

INTELLIGENCE ARTIFICIELLE ET CONSENTEMENT

Di Michel Cannarsa

| 389

SOMMARIO: 1. *Le risque de l'IA «dolosive».* – 1.1. *La capacité augmentée par l'IA de vicier le consentement d'autrui.* – 1.1.1. *Une capacité maîtrisée* - 1.1.2 *Une capacité incontrôlée.* - 1.2 *La capacité autonome de l'IA de vicier le consentement d'autrui.* – 2. *L'espoir d'une IA protectrice.* - 2.1 *La protection augmentée du consentement.* - 2.2 *La protection automatisée du consentement.* – 3. *Conclusion.*

ABSTRACT. *L'article discute la question du consentement contractuel dans l'environnement numérique, et plus spécifiquement dans les situations où sont impliqués les systèmes d'intelligence artificielle. En particulier, l'auteur discute si le consentement dans ce contexte doit être compris de la même manière que le droit contractuel ou si il y a des spécificités dans la relation entre le consentement et les systèmes d'IA.*

The article discusses the issue of contractual consent in the digital environment, and more specifically in situations in which artificial intelligence systems are involved. In this context, the author discusses whether consent in this environment should be understood in the same way as traditional contract law or whether there are some specificities in the relationship between consent and AI systems.



La question de la qualité du consentement dans le contexte d'une diffusion assez rapide et massive des technologies d'intelligence artificielle (IA), est un sujet de préoccupation croissante. En effet, l'IA semble de prime abord autoriser des manipulations, qui sont traditionnellement l'une des cibles principales des remèdes juridiques destinés à protéger le consentement. Dans le domaine contractuel plus particulièrement,¹ le consentement est en effet protégé contre les vices, spontanés ou provoqués, pouvant l'affecter, ce consentement devant être libre et éclairé. L'interaction entre les systèmes d'IA et le domaine contractuel pourrait dès lors voir émerger assez rapidement un nouveau domaine du droit de la formation du contrat et des vices du consentement, à savoir celui de la formation assistée, voire automatisée des contrats, et son corollaire des vices causés par l'IA.

Le souci de ne pas caricaturer les effets de l'IA commande de ne pas négliger son potentiel positif, et plus précisément la capacité de certains outils intelligents de protéger le consentement contractuel des individus. Ainsi, d'aucuns pourraient voir dans l'IA un outil d'*empowerment* des individus, particulièrement de ceux qui en règle générale ne sont pas en position de force dans les négociations contractuelles et dans le processus contractuel en général (personnes vulnérables, consommateurs, petites entreprises, etc.). Des outils performants pourraient par exemple alerter le contractant du risque créé par des dispositions contractuelles trop en contradiction avec ses intérêts. Ceci serait d'autant plus pertinent dans un contexte où les transactions interviennent de plus en plus virtuellement, à distance.² En effet, le monde digital, tout exaltant qu'il puisse être, a tendance à abaisser le niveau habituel de méfiance des individus. Ceci est lié à bien des égards à la dimension intangible de cet environnement, qui déjoue en partie l'attention et la vigilance des individus, qu'ils auraient pourtant exercées davantage dans un environnement physique. A cet égard, l'utilisation de l'IA, et plus particulièrement la capacité de générer des *deep fakes*, peut produire des effets assez dévastateurs dans un environnement virtuel où la réalité est justement de plus en plus virtuelle et peut avoir tendance à se substituer à la réalité physique et tangible.

Lorsque l'on aborde la question du consentement contractuel dans l'environnement numérique et plus précisément dans des situations où des systèmes d'IA sont impliqués, il convient de s'interroger quant à la signification du consentement dans ce contexte plus particulier. Plus précisément, le consentement dans ledit contexte doit-il être appréhendé de la même façon qu'en droit des contrats classique ? existe-t-il des spécificités dans le rapport entre consentement et les systèmes d'IA ? A l'évidence, la réponse doit être positive, ne serait-ce qu'en raison de l'automatisation, au moins

¹ Cette courte étude se focalisera sur la notion de consentement dans le contexte contractuel classique, en tant que condition de validité du contrat et en tant qu'objet de protection au travers des sanctions des vices du consentement.

² Il suffit de penser aux transactions intervenant sur les plateformes de E-commerce.

partielle, qu'engendre le recours aux systèmes d'IA dans le processus contractuel.³

Sur la base des propos qui précèdent et dans une organisation dialectique assez simple, la présente contribution envisagera, d'une part, la capacité des technologies d'IA de vicier le consentement, potentiellement à grande échelle (Partie 1), tout en ménageant, d'autre part, l'alternative plus positive d'une IA susceptible de protéger le contractant et son consentement (Partie 2).

1. Le risque de l'IA «dolosive»

L'une des caractéristiques des technologies d'IA est la capacité d'auto-apprentissage. Cet auto-apprentissage est souvent couplé à l'autonomie. Ainsi, certains systèmes d'IA développent assez rapidement leurs propres trajectoires, sur la base d'instructions initiales potentiellement assez larges, et ceci au gré des données entrantes de diverses natures, que ces systèmes vont collecter de façon assez continue, notamment en étant interconnectés avec un écosystème « intelligent » (il s'agit-là de l'environnement de l'IoT, pour Internet of Things). Ce constat suppose de s'intéresser à la capacité des systèmes d'IA de vicier le consentement d'un contractant de façon autonome (1.2). Toutefois, sans aller jusqu'à cette hypothèse extrême, les systèmes d'IA peuvent plus classiquement accentuer les moyens de certains protagonistes d'altérer le consentement, ou à tout le moins la réflexion, de leurs co-contractants (pensons par exemple aux algorithmes de recommandation...). Dans cette dernière configuration, le système d'IA ne vicie pas en autonomie le consentement d'un co-contractant, mais autorise et accroit cette faculté chez l'individu déployant des systèmes d'IA à l'aune de ses intérêts (1.1).

1.1. La capacité augmentée par l'IA de vicier le consentement d'autrui

Quel que soit le degré d'autonomie des systèmes d'IA, leurs effets sont supposés être attribués à l'individu, la personne physique ou morale qui a recours à ces systèmes.⁴ En effet, dans la mesure où les systèmes d'IA, pour

³ La question du recours aux systèmes d'IA dans le processus contractuel est souvent appréhendée plus largement au travers de sa dimension d'automatisation, cette dernière altérant en effet significativement le paradigme classique du contrat: v. UNCITRAL Model Law on Automated Contracting (2024), disponible à l'adresse suivante : <https://uncitral.un.org/en/mlac>.

⁴ V. *ELI Guiding Principles and Model Rules on Digital Assistants for Consumer Contracts*, disponible à l'adresse suivante : https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Projects/Algorithmic_Contracts/Guiding_Principles_and_Model_Rules_on_Digital_Assistants_for_Consumer_Contracts.pdf, p. 55.

l'heure en tout cas, ne sont pas dotés de personnalité juridique,⁵ il convient d'identifier l'entité dotée de personnalité juridique qui endossera les « actions » de l'IA. Ce principe, qui semble relever de l'évidence, n'est pourtant pas encore fermement établi, ce qui manifeste le caractère encore flottant de l'encadrement juridique de l'IA. Si l'on part du principe d'une attribution des effets générés par les systèmes d'IA à son utilisateur, alors les effets de ces systèmes sur le consentement des contractants peuvent s'envisager comme des multiplicateurs de la capacité traditionnelle de certains agents d'induire en erreur leurs cocontractants (1.1.1). Cette première hypothèse est celle d'une utilisation maîtrisée et dolosive d'un système d'IA, dans l'intention caractérisée de tromper le cocontractant. Pourtant, la complexité et l'opacité fréquentes, voire inhérentes, aux systèmes d'IA rendent parfois ou souvent illusoire la maîtrise des effets, des *outputs*, desdits systèmes. Les effets en question, produits non-intentionnellement, n'en seront pas moins attribués (en tout cas c'est hautement souhaitable) à un agent juridique (personne physique ou personne morale), au même titre qu'un dol classique (1.1.2).

1.1.1. Une capacité maîtrisée

Dans cette première configuration, il est fait l'hypothèse d'une configuration de systèmes d'IA avec l'intention d'altérer le consentement du cocontractant. Il s'agit d'un stade assez rudimentaire de l'utilisation de la technologie pour tromper son interlocuteur. De «simples» systèmes algorithmiques, n'intégrant pas nécessairement des technologies d'IA, peuvent permettre de telles manœuvres. L'IA permettra dans ces cas-là d'augmenter la logique sous-jacente, destinée à altérer le consentement. On peut penser à la génération de faux contenus (sons, images), de faux avis, d'informations ciblées pour un certain type de public, etc. Parfois, ces pratiques pourront d'ailleurs passer les confins du dol pour s'établir dans le territoire de la violence et des pratiques commerciales agressives (cela pourrait être le cas, par exemple, des pratiques sur certaines plateformes de réservation, consistant à imposer un délai court de décision arguant d'un nombre résiduel limité des biens ou des services objet du contrat à conclure).

Dans ce type de configuration, le déclenchement des remèdes classiques des vices du consentement, et particulièrement celui du dol, s'impose sans grande difficulté, puisque ses éléments constitutifs (intention dolosive et effet d'altération du consentement du cocontractant) sont réunis de façon assez similaire aux situations dans lesquelles aucune technologie n'est utilisée pour induire en erreur.

1.1.2. Une capacité incontrôlée

⁵ V. notamment la résolution du Parlement européen du 16 février 2017, 'Civil Law Rules on Robotics', disponible à l'adresse suivante : https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html

La complexité et l'opacité des systèmes d'IA évoquée plus haut et assez unanimement admise,⁶ peut rapidement rendre illusoire le contrôle par l'agent juridique (personne physique ou morale) des *outputs* générés par l'IA. Qu'il y ait eu à l'origine intention trompeuse ou pas, l'évolution potentiellement imprévisible des systèmes auto-apprenants et autonomes rend illusoire la maîtrise des «actions» des systèmes. Pourtant, les actions en question seront bien attribuées aux «déployeurs» des systèmes en question. Ainsi, la vocation traditionnelle de sanction du dol pourrait s'en trouver altérée puisque l'intention de l'auteur pourrait englober des hypothèses où cette intention est bien caractérisée et d'autres où elle résulte du déploiement de systèmes dont on savait qu'ils ne seraient pas, ou pas pleinement, contrôlables et qu'ils étaient susceptibles de tromper des cocontractants. En effet, bien que les effets des systèmes d'IA soient supposés attribués aux personnes les déployant, il s'agit et il s'agira de plus en plus d'une forme de fiction juridique. Un parallèle pourrait d'ailleurs être fait avec la responsabilité du fait d'autrui à son origine où l'idée sous-jacente de faute dans le choix du préposé (n'offrant pas les qualités et compétences qui auraient permis d'éviter un dommage), rend d'autant plus admissible la responsabilité du commettant pour les dommages causés par son préposé. Avec les systèmes d'IA toutefois, la maîtrise de leurs paramètres, d'une part, et de leur comportement, d'autre part, est nettement plus complexe et presque illusoire. Alors, il n'est pas excessif de penser qu'il y a une faute dans le recours à des systèmes dont on ne peut ignorer qu'ils risquent de développer un potentiel dolosif pour répondre à des instructions initiales, comme celle par exemple assez prévisible et basique, à savoir attirer le plus de clients (cocontractants) possibles.

1.2. La capacité autonome de l'IA de vicier le consentement d'autrui

Le *deep learning* est un aspect prépondérant des systèmes d'IA. Cette capacité permet au système d'IA concerné d'ajuster son «comportement», ses *outputs* au gré des *inputs* qu'il recueille de différentes sources externes (des individus, d'autres systèmes d'IA, des objets connectés, etc.). Signalons incidemment, même s'il ne s'agit pas du cœur de cette contribution, que l'individu doit pouvoir consentir à interagir avec l'IA et donc être informé d'une interaction avec un système d'IA.⁷ En effet, les personnes physiques ne savent pas forcément qu'elles interagissent avec des systèmes d'IA, alors que cette information est indispensable pour consentir à cette interaction. Cette préoccupation est d'autant plus cruciale dans un contexte où le réalisme de voix et/ou d'images générées par les systèmes d'IA dépasse la capacité sensorielle humaine à détecter l'origine synthétique de la communication et de l'interaction. Les individus doivent pouvoir disposer d'un droit à une alternative à l'interaction avec un système d'IA.

⁶ V. par ex. dans le règlement (UE 2024/1689) IA du 13 juin 2024, considérant 31.

⁷ V. à ce sujet la question de la notification de l'usage d'un système d'IA, art. 19 des ELI Guiding Principles and Model Rules on Digital Assistants for Consumer Contracts, n. 4.

Bien que les *outputs* d'un système d'IA soient attribués à son déployeur, en tout cas en l'état actuel des réflexions sur ces questions,⁸ il faut considérer que l'autonomie grandissante de certains systèmes d'IA, particulièrement ceux opérant par auto-apprentissage, crée les conditions d'une faculté « autonome » à vicier le consentement du cocontractant du déployeur, notamment afin d'atteindre les objectifs qui ont été fixés au système, comme par exemple celui de maximiser des profits et de limiter les risques, économiques comme juridiques. Ce type d'objectif est assez vraisemblable dans le contexte commercial, domaine où les consommateurs sont pourtant en règle générale dans une position défavorable par rapport au professionnel.

Des instructions précises et l'intégration d'un certain nombre de garde-fous par le déployeur permettront-elles d'éviter des actions? le fossé existant entre le potentiel technologique et la capacité à le canaliser peuvent suggérer un certain scepticisme à cet égard. Le processus contractuel peut ainsi être compromis de différentes façons, que cela soit en excluant des contractants (on peut penser par exemple à des errances du *credit scoring* ou du *social scoring*) ou en les amenant par différentes techniques à prendre une décision qu'ils n'auraient pas prise autrement. Or, plus un système d'IA développe son autonomie, plus le processus contractuel est automatisé, et plus l'altération du consentement d'autrui sera le « fait » du système d'IA. Se posera alors à nouveau la question, dans un contexte de forte automatisation, de cette capacité autonome de l'IA d'altérer le consentement. Sans doute la logique juridique sera-t-elle sauvée en considérant que le déployeur (personne physique ou morale) d'un système d'IA susceptible d'évoluer de la sorte doit en tout état de cause en supporter les errements...

Sombre tableau pour les contractants que celui qui vient d'être dressé, alors même que les technologies sont souvent présentées comme neutres. Partant de ce dernier postulat, faisons alors l'hypothèse que les systèmes d'IA puissent, tout comme ils le menacent, renforcer et sublimer le consentement.

2. L'espoir d'une IA protectrice

En miroir à la première partie de cette contribution, le degré variable d'autonomie des systèmes d'IA suggère de distinguer la contribution annexe de ces systèmes à la protection du consentement (2.1), de la contribution autonome à ladite protection (2.2)

2.1. La protection augmentée du consentement.

Il est concevable que des systèmes d'IA soient conçus pour permettre à des individus, particulièrement des consommateurs, de mieux défendre leurs intérêts, notamment face à des co-contractants structurellement mieux in-

⁸ Ibid., p. 56, signalant que ces sujets sont encore largement débattus en doctrine et citant om Allen and Robin Widdison, 'Can computers make contracts?' (1996) 9 Harvard Journal of Law & Technology 26.

formés. Plus largement, il s'agit-là d'une déclinaison d'une des tendances importantes du développement de l'IA, à savoir l'«*agentic AI*». Les systèmes d'IA en question sont des « agents » (des préposés dans un langage juridique plus classique) qui assistent donc le contractant principal dans le processus contractuel, principalement bien entendu afin de garantir au mieux les intérêts dudit contractant. Cette tendance actuelle et prometteuse des différentes technologies d'IA a d'ailleurs donné lieu récemment à une étude par le European Law Institute,⁹ en témoignage à son potentiel.

Vu de la fenêtre des consommateurs, et des contractants vulnérables en général, des systèmes d'IA paramétrés et instruits pour protéger au mieux les intérêts de cette catégorie de contractants, dans une logique plus générale régulièrement vantée, à savoir celle de l'«*empowerment*», ne peuvent être accueillis que favorablement. Concrètement, de tels systèmes peuvent alerter le contractant face à des clauses contractuelles suspicieuses, ou bien plus largement face à des communications commerciales, pour utiliser le langage du droit des pratiques commerciales déloyales, qui ne semblent pas correspondre à la réalité. Il pourrait être imaginé, par exemple, que lesdits systèmes soient interconnectés et alimentés régulièrement en *inputs* par des bases de données, notamment juridiques, leur permettant de détecter le caractère potentiellement frauduleux de messages publicitaires et de sollicitations de la part de cocontractants potentiels.

Plutôt que de se substituer au contractant, ce type d'IA est une source de conseil dans le processus décisionnel et permettrait d'éviter la conclusion de contrats fruits d'un consentement altéré. Toute louable que puisse être cette perspective, il n'en demeure pas moins qu'elle n'est pas exempte de risques. En effet, en se focalisant sur le conseil au contractant dont le consentement risque d'être vicié, on se détourne quelque peu de la logique traditionnelle du droit de la consommation où les déséquilibres structurels sont corrigés notamment par des obligations multiples d'information à la charge du professionnel et par des sanctions sévères des pratiques commerciales trompeuses, voire agressives. Le mythe d'un consommateur *empowered* ne résisterait pas longtemps au retour à une logique plus classique des rapports contractuels.

2.2. La protection automatisée du consentement

Tout comme un système d'IA pourrait « spontanément » altérer le consentement d'un cocontractant, il est concevable de déléguer intégralement ou presque intégralement le processus contractuel à ce type de système et lui permettre de déceler les tentatives d'autrui de vicier le consentement du contractant principal. Il s'agit-là d'un degré supplémentaire d'intervention par rapport à l'hypothèse faite dans la section précédente. Le système d'IA est, dans cette configuration, l'agent du contractant et répond à des instructions initiales dont on suppose qu'elles prévoient la protection des intérêts dudit contractant. Ce type de système plus « intelligent » servirait alors

⁹ N. 4.

d'écran entre le contractant et son environnement juridique et pourrait être conçu comme un outil de rééquilibrage structurel des rapports contractuels, pour autant qu'il soit performant. Or, bien qu'il y ait un marché pour ce type d'outil, il est probable que différentes gammes soient développées, les plus performantes étant réservées à des contractants plus fortunés, qui pourraient potentiellement reproduire les déséquilibres existants dans le marché juridique classique.

On pourrait alors imaginer que la puissance publique développe des systèmes d'IA, gratuits et performants, que les individus, particulièrement ceux dont le consentement est le plus vulnérable, pourraient utiliser comme interface dans leur commerce juridique quotidien. Un tel système, déployé sur une plateforme publique unique, serait en entraînement constant pour tenir compte des évolutions législatives et des pratiques de marché, et demeurer «*compliant by design*». Cette approche procurerait également l'effet bénéfique de susciter la confiance dans l'économie numérique et de renforcer la sécurité juridique, qui constituent deux ingrédients parmi les plus importants pour un fonctionnement efficace d'une économie. Il est permis de fonder cet espoir, bien que les tendances de marché actuelles puissent rendre moyennement optimiste.

Mais alors, dans l'hypothèse où un système d'IA se substitue au contractant, est-ce finalement vraiment le consentement du contractant personne physique ou morale qui est protégé, ou bien ne le prive-t-on pas de cette expression originaire de liberté et d'autonomie? l'autonomie des systèmes d'IA hypothèque inversement l'autonomie de l'individu. Il faudrait alors choisir de deux maux le plus doux, la protection du consentement de l'individu, ou à tout le moins de ses intérêts, ou bien la préservation de son autonomie. Le format de cette contribution ne permet malheureusement pas d'explorer plus loin ce dilemme, qui se niche pourtant au cœur de la société numérique.

3. Conclusion

La question de la relation entre IA et consentement relève, comme l'ensemble de la discussion autour des nouvelles technologies digitales et particulièrement de l'IA, d'un débat de politique publique et de politique juridique. En effet, le traitement des effets structurels, disruptifs et très rapides, de la quatrième révolution industrielle ne peuvent être laissés aux forces de marché et à l'évolution « naturelle » dudit marché. L'un des risques principaux des évolutions technologiques et juridiques récentes et à venir est celui d'un façonnage excessif des règles juridiques par la technologie, dans une logique de normalisation technique. C'est déjà ce qui est constaté avec l'AI Act européen, dont l'approche *ex ante* consiste à réguler les produits (les systèmes d'IA) afin qu'ils puissent être commercialisés, sans s'intéresser suffisamment à ce stade aux droits des citoyens, dont ces derniers pourraient se prévaloir *a posteriori*, notamment pour obtenir du juge qu'il rectifie le manquement à des règles traditionnelles du droit.

Dans le domaine contractuel et en ce qui concerne le consentement, la question de politique juridique consiste principalement à choisir entre plusieurs options possibles, dont chacune comporte des coûts, qu'ils soient juridiques ou économiques. Tout d'abord, laisser prospérer des systèmes de plus en plus performants, sans encadrer strictement leur conception et leur utilisation, pourrait assez rapidement sonner le glas du consentement de catégories entières de contractants, particulièrement les plus vulnérables. En effet, soumis à des algorithmes puissants, embarqués dans des systèmes auto-apprenants et évolutifs, les individus seraient manipulés à une échelle encore plus vaste que ne l'ont permis les techniques traditionnelles de marketing et de conviction. Ensuite, permettre le développement d'un écosystème vertueux, où se développent des outils performants au service des contractants les plus vulnérables, suppose selon toute vraisemblance une intervention publique mobilisant les finances publiques et un interventionnisme qui n'est plus vraiment au goût du jour, pour des raisons souvent budgétaires, et qui entame la conception traditionnelle et libérale du processus contractuel.

Il convient à tout le moins que ces sujets soient davantage placés dans le débat public, afin que des choix soient faits le plus démocratiquement possible, considérant les effets massifs sur un certain nombre de paradigmes de l'ordre juridique et de la société dans son ensemble. Gageons qu'il en soit ainsi, dans une époque où la notion de consentement n'a aucunement perdu de sa vigueur, en tout cas dans les nombreuses incantations dont elle l'objet.



