude en raison de la véracité augmentée des productions de l'IA

Les outils d'intelligence artificielle permettent de produire des contrefaçons de plus en plus réalistes, augmentant le taux de réussite des manipulations du type des « fraudes au président ». Une société multinationale britannique établie, la société d'ingénierie Arup[4], en a été victime, un employé ayant procédé au transfert de $ 25 millions en réponse à l'instruction en vidéo de son faux Directeur Financier.

On pense d'abord à l'utilisation de ces « deepfakes » pour détourner des actifs financiers mais les entreprises doivent être vigilantes à la protection de tous les actifs numériques constitutifs de leur valeur (clients, prix, savoir-faire, travaux de recherche, …). Les mises en garde sur les risques de détournement d'actifs matériels et immatériels ne doivent donc pas se limiter aux collaborateurs occupant des postes financiers. Tous les salariés ayant accès aux systèmes et aux données de l'entreprise doivent faire preuve de vigilance et être formés.

Parallèlement à la protection de leurs propres valeurs, elles doivent veiller à l'utilisation de leur identité, ou celles de leurs collaborateurs, vis-à-vis des tiers. La Banque BNP Paribas[5] a été tenue d'indemniser un client victime d'une fraude au « faux conseiller bancaire » fondée sur l'usurpation de numéros de téléphone de la banque. On peut projeter de nombreux scenarii frauduleux ou malveillants utilisant des faux de plus en plus crédibles façonnés par l'IA, et s'inquiéter à l'idée que ces manipulations puissent viser des infrastructures critiques.

4. *Financial Times 16/5/2024 / ARUP and Hong Kong deepfake*
5. *Les Echos 24/10/2024 BNP Paribas et faux conseiller bancaire.*

Attention aux agents dormants

Avec l'IA générative, l'IA cesse d'être un simple outil pour devenir un agent capable d'apprendre et de prendre des décisions de manière indépendante sans intervention humaine. Une attention particulière doit être portée au suivi de ces systèmes pour s'assurer qu'ils remplissent la fonction qui leur est assignée.

L'Acte sur l'Intelligence Artificielle adopté au plan Européen prévoit que les applications à haut risque, déployées dans des domaines aussi critiques que les infrastructures vitales, l'éducation, l'emploi, et la justice, devront répondre à des normes particulièrement strictes concernant la gestion des risques, la qualité des données, la documentation technique, et la surveillance humaine.

Plus pernicieuses sont les applications qui, au niveau d'une entreprise individuellement, ont vocation à être des aides à la décision et qui, utilisées par un grand nombre d'acteurs d'un même secteur, génèrent des comportements moutonniers dont l'effet pratique peut être de réduire la concurrence.

Une trentaine d'actions civiles, relayées par une action de la section Antitrust du Department of Justice[6], ont été engagées aux Etats-Unis sur le fondement d'entente illicite en matière de fixation des loyers. Ces litiges mettent en cause des administrateurs de biens et des éditeurs de logiciel à qui il est reproché une collusion de fait aboutissant à un alignement des loyers par le haut. Les logiciels, formatés pour atteindre un même objectif, la maximisation des loyers, fonctionnant avec les mêmes algorithmes, et se nourrissant des mêmes données, aboutissaient aux mêmes préconisations.

Dans d'autres cas, assez nombreux[7], les applications produisent des résultats inexplicables, désignés sous le terme d'hallucinations et qui, sans contrôle humain, peuvent se traduire par des recommandations dommageables, voire dangereuses, susceptibles de porter atteinte à la confiance des consommateurs et à la réputation de l'entreprise. Les entreprises, et notamment celles fournissant des prestations intellectuelles, seraient donc avisées de prévoir la validation des documents produits par l'IA par un lecteur expérimenté.

La protection des données, la mère des batailles (juridiques)

Les données sont au cœur de la chaîne de valeur de l'IA. Pour preuve, la plupart des litiges touchant à l'IA est liée à la collecte, au traitement et à l'utilisation des données. Les actions en justice sont nombreuses[8]. Quasiment tous les développeurs de modèles d'IA (les modèles de fondation) sont mis en cause, les fondements invoqués étant, selon les cas, l'atteinte à la vie privée, la divulgation d'informations confidentielles, le non-respect de la propriété intellectuelle et des droits d'auteur. Dans l'affaire Cambridge Analytica (dont le slogan était « Data drives all we do »), Facebook s'est vu infliger une amende de $5 milliards imposée par la Federal Trade Commission confirmée par une décision de justice (Avril 2020) et a conclu un accord transactionnel de $ 725 millions pour clore l'action collective engagée au nom des utilisateurs de Facebook pour avoir permis à des tiers d'avoir accès à leurs données privées[9].

6. *Justice Department 23/8/2024 RealPage algorythmic pricing schemes*
7. *New York Times 16/11/2023 Chatbot and hallucination rates*
8. *ChatGPT eating the world 27/8/2024 Master list of lawsuits*
9. *Les Echos 23/12/2022 Cambridge Analytica – Facebook*

Prévention

Un grand nombre d'instances sont aujourd'hui pendantes dont l'issue, du fait de leurs complexité et nouveauté, reste incertaine. Dans cet environnement juridique évolutif, les entreprises doivent au moins veiller à ce qui est sous leur contrôle, notamment :

- A l'usage que leurs équipes font des outils d'IA en rappelant la nécessité de protéger les données propres à leur entreprise ou celles qui lui sont confiées. Un article de Business Insider du 11 Juillet 2023[10] citait plusieurs grands groupes (parmi lesquels Amazon, Apple, Deutsche Bank, Samsung, Verizon ou Wells Fargo) qui avaient décidé de prohiber l'utilisation de chatbots du type ChatGPT à partir de leurs systèmes informatiques pointant le risque de fuite de données confidentielles.
- A la conformité des pratiques de l'entreprise avec les divers règlements auxquels elle est soumise (AMF, CNIL, Règlement européen…), les manquements pouvant être lourdement sanctionnés.
- A la diffusion de bonnes pratiques au sein de l'entreprise (« Acceptable Use Policies ») parmi lesquelles la validation par un responsable expérimenté, avant transmission à l'extérieur, de toute production issue d'un modèle d'IA.
- Dans la gestion de leurs responsabilités, à ne pas accepter d'obligations contractuelles qui pourraient être qualifiées d'exorbitantes et en conséquence remettre en cause le bénéfice des contrats d'assurance souscrits par l'entreprise.

L'assurance

Après le premier filet de protection, constitué par les règles de prudence mises en place dans l'entreprise dont la vocation est de prévenir la survenance du risque, le programme d'assurance de l'entreprise est le second niveau de sécurité. Il a pour objet d'indemniser l'entreprise des conséquences financières des risques s'ils se produisent ou, pour certains contrats, s'ils menacent de se produire (« mitigation costs »).

Revue des contrats

Le premier exercice doit être de relire, ou faire relire, chaque contrat d'assurance souscrit par l'entreprise en s'interrogeant sur l'applicabilité du contrat dans le cas où la survenance du risque défini dans l'objet des garanties trouverait son origine dans une application d'IA. Différents scenarii d'incidents doivent être imaginés comme, à titre d'exemples, un incendie déclenché par le dysfonctionnement de thermostats actionnés par un logiciel d'intelligence artificielle, la décision d'un client de l'entreprise prise sur la base d'une préconisation émise par un outil d'IA mis à sa disposition et qui se révèle préjudiciable ou encore une fraude perpétrée contre l'entreprise.

Les principaux contrats à étudier sont ceux souscrits en couverture :

- Des risques de Dommages aux Biens – Pertes d'exploitation, Bris de Machine et Tous Risque Informatique ;
- Des risques de responsabilité couverts par les contrats de RC générale, RC Employeur, RC Produits, RC Professionnelle, RC des Dirigeants et des Mandataires Sociaux ;
- Des risques de cybersécurité.

10. Business Insider 11/07/2023 – Large groups issue restrictions for use of AI software

Ce travail fera ressortir :

- Soit, la **bonne application des contrats**, notamment ceux construits sous forme de « Garanties Tous sauf » ;
- Soit des **zones grises** qui devront être discutées avec les conseils de l'entreprise pour savoir s'il convient d'intervenir auprès des assureurs du programme pour faire préciser les garanties en ayant en tête la maxime du Cardinal de Retz, selon laquelle on sort souvent de l'ambiguïté à son détriment. Peut-être, dans certains cas, sera-t-il plus avisé de se satisfaire d'une couverture implicite (« silent cover ») ;
- Soit des **« trous » de garantie** résultant d'exclusions ou de libellés trop restrictifs. Certaines exclusions sont non négociables, d'autres sont prévues pour inciter à la souscription de garanties spécifiques. Ainsi, les exclusions touchant aux cyber-risques ont été introduites pour encourager les entreprises à souscrire des garanties dédiées à ces risques et permettre aux assureurs de mener un travail de souscription adapté à ces risques émergents.

Les contrats d'assurance Cyber sont le réceptacle naturel pour la couverture des risques découlant de l'IA qui ne sont pas couverts par les garanties standards. Les définitions devront être ajustées par l'introduction de termes propres à l'écosystème de l'IA et faire référence aux textes réglementaires auxquels l'entreprise doit se conformer.

Il faut s'attendre à ce que les assureurs, en contrepartie des garanties étendues à l'IA, renforcent leurs exigences en termes d'informations (formalisation de la gouvernance IA, liste des applications IA, liste des prestataires utilisés par l'entreprise, vérification des assurances RCP souscrites par les prestataires…).

Conclusion

De même qu'on a inventé le terme «d'agriculture raisonnée» pour concilier, dans le domaine de la production agricole, finalité de rentabilité et limitation de l'impact sur l'environnement, de même est-il prudent, à ce stade précoce du développement de l'IA et de la compréhension de ses conséquences, d'expérimenter une **IA raisonnée** c'est-à-dire :

- Encadrée dans une réflexion globale pilotée au niveau de la direction de l'entreprise en y associant le conseil d'administration ;
- Déployée par petites touches, sous le contrôle direct de collaborateurs formés rendant eux-mêmes compte à une équipe de management pluridisciplinaire chargée d'en suivre les effets positifs ou négatifs sur tous les plans, notamment juridique, social mais aussi environnemental, et s'agissant des effets négatifs, de proposer des mesures appropriées pour en supprimer ou en limiter l'impact ;
- Dans le respect des intérêts de toutes les parties prenantes à commencer par les collaborateurs de l'entreprise.

Eric REMUS | Avril 2025

Eric REMUS a dirigé le département des Lignes Financières du Groupe SIACI SAINT HONORÉ. Il avait auparavant fondé et dirigé Assurance & Capital Partners, société de courtage spécialiste des risques financiers.

DIOT - Groupe DIOT-SIACI - Société de Courtage d'Assurance et de Réassurance.

Siège social : Siège social : Siège social : Season - 39, rue Mstislav Rostropovitch - 75815 PARIS - FRANCE - Tél. : +33 (0)1 44 79 62 00.
SAS - Capital : 1 831 008 € - RCS Paris 582 013 736 - N° TVA : FR 92 582 013 736.
N° ORIAS : 07 009 129 (www.orias.fr) - Sous le contrôle de l'ACPR - 4 place de Budapest - CS 92459 - 75436 PARIS CEDEX 09 - FRANCE.
Réclamations : reclamations@diot.com - www.mediation-assurance.org

SIACI SAINT HONORE - Groupe DIOT-SIACI - Société de Courtage d'Assurance et de Réassurance.

Siège social : Season - 39, rue Mstislav Rostropovitch - 75815 Paris cedex 17 - France - Tél. : +33 (0)1 4420 9999.
SAS - Capital : 179.056.753,60 € - RCS Paris 572 059 939 - APE 6622 Z - N° TVA : FR 54 572 059 939.
N° ORIAS : 07 000 771 (orias.fr) - Sous le contrôle de / Regulated by ACPR - 4 place de Budapest - CS 92459 - 75436 Paris cedex 09 - France.
Réclamations / Complaint : SIACI SAINT HONORE - Service réclamations - 23, allées de l'Europe - 92587 Clichy cedex - France.