INTELLIGENCE ARTIFICIELLE ET MANIPULATIONS DES COMPORTEMENTS DE MARCHÉ : L'ÉVALUATION *EX ANTE* DANS L'ARSENAL DU RÉGULATEUR

Nathalie de Marcellis-Warin, Frédéric Marty, Eva Thelisson, Thierry Warin

De Boeck Supérieur | « Revue internationale de droit économique »

2020/2 t. XXXIV | pages 203 à 245

ISSN 1010-8831
ISBN 9782807394063

Article disponible en ligne à l'adresse :

https://www.cairn.info/revue-internationale-de-droit-economique-2020-2-page-203.htm

INTELLIGENCE ARTIFICIELLE ET MANIPULATIONS DES COMPORTEMENTS DE MARCHÉ : L'ÉVALUATION *EX ANTE* DANS L'ARSENAL DU RÉGULATEUR

Nathalie DE MARCELLIS-WARIN
Polytechnique Montréal, Cirano & Obvia

Frédéric MARTY
CNRS – Université Côte d'Azur & Cirano

Eva THELISSON
AI Transparency Institute, MIT Connection Science

Thierry WARIN
HEC Montréal, Cirano & Obvia

Résumé : *Le développement de l'économie numérique pose des problèmes inédits par leur ampleur en matière de possibles manipulations de marché et manipulations des choix des consommateurs. Des stratégies trompeuses et déloyales dans le champ du droit de la consommation peuvent coexister et se renforcer mutuellement, avec des infractions dans le champ de la concurrence, qu'il s'agisse de collusion algorithmique ou d'abus de position dominante. Face à la difficulté de détecter et sanctionner ces pratiques, l'effet dissuasif de la sanction, notamment pour des dommages possiblement irréversibles, est à questionner. À cette fin, cet article envisage les outils de supervision disponibles tant pour les autorités responsables de la supervision des marchés, les consommateurs ou les parties prenantes des entreprises concernées.*

Revue Internationale de Droit Économique – 2020 – pp. 203-245 – DOI: 10.3917/ride.342.0203

1 INTRODUCTION : ÉCONOMIE NUMÉRIQUE ET RISQUES POUR LES CONSOMMATEURS – ILLUSTRATIONS PAR LES CAS ZOOM ET APPLE

Les difficultés rencontrées dans l'identification de pratiques mises en œuvre par les firmes numériques porteuses de dommages au consommateur et à la concurrence peuvent être illustrées par deux procédures négociées (des *consent decrees*) conclues aux États-Unis en novembre 2020. Celles-ci ont concerné Apple et Zoom.

Le 18 novembre 2020, l'Avocat général de Californie, Xavier Becerra, annonce un accord de 113 millions de dollars américains entre l'État de Californie et Apple à propos de l'algorithme de gestion des batteries et de la performance des appareils fonctionnant sous iOS (Becerra, 2020). La raison tient en une fausse représentation de la durée de vie des batteries et du mécanisme algorithmique de correction qui a été mis en place pour gérer leurs défaillances (pouvant conduire à l'arrêt inopiné des terminaux). Néanmoins, ce dernier altérait la performance des appareils. L'argument est donc un argument de fausse représentation de la qualité et d'une correction qui en réalité masquait le problème initial plutôt que de le corriger[1]. Apple a dû s'engager à fournir une information à ses clients leur permettant de prendre leurs décisions d'achat en connaissance de cause et de les informer de façon transparente sur les problèmes affectant leurs terminaux et des conséquences de ses actions

1. https://oag.ca.gov/news/press-releases/attorney-general-becerra-announces-113-million-multistate-settlement-against.

correctrices sur leurs performances. Le défaut d'information au client ou la mise en œuvre de stratégies trompeuses ou déloyales peut passer par de nombreux canaux dans le monde numérique. Cette même préoccupation quant à la transparence et à la véracité de l'information délivrée aux consommateurs se retrouve dans la procédure négociée qui a mis fin à la procédure engagée par la Federal Trade Commission (ci-après FTC) contre Zoom[2].

Zoom est l'outil de visioconférence qui a connu la plus forte croissance durant la crise de la covid-19. Créée en 2011, la société avait 10 millions d'utilisateurs par jour en 2019, essentiellement des indépendants et des PME. En jouant sur son modèle *freemium* et en étendant les conditions de gratuité au profit des établissements scolaires et universitaires, elle a atteint le seuil de 300 millions d'utilisateurs quotidiens dès le mois d'avril 2020. Son chiffre d'affaires, qui était de 622,7 millions de dollars américains en 2019, a atteint 328,2 millions au premier trimestre 2020, 663,5 au deuxième trimestre[3] et devrait s'établir entre 685 et 690 au troisième trimestre.

En raison de son activité, Zoom collecte des données sur ses utilisateurs, sur les participants aux réunions. Il stocke également les fichiers correspondants aux échanges audio et vidéo, aux messages écrits instantanés, etc. Il s'agit donc de traiter une communication trompeuse vis-à-vis des utilisateurs quant au degré de sécurité attaché au service conjugué à l'absence de mesures préventives contre des menaces « *commonly known and reasonably foreseable* »…, en d'autres termes, la communication délivrée aux utilisateurs est à la fois trompeuse et cache l'absence de mise en œuvre d'un standard de précaution raisonnable.

Les pratiques dénoncées par la FTC étaient au nombre de quatre. Premièrement, Zoom met en œuvre une communication trompeuse quant au cryptage des vidéoconférences. Un cryptage point à point doit permettre d'éviter qu'un tiers puisse accéder aux échanges. Or, seule une partie des échanges (ceux via Zoom *connect*) est conforme à cette revendication. Deuxièmement, l'entreprise annonce un cryptage à 256 octets alors que ce dernier ne correspond qu'à une clé à 128 octets. Le troisième grief tient à une communication trompeuse sur le niveau de sécurité du stockage en ligne des vidéos. Alors que la compagnie annonce un stockage crypté dès la fin de la réunion, ce cryptage peut prendre dans les faits jusqu'à 60 jours, le temps que Zoom rapatrie les fichiers vers ses propres infrastructures de stockage. Le quatrième grief porte sur une stratégie déloyale de contournement des mesures de sécurité et de protection de la vie privée et de la confidentialité des utilisateurs mises en place par

2. https://www.ftc.gov/system/files/documents/cases/1923167zoomcomplaint.pdf.
3. https://investors.zoom.us/news-releases/news-release-details/zoom-reports-second-quarter-results-fiscal-year-2021.

des entreprises tierces. Il s'agit en l'espèce de l'installation d'un serveur web sur les ordinateurs, tablettes et téléphones intelligents des utilisateurs permettant une connexion en un clic (avec activation automatique de la caméra sans demande explicite de consentement). Cette automaticité crée une vulnérabilité significative en cas d'intrusion malveillante (notamment des attaques de type *remote control execution – RCE*). Si un correctif d'Apple, appliqué au travers d'une mise à jour, a supprimé à distance ce serveur web des appareils concernés, ce dernier peut se réinstaller seul… et ce, même si l'utilisateur a désinstallé l'application Zoom.

La procédure entamée par la FTC a trouvé un terme dans un *Consent Decree* prononcé le 9 novembre 2020. Selon les termes de l'accord négocié entre la FTC et Zoom, la compagnie s'engage à mettre en œuvre un programme de sécurisation de son service et de mettre fin aux manœuvres trompeuses et déloyales (*deceptive and unfair practices*) qui ont compromis la sécurité des utilisateurs et la solidité des bases sur lesquelles ces derniers prennent leurs décisions. Les faux engagements en matière de cryptage ont pu donner un faux sentiment de sécurité et donc altérer les termes des arbitrages des consommateurs entre les différents services en concurrence.

Le fait que la procédure ait trouvé une issue au travers d'un programme de conformité a donné lieu à deux opinions dissidentes au sein du collège de la FTC (composé de cinq membres). Ces deux opinions, de Rebecca Slaughter et de Rohit Chopra, sont d'autant plus intéressantes qu'elles s'interrogent sur la stratégie de croissance de la firme et sur ses possibles conséquences sur ses clients en matière de sécurité et de protection de la vie personnelle. Rebecca Slaughter met en regard, dans son opinion dissidente, l'impact de cette stratégie sur la protection du consommateur. Rohit Chopra insiste sur les conséquences des agissements de Zoom en termes concurrentiels.

Rebecca Slaughter souligne le fait que l'information trompeuse donnée aux utilisateurs a dissimulé un arbitrage entre la rapide augmentation d'échelle (visant à atteindre une taille critique) et la protection des utilisateurs. La qualité de l'expérience utilisateur « visible » (la connexion en un clic) a été préférée à la qualité moins aisément apparente. Il ne s'agit pas, pour Rebecca Slaughter, seulement d'une simple tromperie, mais de manœuvres volontaires visant à contourner et donc à affaiblir les mesures de protection mises en place par d'autres entreprises au profit de leurs utilisateurs[4]. Son opinion dissidente montre également que la question de la sécurité seule ne permet pas de résoudre

4. Rebecca Slaughter ajoute notamment que l'accès via le connecteur LinkedIn sur Zoom donnerait accès aux profils LinkedIn des participants à une réunion même si ces derniers ont souhaité y accéder sans dévoiler leur identité. De la même façon, les différentes données (enregistrements des visioconférences) stockées sur les serveurs seraient d'autant plus accessibles que les structures url permettant d'y accéder sont prédictibles.

tous les problèmes reliés à la protection de la vie privée. En d'autres termes, la sécurité est une condition nécessaire, mais non suffisante à la protection de la confidentialité. D'où son regret quant à l'absence d'un programme de protection de la vie privée dans les engagements de Zoom :

> « *A more effective order would require Zoom to engage in a review of the risks to consumer privacy presented by its products and services, to implement procedures to routinely review such risks, and to build in privacy-risk mitigation before implementing any new or modified product, service or practice* »[5].

L'opinion dissidente du commissaire Rohit Chopra se dissocie de l'accord trouvé avec la FTC en insistant de son côté sur les dimensions concurrentielles (FTC, 2020). Pour lui, « *deception distorts competition* » (Chopra, 2020). L'idée est que la course à la taille critique (et donc à la part de marché à partir de laquelle un service numérique bascule vers le degré de dominance qui fait de lui l'application par défaut pour les consommateurs) supposait la divulgation des informations trompeuses quant à son niveau de sécurité. En d'autres termes, « *when companies need to act quickly to exploit an opportunity, deploying deception to steal users or sales from competing players is tantalizing* ».

La conquête de la dominance se fait sur la base d'une concurrence déloyale vis-à-vis des autres opérateurs. Cette dominance trouve son fondement dans la diffusion d'une information trompeuse aux consommateurs. Comme l'indique Rohit Chopra : « *when companies deploy deception, this harms customers and honest competitors, and it distorts the market place* ». Rebecca Slaughter reliait sécurité des données et protection des utilisateurs, Rohit Chopra relie la protection du consommateur à la protection d'une concurrence équitable (*fair competition*[6]). Le problème qu'il met en exergue tient au fait qu'avant la pandémie la majeure partie des utilisateurs payants de Zoom étaient des PME et des indépendants. Leurs décisions peuvent être d'autant plus significativement biaisées par une communication trompeuse qu'ils ne disposent pas des services informatiques des grandes entreprises : « *that's why they rely on representations made by those they purchase software and services from* ».

5. L'exemple mis en exergue est celui des engagements pris par Facebook vis-à-vis de la FTC en juillet 2019. https://www.ftc.gov/news-events/blogs/business-blog/2019/07/ftcs-5-billion-facebook-settlement-record-breaking-history.

6. De façon significative, Rohit Chopra adopte un vocabulaire proche de celui utilisé au début du siècle dernier par Louis Brandeis : « *We should all be questioning whether Zoom and other tech titans expanded their empire through deception* ». Pour Brandeis, il ne pouvait y avoir de situation de monopole qui ne procède de stratégies de monopolisation… Dans son opinion dissidente dans l'affaire *FTC v. Graz* (*FTC v. Gratz*, 253 U.S. 421, 1920), Louis Brandeis, alors juge à la Cour suprême des États-Unis, soulignait que, lorsque le *FTC Act* avait été promulgué en 1914, « *the belief was widespread that the great trusts had acquired their power, in the main, through destroying or overreaching their weaker rivals by resort to unfair practices* ».

Dans les cas Apple et Zoom, l'argument est un argument de différenciation verticale. L'annonce d'un niveau de sécurité à 256 octets, par exemple, positionne l'entreprise à un certain niveau de différenciation verticale aux yeux des consommateurs. En offrant un produit de qualité intrinsèque inférieure à ce niveau de différenciation verticale, Zoom bénéficie d'avantages vis-à-vis de ses concurrents en ne supportant pas les coûts de sécurité qui auraient pu entraver sa croissance. L'exemple de Zoom est un exemple bien particulier qui porte sur la fausse information et le manque d'incitations proposées aux entreprises technologiques. Dans un tel cas, les incitations à améliorer la qualité du produit peuvent être discutées.

À la vue de ces deux exemples récents, la question du positionnement en matière de différenciation verticale est extrêmement importante pour l'industrie des entreprises technologiques. Rappelons que ces entreprises sont des entreprises qui développent des innovations d'architecture et qui en conséquence contribuent à mettre en place la nouvelle infrastructure technologique mondiale. Les technologies et les modèles d'affaires reposant sur ces technologies font aussi que tout va très vite. Par conséquent, si le diagnostic selon lequel la fausse information crée de la distorsion de concurrence est vrai, il n'en reste pas moins que le contrôle *ex post* peut montrer des limites dans la dissuasion de pratiques ne correspondant pas aux standards d'une concurrence à égalité des armes et d'une protection des consommateurs.

Nous allons nous intéresser, dans cet article, au cas plus général des manipulations de marché par ces entreprises. Nous évoquerons le dilemme actuel : les entreprises technologiques offrent des produits et services et peuvent faire de la fausse représentation. Le risque illustré par ces deux affaires est que les conséquences d'une détection *ex post* (procédures négociées et éventuellement réparations) soient insuffisantes à dissuader, *ex ante*, ces pratiques.

Cependant, à une époque où les entreprises technologiques reposent sur des modèles d'affaires reliés aux données massives et à l'utilisation de modèles de décision reposant sur les techniques d'intelligence artificielle, tout va exponentiellement plus vite, y compris les stratégies de développement et de commercialisation des produits et services de ces entreprises. Dans ce contexte, nous pensons que le système actuel doit développer de nouveaux outils pour éviter des comportements anticoncurrentiels, voire non éthiques.

Aux mécanismes de régulation *ex post*, nous allons proposer un argument en faveur d'un mécanisme *ex ante* : un indice de transparence pour les entreprises technologiques. Une approche *ex ante* nous semble être un outil important à ajouter à l'arsenal du régulateur pour plusieurs raisons. Il s'agit de garantir sur le marché des comportements responsables évitant l'occurrence de risques majeurs et irréversibles. Ces derniers peuvent procéder de dommages irréparables à la concurrence (dominance irréversible, collusions tacites impossibles

à sanctionner…) ou de conséquences majeures pour les consommateurs (lesquelles sont liées à des discriminations ou à de graves violations de la vie privée). Ce point permet de mettre en exergue la notion de décisions hautement conséquentes (*high stakes*) qui décrit les conséquences de choix algorithmiques qui peuvent avoir des conséquences majeures sur la concurrence ou sur les consommateurs. Cela conduit enfin à la prise en compte de mesures *ex ante* permettant aux régulateurs des marchés de détecter certaines de ces pratiques avant que leurs effets soient irréversibles et aux entreprises elles-mêmes de prévenir de telles difficultés, à la fois en matière de gestion de leurs risques juridiques et de mise en œuvre de leur stratégie propre de responsabilité sociale.

Notre propos dans la suite de cet article tient à l'analyse des pratiques des entreprises sur les marchés en ligne qui peuvent à la fois manipuler les choix des consommateurs et fausser la concurrence. Ces pratiques, qui peuvent reposer sur des manipulations algorithmiques, sur des architectures de présentation des choix volontairement trompeuses ou sur la diffusion de fausses informations sur la qualité des services, présentent deux caractéristiques particulièrement problématiques. Premièrement, elles ne sont pas perceptibles *ex ante* par les consommateurs qui font face à une information incomplète et asymétrique. Même après la consommation du service, le client ne peut réellement évaluer sa qualité. Il s'agit donc bien plus de « biens de confiance » que de « biens d'expérience » (Karpik, 2013). Deuxièmement, la détection par les autorités de supervision des marchés est particulièrement difficile pour deux raisons : une tient aux difficultés de détection des stratégies algorithmiques, une autre au fait que certaines de ses pratiques se situent à la confluence de la protection des données, de la protection du consommateur et de celle de la concurrence.

Des autorités comme la FTC aux États-Unis ou la CMA au Royaume-Uni, dont le périmètre des attributions est particulièrement large, peuvent être mieux placées que d'autres autorités spécialisées dans tel ou tel domaine. À ce titre, le recours aux procédures négociées, comme moyen de règlement des procédures ou dans le cadre d'enquêtes de marché sur le modèle des *market investigations* britanniques instaurées par l'*Entreprise Act* de 2002, peut être une piste intéressante à considérer. En effet, se pose la question de la capacité des seules sanctions *ex post* à répondre à ces risques. Les deux exemples cités *supra* peuvent en effet illustrer quelques-unes des voies possibles pour prévenir ces pratiques avant qu'elles ne créent des dommages ou pour permettre aux parties prenantes elles-mêmes de modifier leurs comportements pour en limiter les conséquences.

La première de ces voies est celle de la régulation par coup de projecteur (*sunshine regulation*). Révéler aux consommateurs les pratiques d'une entreprise donnée peut permettre de l'amener à modifier son comportement et conduire à une sanction par le marché. Cette voie est à la source d'une des divergences entre les commissaires de la FTC majoritaires et Rohit Chopra. Ce dernier aurait

souhaité que Zoom soit obligé d'informer tous ses utilisateurs quant à ses agissements passés. Pour la majorité, « *the conduct at issue was broadly publicized and we believe the Commission's press release and business and consumer education will provide ample information to consumers to learn more* ».

La seconde voie est celle du contrôle de la firme par ses différentes parties prenantes, notamment les investisseurs. Un manquement constaté *ex post* par les autorités de supervision des marchés peut conduire à une sanction additionnelle au travers du retrait de certains financeurs engagés dans des politiques éthiques et responsables. Cependant, ces derniers peuvent légitimement vouloir s'assurer de la conformité de la stratégie de la firme à leurs valeurs en dehors de procédures publiques. Cela suppose la réalisation d'indicateurs permettant à l'ensemble des parties prenantes de détecter de tels risques. La responsabilité sociétale des firmes suppose qu'elles se dotent d'instruments d'auto-évaluation de leurs pratiques, mais également que soient développés par des institutions indépendantes des outils d'évaluation de la responsabilité algorithmique des firmes.

Ces outils sont d'autant plus importants que le développement de l'espace des décisions algorithmiques pose des problèmes inédits de compréhension des prédictions que ces derniers font (notamment du fait de la diffusion des outils d'intelligence artificielle). L'enjeu est celui de la redevabilité des choix et de la détection de pratiques manipulatoires.

Nous centrons notre article sur les manipulations algorithmiques et sur les conséquences en matière de protection de la vie privée et de la concurrence.

L'article se structure comme suit. Une deuxième section présente les risques que peut induire un recours croissant à l'I.A. dans les dispositifs de recommandations algorithmiques pour les consommateurs au travers de la réduction de leur liberté de choix, des manipulations comportementales ou encore par des conditions contractuelles personnalisées certes, mais déséquilibrées. Une troisième section montre que le dommage au consommateur peut aussi passer de façon indirecte par une atteinte à la concurrence, au travers de la consolidation de positions dominantes individuelles ou de l'augmentation des risques de collusion algorithmique. Une quatrième section s'attache aux solutions de régulation des algorithmes par d'autres algorithmes ou du moins par des procédures de suivi permettant une redevabilité quant au fonctionnement même des algorithmes mis en œuvre. Une première voie tient à l'utilisation d'outils algorithmiques par les autorités de supervision des marchés elles-mêmes pour détecter d'éventuels « *patterns* »[7] anormaux pouvant conduire à l'ouverture de

7. Le terme de *pattern* désigne la configuration qui ressort de l'observation des différentes données collectées. La forme prise par le nuage de points peut être un indice de pratiques contraires aux règles de concurrence.

procédures. Une seconde voie peut tenir à l'utilisation de contre-mesures algorithmiques par les consommateurs. Une cinquième section envisage la façon dont les « décisions hautement conséquentes » peuvent faire l'objet d'une attention particulière par les différentes parties prenantes de l'entreprise. La prise en compte de ces décisions peut conduire à la conception de dispositifs permettant aux firmes de garantir l'intégrité et l'éthique de leurs algorithmes au travers de méthodes de scores[8]. Nous présentons, à ce titre, l'indice de confiance, développé par l'AI Institute[9], en l'adaptant dans le cadre de cet article aux enjeux reliés aux droits de la consommation et de la concurrence.

2 L'INTELLIGENCE ARTIFICIELLE COMME VECTEUR DE RISQUES POUR LES CONSOMMATEURS

Trois types de dommages pour le consommateur peuvent être considérés[10]. Un premier dommage tient à la réduction des choix qui lui sont ouverts. Un deuxième dommage réside dans la possibilité de manipulation des choix. L'éventail des solutions disponibles n'est pas artificiellement refermé, mais le comportement du consommateur est altéré par la production de stimuli destinés à biaiser sa décision. Un troisième type de dommage correspond à ce que nous appelons un abus d'exploitation. La capacité à prédire les caractéristiques du consommateur (expertise technique, capacité à payer, etc.) permet de faire des offres conduisant à extraire la totalité de son surplus (ce que ne permettraient pas de faire des prix uniformes ou des prix imparfaitement différenciés) ou conduisant à des offres discriminatoires, que cela soit en matière de prix ou de qualité des produits et services proposés.

8. Dans le secteur bancaire, la méthode de score désigne une technique d'analyse destinée à diagnostiquer préventivement les difficultés des entreprises. Par extension, la méthode recouvre un ensemble d'outils permettant d'identifier, domaine par domaine, les forces et les faiblesses relatives d'une entreprise par rapport à un objectif donné par les actionnaires, une entreprise concurrente cible ou un indice composite traduisant la performance moyenne des entreprises du marché concerné.
9. AI Transparency Institute est un organisme à but non lucratif dont les travaux portent sur la gouvernance de l'intelligence artificielle, sur sa transparence et son caractère explicable, et sur la responsabilité numérique sociétale des firmes.
10. Dans le même temps, le recours à l'I.A. est indubitablement porteur de gains d'efficience (de Marcellis-Warin et Warin, 2020).

2.1 Réduction de l'espace des choix possibles

L'I.A. permet des changements de rupture des paradigmes commerciaux habituels. L'I.A. peut être utilisée dans le cadre de stratégies conduisant, au travers de préconisations ou de recommandations de plus en plus finement ciblées, à une limitation de la liberté de choix des consommateurs. Ces derniers peuvent voir l'éventail de leurs choix se réduire en fonction de leur consommation passée, ou encore en fonction du segment de clientèle auquel l'algorithme les rattache. L'I.A. est, sur la base de l'apprentissage machine, d'abord un outil de prédiction (Agrawal, Gans et Goldfarb, 2018). Ils peuvent être, en d'autres termes, enfermés dans l'équivalent d'une bulle de filtre. Un tel effet peut être aggravé par le passage de certaines plateformes d'une logique de *shopping-then-shipping* à une logique de *shipping-then-shopping* (Agrawal, Gans et Goldfarb, 2017). Le client peut subir un coût pour réexpédier le produit quand bien même le retour serait gratuit. Ce coût peut s'apprécier en termes de temps perdu par exemple. La nature même du produit reçu mais non sollicité pourrait également le heurter.

Le recours à l'I.A. peut également faciliter des pratiques pouvant conduire à manipuler les choix des consommateurs au travers d'une compréhension fine de leurs comportements ou d'une estimation précise de leur capacité maximale de paiement. En effet, comme le notent Ezrachi et Stucke (2020) : « *[…] in a data driven economy, personal data on user behavior, preferences, weaknesses, and habits is the new currency for the advertising – and marketing dependent – business models* ». Ces capacités passent par le contrôle de données massives, diversifiées et sans cesse renouvelées et par la maîtrise d'instruments analytiques permettant de personnaliser les offres vis-à-vis des cadres décisionnels des consommateurs ou de mieux prédire les stratégies des autres opérateurs (Marty et Warin, 2020b, 2020a). La prise de contrôle d'Onavo par Facebook ou celle de Looker par Google témoigne de l'importance des capacités de maîtrise de l'environnement concurrentiel par la possibilité technique de prévoir de mieux en mieux le présent.

Plusieurs exemples de limitation des capacités de choix des consommateurs pourraient être ajoutés à la bulle de filtre et aux modèles de *shipping-then-shopping* présentés *supra*. Certains sont liés à des facteurs de coûts. Les complémentarités entre équipements accroissent fréquemment les coûts de changements entre écosystèmes. Des biais peuvent ensuite procéder des modes de proposition des choix. Le cas des assistants personnels montre comment les options proposées peuvent se réduire à un très faible éventail. Le dernier exemple, développé par Ezrachi et Stucke (2020), tient à la possibilité de maîtrise de la diffusion des innovations dans les écosystèmes par l'opérateur pivot de chacun de ses derniers.

Étapes du processus de diffusion	Stratégie favorable du pivot	Stratégie défavorable du pivot
Connaissance	Capacité de proposition, de mise en avant.	Réduire les possibilités d'information quant à une innovation potentiellement disponible ou l'accès à des informations quant à son fonctionnement (par manipulation algorithmique du moteur de recherche par exemple, par déréférencement des sites…).
Persuasion	Capacité de ciblage, de mise en évidence de l'adéquation aux besoins personnalisés ; stratégies vis-à-vis de l'attention ; identification de possibles adopteurs précoces et diffusion d'informations personnalisées vis-à-vis de suiveurs potentiels.	Production d'avis négatifs ou création de frictions de nature à rendre plus difficile le téléchargement ou l'interopérabilité avec les différents services rendus par l'écosystème[11].
Décision	Marketing personnalisé ; essais gratuits ; jeu sur les recommandations d'amis.	Blocage par friction : jeu sur le biais comportemental de statu quo – les paramètres par défaut sont rarement modifiés par les agents, quelles que soient leurs préférences[12].
Mise en œuvre	Facilitation des adaptations, des corrections des bogues.	Les utilisateurs peuvent être sans cesse redirigés vers des options moins performantes, mais dépendantes de l'écosystème.
Confirmation	Redirections par les outils d'assistance vers l'innovation en question.	La firme pivot peut dégrader la performance des services complémentaires fournis par le concurrent pour réorienter les consommateurs vers un service mieux contrôlé.

Pour illustrer leur propos, Ezrachi et Stucke (2020, p. 42) s'appuient sur le modèle de diffusion de l'innovation de Rodgers (2003). L'adoption d'une innovation par un individu donné est décrite comme suivant un processus en cinq phases. La première est celle de la connaissance. Il faut que l'individu soit informé de la disponibilité d'une solution innovante et de ses fonctions. La deuxième est celle de la persuasion. C'est à travers elle que l'individu forme

11. Il s'agit de la notion de *bad sludges* que nous détaillerons *infra*.

12. « *As behavioral economics littérature shows, the setting of the default can often determine the outcome (even when transaction costs are minimal)* » (Ezrachi et Stucke, 2020, p. 48). Comme le relèvent Ezrachi et Stucke, l'inertie (le biais du statu quo) n'est pas la seule raison pour laquelle des consommateurs peuvent conserver des options par défaut qui ne sont pas les meilleures pour eux. Si ces derniers ont des compétences standard, ils peuvent considérer que les choix par défaut sont ceux les plus favorables en termes de qualité et de performance du service.

des anticipations favorables ou défavorables vis-à-vis de celle-ci. La troisième est celle de la décision d'adoption ou de non-adoption. La quatrième celle de la mise en œuvre. La cinquième, enfin, celle de la confirmation de l'adoption. Elle peut être confortée par l'observation du choix des tiers ou au contraire être négativement affectée par des messages négatifs.

L'action stratégique des plateformes – si elles disposent de capacités de détection des comportements de leurs utilisateurs – peut jouer favorablement pour l'adoption d'une innovation développée en interne (ou par un complémenteur privilégié) et défavorablement pour une innovation développée par une firme indépendante. La première variante peut expliquer pourquoi les écosystèmes numériques favorisent des adoptions précoces et massives des innovations. La seconde peut expliquer pourquoi des innovations n'arrivent pas à se diffuser.

2.2 La manipulation du comportement du consommateur

L'ADN des marchés est en fin de compte le mécanisme des prix. Ce dernier joue une fonction importante : il informe les acteurs du marché avant qu'ils prennent une décision. Il existe une abondante littérature sur le concept de « valeur de l'information », et certains auteurs l'ont examiné dans le contexte des stratégies de prix sur les marchés numériques (Warin et Leiter, 2012 ; Warin et Troadec, 2016) ainsi que sous un angle d'économie du droit (Marciano, Nicita, et Ramello, 2020).

Avec l'I.A., certaines entreprises ont désormais accès à l'information agrégée et à la valeur de l'information pour le client, grâce notamment à des systèmes de recommandation. Grâce à ce riche accès à la valeur de l'information, l'I.A. peut être utilisée pour modéliser le comportement des consommateurs et créer une incitation à l'achat (*pitch* émotionnel, *dark nudge*…) au bon moment. Ces problèmes dépassent le cadre de la seule I.A. dans la mesure où ils peuvent être observés dans le cadre des algorithmes traditionnels. Par exemple, de nombreux sites marchands peuvent mettre en place des stratégies de tarification au goutte-à-goutte (Rasch, Thöne et Wenzel, 2020) ou de partitionnement des prix. Le client peut être engagé dans un processus d'achat par un prix d'appel attractif et ne « découvrir » le prix complet que plus tard. Le temps passé à remplir les pages suivantes lui fera oublier le prix des concurrents consulté au début de sa recherche, ou il hésitera.

La notion de *dark patterns* illustre ces pratiques qui peuvent être aggravées par les performances de l'I.A. (Stigler Center, 2019). Elle recouvre l'ensemble des méthodes de profilage, de propositions algorithmiques ou encore d'interfaces utilisateurs qui peuvent restreindre la capacité de faire un choix libre et

éclairé de la part du consommateur. Les *dark patterns* rassemblent *dark nudges* et *bad sludges*. Ils recouvrent les stratégies qui accroissent l'opacité des choix pour les consommateurs, qui rendent plus difficile l'expression libre de leurs préférences ou qui les conduisent à prendre des décisions qu'ils n'auraient pas prises spontanément.

Des *dark patterns* peuvent être produits pour conduire le consommateur à prendre des décisions qui ne vont pas dans le sens de ses préférences. Alors qu'un *nudge* (positif) participe théoriquement d'une logique de *parternalisme libéral* – conduire l'individu à se comporter dans un sens conforme à son intérêt et/ou à l'intérêt général –, les *dark nudges* visent à le conduire à agir dans un sens non conforme à ses intérêts (Thaler, 2018 ; Sustein, 2019). Il s'agit donc de manipuler les choix des consommateurs en altérant volontairement leurs préférences, voire en créant celles-ci, au travers de l'exploitation de biais cognitifs (*framing effect, sunk cost fallacy, anchoring…*). Les *dark sludges* peuvent dès lors se définir comme « *an evil nudge […] that can exploits [online consumers] cognitive biases to persuade them to do something that is undesirable, typically by introducing exessive friction into choice architecture* » (Sunstein, 2020).

Il convient de distinguer au sein des *dark patterns,* la nuance entre *bad nudges* et *bad sludges*. Un *nudge* peut se définir comme un encouragement, un petit coup de coude qui conduit l'agent à agir dans un sens donné. C'est une poussée vers une action. Elle est souvent présentée comme positive (l'agent est incité à agir dans son intérêt dans la mesure où il ne le ferait pas spontanément). Elle peut cependant être négative. C'est par exemple le *bad nudge* qui est mobilisé dans le cadre d'un argument de vente émotionnel (*emotional pitch*) : on fait apparaître par une fenêtre intempestive une bannière conduisant à cliquer pour accéder à un service donné vis-à-vis duquel on sait que le consommateur a développé une assuétude. Il s'agit donc d'une poussée – dans le sens d'un *stimulus* – visant à faire « tomber » le consommateur du côté dont on sait qu'il a tendance à pencher.

La conception de la notion de *nudge*, dans la littérature d'économie comportementale, tenait principalement à une réflexion sur l'environnement de la décision des acteurs de façon à leur permettre de faire des choix plus avisés sans pour autant restreindre leur liberté. Il s'agissait donc de promouvoir une architecture de choix conciliant autonomie et « signalement » des meilleures options pour l'agent lui-même (Thaler et Sunstein, 2008).

Cependant, cette poussée – au travers de la conception de l'architecture de choix – peut également être exercée dans une visée bien moins altruiste. Elle peut, comme nous l'avons noté *supra*, être utilisée dans l'intérêt de la firme et au détriment du consommateur. Il ne s'agit plus de l'inciter à prendre une bonne décision pour lui, mais à le pousser à prendre une décision conforme aux

intérêts de la firme. Un *nudge* peut donc être aussi bien positif que négatif et peut donc dans ce cas-là participer d'un *dark pattern*[13].

À l'inverse, le terme *sludge* évoque une friction. Il s'agit, au sens premier, du fait de s'embourber, de perdre en mobilité. Dans notre contexte, la firme crée des difficultés artificielles pour empêcher les consommateurs d'exercer leur liberté de choix. Elle fait obstacle à l'identification des options les plus intéressantes et à leur sélection. Thaler (2019) donne un exemple simple, mais évocateur. Le consommateur prend la décision d'acquérir un bien ou de s'abonner à un service en ligne en prenant en compte l'offre d'un remboursement différé. Cependant, ce bénéfice est conditionné à l'envoi d'une preuve d'achat par courrier postal dans une période donnée (ou à la création fastidieuse d'un compte en ligne). Bon nombre de consommateurs, qui pourtant auront fondé leurs décisions sur ce rabais, ne le réclameront pas[14]. Comme le note Thaler (2019, p. 431) : « *because of this thick sludge, redemption rates for rebates tend to be low, yet the lure of the rebate still can stimulate sales – call it 'buy bait'* ».

Une *sludge* se définit comme une « *kind of friction, large or small, that people face when they want to go in one or another direction* » (Sunstein, 2019). Elle peut aussi bien permettre de prévenir une attitude dommageable de la part même du consommateur (freiner une frénésie d'achat, s'assurer de ses conditions d'éligibilité ou de ses caractéristiques, l'obliger à bénéficier d'un délai de réflexion) qu'au contraire faire obstacle à l'accès à des droits légitimes[15].

Tout comme un *nudge*, une *sludge* repose sur l'exploitation – par l'établissement stratégique d'une architecture de choix – des biais comportementaux des agents économiques. Quels peuvent être les biais qu'un mécanisme de friction peut mettre à profit ? Il peut s'agir, par exemple, des biais d'inertie

13. Sur la notion de *bad nudge*, voir Akerlof and Shiller (2017).

14. Sunstein (2019a) montre que les demandes effectives de remises différées s'établissent communément dans une fourchette de 10 à 40 %. Celles-ci sont pourtant prises en compte dans les décisions d'achat. Les consommateurs tendent à surestimer leur capacité future à prendre le temps nécessaire pour activer leurs droits. En conséquence, des demandes de remboursement par envoi d'un formulaire « papier » – *mail-in-forms* – jouent comme des *dark sludge*s (Edwards, 2007 ; Tasoff and Letzler, 2014). Une expérimentation menée par Tasoff et Letzler (2014) montre que les relances n'ont pas d'impact significatif sur le comportement des agents à l'inverse de la dématérialisation de la procédure (pas de formulaire à imprimer). Ce n'est pas une question d'attention, mais de simplification. Le problème est celui du frottement lié à la procédure d'activation de la ristourne. Il est possible de rapprocher cette expérimentation du modèle du *roach motel* que nous présenterons *infra* et qui explique les renouvellements tacites non désirés d'abonnement liés à la dissymétrie techniques entre les procédures d'abonnement et celles de désabonnement. À la simplicité et à la dématérialisation des premières répondent les délais et les exigences matérielles des secondes. La tendance à la renonciation à un avantage quelconque lié à une contrainte « administrative » – *i.e.* remplir un formulaire – a également été mise en évidence par Bettinger *et al.* (2012).

15. Les (*bad*) *sludges* peuvent à cette aune être aussi bien privées que publiques. En d'autres termes, les contraintes administratives peuvent agir comme autant de barrières à l'entrée pour certaines personnes pour accéder à des droits ou pour activer certaines procédures en leur faveur.

(Madrian et Shea, 2001), de procrastination (Akerlof, 1991) et de préférence pour le présent (O'Donoghue et Rabin, 2015).

La notion de *bad sludge* se retrouve d'ailleurs dans des contentieux entre les firmes pivots des grands écosystèmes numériques et leurs complémenteurs. C'est, par exemple, comme l'a relevé le rapport sur la concurrence dans le secteur numérique publié en octobre 2020 par la sous-commission à l'antitrust de la Chambre des représentants des États-Unis (House of representatives, 2020, p. 218), l'un des arguments soulevés par EPIC Games dans le cadre de sa plainte contre Google aux États-Unis[16]. Le développeur du jeu Fortnite insiste sur le fait que le contournement – techniquement possible et aisé – du magasin d'applications (Play Store) est rendu plus difficile et stressant pour le consommateur pour le dissuader de télécharger directement le jeu :

> « *Direct downloading on Android mobile devices, however, differs dramatically. Google ensures that the Android process is technically complex, confusing and threatening, filled with dire warnings that scare most consumers into abandoning the lengthy process. For example, depending on the version of Android running on a mobile device, downloading and installing Fortnite on an Android device could take as many as 16 steps or more, including requiring the user to make changes to the device's default settings and manually granting various permissions while being warned that doing so is dangerous. Below are the myriad steps an average Android user has to go through in order to download and install Fortnite directly from Epic's secure servers* ».

La friction ne se limite pas au téléchargement initial, mais également aux mises à jour. Toujours selon la plainte déposée en août dernier par EPIC :

> « *As if this slog through warnings and threats were not enough to ensure the inferiority of direct downloading as a distribution method for Android apps, Google denies downloaded apps the permissions necessary to be seamlessly updated in the background – instead allows such updates only for apps downloaded via Google Play Store. The result is that consumers must manually approve every update of a "sideloaded" app. In addition, depending on the OS version and selected settings, such updates may require users to go through many of the steps in the downloading process repeatedly, again triggering many of the same warnings. This imposes onerous obstacles on consumers who wish to keep the most current version of an app on their mobile device and further drives consumers away from direct downloading and toward Google's monopolized app store wish to keep the most current version of an app on their mobile device and further drives consumers away from direct downloading and toward Google's monopolized app store* »[17].

16. Complaint for injunctive relief, *Epic Games v. Google* LLC, n°3:20-cv-05671 – ND Cal., Aug., 13, 2020, § 96.
17. *Ibid*., § 98.

Quelle est ici l'utilité potentielle de ces *dark nudges* ? Maintenir le magasin d'applications en ligne comme un verrou d'accès à l'écosystème (*gatekeeper-side downloading*) pour assurer le pouvoir de régulation privé de l'entreprise pivot (*structuring power*). La maîtrise de l'écosystème passe donc par le fait d'imposer des frictions techniques et psychologiques pour faire pièce à une menace de perte de contrôle par le téléchargement direct d'applications[18] :

> « *As if this slog through warnings and threats were not enough to ensure the inferiority of direct downloading as a distribution method for Android apps, Google denies downloaded apps the permissions necessary to be seamlessly updated in the background – instead allows such updates only for apps downloaded via Google Play Store. The result is that consumers must manually approve every update of a "sideloaded" app* ».

La notion de *dark pattern* dépasse la manipulation des choix et des comportements. Ces dispositifs peuvent conduire les internautes à révéler déraisonnablement leurs informations personnelles. Dans un tel cas de *nudge*, la conception du site ou les modes de présentation des choix font que l'utilisateur va aller au-delà de ce qui est nécessaire ou de ce qu'il aurait accepté si son choix avait répondu à une rationalité de type 2 (Acquisti *et al.*, 2020). Il ne s'agit pas seulement d'exploiter les vulnérabilités des consommateurs ou des utilisateurs pour les inciter à effectuer des choix qui correspondent à des tendances qu'ils pourraient rationnellement essayer de réfréner, mais à l'extrême de susciter (*i.e.* de construire) ces préférences (Mulligan *et al.*, 2020). Le *dark pattern* peut donc résulter de la conception même d'un site. Il en est ainsi des *clickwraps* qui conduisent le consommateur à faire des choix en blocs pour des questions d'importance et de nature fort différentes (Obar and Oeldorf-Hirsch, 2018). Il existe donc des dispositifs « manipulatoires par conception ». Leur analyse n'est pas nouvelle, que cela soit dans le monde en ligne (Calo, 2013) ou même dans le monde hors ligne (Hanson et Kysar, 1999). Cependant, la nature du

18. Il peut être intéressant, pour notre propos, de comparer les termes du contentieux opposant Epic Games à Google à ceux l'opposant à Apple. La *complaint for injuctive relief* déposée devant l'US District Court for the Northern District of California, le 13 août 2020, décrit également une impossibilité d'accès aux clients d'Apple en dehors du magasin d'applications de la firme. Cependant, les pratiques en cause ne tiennent pas au fait d'imposer des frictions pour les utilisateurs qui seraient désireux de contourner le magasin d'applications, comme cela est le cas pour Google, mais à une impossibilité de le faire. Cette impossibilité tient alors à des restrictions techniques et contractuelles. Au point de vue technique, les clients ne peuvent pas installer un magasin d'applications alternatif à l'App Store sur leurs terminaux (pt. 58). Au-delà de ce verrouillage, le système d'exploitation iOS empêche, au travers de restrictions techniques, le téléchargement direct d'applications depuis des sites Internet en contournant le magasin d'applications (pt. 66). Enfin, le verrouillage est également contractuel. Les développeurs ne peuvent être présents sur l'App Store que s'ils s'engagent à ne pas rendre possible des téléchargements par des voies alternatives : « […] *to access the iOS userbase, app developpers must agree not to distribute or create app stores that could compete with Apple's App Store – whether they intend to distribute their or through the developer's own website* » (pt. 80).

parcours en ligne et les capacités de capter, traiter, déduire et créer des *stimuli* ciblés font que les effets sont d'une tout autre envergure.

Les modalités de mise en œuvre de ces pratiques peuvent être des plus basiques. Un internaute à qui il est demandé *n* fois ses préférences en termes de protection de ses données personnelles après chacun de ses refus aura tendance à accepter, soit par inadvertance, soit pour mettre un terme aux demandes[19]. Il en va de même pour les techniques décrites *supra* dans lesquelles le prix n'est révélé qu'à la fin du processus d'achat ou encore des transactions dans lesquelles sont « présélectionnées », de façon non évidente, des options que le consommateur n'aurait *a priori* pas souhaité souscrire[20].

De la même façon, la notion de *dark pattern* peut recouvrir les pratiques d'*emotional pitch*. Il s'agit de produire un *stimulus* qui ferait basculer le consommateur vers un acte d'achat. De tels *stimuli* peuvent être l'annonce (sous forme d'une fenêtre intempestive) d'une réduction de prix ou encore celle d'un nombre limité de produits restants couplée avec des messages indiquant qu'un autre internaute vient d'en acheter un[21]. Il s'agit alors de créer un sentiment d'urgence qui va pousser le consommateur à précipiter son achat de crainte de faire face à un épuisement du stock disponible (Mathur *et al.*, 2019). Dans de telles situations, il est fait appel au *système 1* de notre cerveau, celui du choix rapide, instinctif, guidé par des routines, émotionnel… et non plus au *système 2*, celui du choix rationnel[22] (Kahneman, 2011).

Qu'il s'agisse de « poussées ou de frictions », l'architecture des choix et du parcours même de l'internaute (et du consommateur en général) exerce une influence déterminante et appelle à ce titre une interrogation sur la responsabilité de la firme qui les met en place.

Le développement de l'I.A. pourrait rendre ces stratégies plus efficaces en permettant une meilleure compréhension du comportement du consommateur, après l'avoir très précisément rattaché à un segment donné à partir des caractéristiques observées et des caractéristiques déduites. En d'autres termes, l'I.A.

19. Luguri et Strahilevitz (2019) montrent, à partir d'un échantillon représentatif de 1 762 internautes américains, l'influence des *dark patterns* sur les choix individuels en matière de protection de la vie privée en ligne. L'utilisation de *mild dark patterns* fait passer le taux de choix d'un régime proposé de protection des données de 11 à 26 % (+228 %). Celle d'*aggressive dark patterns* de 11 à 42 % (soit 371 %). Ces pratiques conduisant un internaute à renoncer inutilement à la protection de ses données personnelles correspondent à la notion de *privacy zuckering*.
20. Ces pratiques peuvent également conduire à l'ajout involontaire d'un produit dans un panier dans le cas d'une place de marché ou encore au choix d'une assurance inutile dans le cadre d'une réservation d'un voyage en avion.
21. L'offreur joue ici sur un biais comportemental lié à l'aversion aux pertes.
22. Dans la même perspective, une *sludge* peut être favorable au consommateur en ce qu'elle lui laisse le temps de prendre une décision « à froid » et permet d'éviter qu'il ne s'engage de façon impulsive.

peut favoriser non seulement une personnalisation des prix, mais également une personnalisation des manipulations. En effet, comme le note le Stigler Center (2019, p. 238), l'utilisation des *sludges* va avoir des effets démultipliés au travers du recours à l'intelligence artificielle : « *Dark patterns are often used to direct users towards outcomes that involves greater data collection and processing. Additionally, the proliferation of data-driven computational methods allows firms to identify vulnerabilities of users and to target specific users with these vulnerabilities* »[23].

Par exemple, le recours à l'I.A. peut permettre de déterminer quel *stimulus* présenter à un consommateur et à quel moment le faire à partir d'une prédiction de plus en plus fine de ses caractéristiques et donc également de ses faiblesses déduites.

Luguri et Strahilevitz (2019) proposent une typologie détaillée des différents mécanismes pouvant se rattacher à la catégorie des *dark patterns*. Nous la reproduisons en partie dans le tableau présenté *infra* en lui ajoutant des éléments développés par Gray *et al.* (2018).

Catégorie de pratiques	Variante	Description
nagging **(ténacité)**	Une même option bien que déclinée est présentée à de multiples reprises sous des formes différentes	demandes répétées de faire un choix ou mise en forme d'un choix qui n'est pas définitif (non versus pas maintenant)
social proof	messages sur l'activité de tiers	fausses annonces que des tiers sont en train d'acheter ou de déposer des commentaires
	avis de contributeurs	faux avis
obstruction	modèle de la friction (*sludge*) visant à entraver une action ou un choix	
	roach motel (punaises de lit)	entrée facile dans une prestation/désabonnement ou renonciation bien plus complexe et chronophage
	obstacles à la comparaison des prix	le consommateur est empêché (au travers d'une interface empêchant un copier-coller des caractéristiques du produit sélectionné par exemple) de comparer le prix sur un site concurrent
	flou sur les devises	des options ou des prestations sont affichées dans des devises différentes ou le site repose sur l'utilisation d'une monnaie virtuelle (un *token*) qui entrave la comparaison des prix
	compte immortel	il est impossible de supprimer définitivement son compte

23. Gray *et al.* (2018) définissent les *dark patterns* comme « *instances where designers use their knowledge of human behavior* (e.g. *psychology) and the desires of end users to implement deceptive functionality that is not in the user's best interest* ».

actions en cachette (*sneaking*)	*sneak into basket* (un ajout discret dans le panier)	un item non demandé est rajouté par défaut à la commande
	coûts cachés	les coûts réels sont cachés ou révélés très retard dans le processus d'achat (ex. taxes, commission de change pour un achat à l'étranger ou encore coûts d'expédition prohibitifs)
	subscription cachée/continuité forcée	renouvellement tacite non anticipé ou non annoncé
	appâter et changer (*bait & switch*)	l'achat ne correspond pas à ce qui était initialement présenté
	désinscription piégée	la désinscription se double de l'acceptation de la réutilisation ou de la revente des données personnelles
manipulation des interfaces	dissimulation d'informations/ manipulation « esthétique »	les informations les plus importantes sont celles qui sont le moins mises en évidence sur l'interface utilisateur ou la présentation rend les options les moins intéressantes, les plus visibles ou attractives la position des options peut également varier dans le temps pour conduire le client à cliquer sur une option non désirée
	présélection	des options (défavorables au consommateur) sont cochées par défaut
	jeu sur les émotions	la présentation des pages et des options manipule les émotions de l'internaute (couleur, vocabulaire, illustrations…)
	fausse hiérarchie, pression sur l'achat	le comportement est manipulé pour conduire l'utilisateur à choisir l'option la plus coûteuse ou celle pour laquelle l'engagement est le plus long
	questions piégeuses	ambiguïté volontaire – l'utilisateur pense répondre à une question simple, alors que les répercussions de son choix sont plus complexes variante : des questions reposant sur des doubles négations ou sur un vocabulaire inutilement complexe et ambigu
	publicité déguisée	l'internaute est conduit à cliquer sur un lien qui n'apparaît pas de façon claire comme étant une publicité
	confirmation et culpabilité (*confirmshaming*)	la présentation de l'option est telle qu'un refus est signalé comme stupide
	attendrissement (*cuteness*)	un assistant numérique est présenté de façon à jouer sur les sentiments de l'internaute

action contrainte	spam sur les actions des tiers, pyramide sociale, exploitation du carnet d'adresses *(address book leeching*)	extraction manipulatrice d'informations liées à des contacts de l'internaute
	zuckérisation de la confidentialité (*privacy zuckering*)	les consommateurs sont conduits à rendre publiques des informations personnelles sans qu'ils en soient réellement conscients
	ludification (gamification)	le retour sur le site ou la sélection de certaines options du site sont présentés comme susceptibles de faire gagner quelque chose à l'utilisateur
	identification forcée	on présente l'identification sur le site comme indispensable
	orientation vers une mauvaise direction	l'utilisateur voit son attention captée vers un élément au détriment d'un second, plus déterminant
sentiment de rareté	message annonçant un faible nombre de produits disponibles	création d'un sentiment d'urgence sur l'acte d'achat (jeu sur l'aversion pour les pertes)
	message indiquant que la demande pour le produit est forte	le consommateur est prévenu que d'autres internautes (les plus personnalisés possibles) sont en train de faire des achats
sentiment d'urgence	compte à rebours	un indicateur mis en évidence sur la page montre que le temps de disponibilité de l'offre se réduit constamment
	fenêtre intempestive informant que l'offre n'est plus disponible pour très longtemps	l'option d'achat va disparaître dans quelques instants

L'existence de stratégies manipulatrices explique la différence entre les préférences des agents, par exemple en matière de protection de leurs données personnelles, et leur comportement effectif. Acquisti *et al.* (2013) ont montré que les agents peuvent accepter de payer pour protéger leurs données personnelles. De la même façon, leur comportement en ligne (adresses de messagerie multiples, spécialisation des différents réseaux sociaux…) atteste également de cette stratégie (Acquisiti *et al.*, 2020). Cependant, le cas des *dark patterns* décrits *supra* montre qu'il convient de distinguer la volonté des opportunités. Des biais psychologiques peuvent contrecarrer l'expression des préférences et ouvrir la possibilité de manipulations telles que celles que nous venons de détailler. Le tableau reproduit *infra* reprend les éléments présentés par Acquisti *et al.* (2020).

Biais psychologique	Description	Conséquence possible	Capacité d'action manipulatrice pour la firme
Asymétries d'information	Les utilisateurs ne peuvent anticiper l'usage qui sera fait de leurs données	Impossible de se couvrir contre un risque que l'on ne peut anticiper ou mesurer	La firme a intérêt à ne pas rendre ses pratiques transparentes
Rationalité limitée	L'utilisateur ne peut intégrer tous les paramètres de choix dans sa décision	Peu sont capables de comprendre les règles générales d'utilisation (et volontaires pour les lire)	Plus les règles sont complexes et techniques, plus elles seront obscures et moins l'utilisateur les évaluera
Biais de présentéisme	Surestimation des gains de court terme par rapport aux coûts de long terme	Une incitation modique pour l'entrée dans un contrat suffit à modifier l'arbitrage du consommateur même si les risques futurs sont élevés	Offrir une incitation tangible et immédiate pour pousser au partage des données
Biais d'évaluation des intangibles	Les paramètres intangibles sont difficiles à isoler, à quantifier et donc à prendre en compte	Les conséquences négatives d'une faible protection des données personnelles sont non seulement diffuses, mais de plus difficiles à rattacher à une décision précise (éloignée dans le temps)	L'opacité des clauses par défaut (relatives à l'utilisation des données) permet de réduire le risque réputationnel lié à la vente des données ou à leur utilisation stratégique (par exemple au travers de prix discriminateurs)
Préférences construites	On raisonne sur des heuristiques ne prenant pas en compte les coûts et bénéfices objectifs	Les individus ne modifient pas les règles par défaut qui leur sont proposées	Les règles par défaut sont plus avantageuses pour la plateforme que pour le client
L'illusion du contrôle	L'impression chez l'agent qu'il exerce un réel contrôle le pousse à des prises de risque excessives	L'illusion peut être d'autant plus forte que la granularité des options est fine	De multiples questions avec de multiples options jouent comme un *nudge* négatif
Panurgisme	Caler son comportement sur celui des tiers	Les décisions de partage d'informations sont liées à celles qui sont divulguées par les tiers	Pousser aux divulgations présentées comme une norme
Sous-adaptation	L'agent n'ajuste pas son comportement face à des circonstances évolutives	Même si la politique de protection des données personnelles se dégrade, les agents ne révisent pas leurs choix initiaux	Dégrader progressivement les conditions pour les utilisateurs
Incitations à divulguer	L'architecture « pousse » les individus à partager des informations personnelles	Risque de diffuser (ex. sur des réseaux sociaux) des informations auto-incriminatoires	Manipulations comportementales pour partager des contenus en ligne (« vos photos ont été vues par XX personnes »)

Certains des effets sont liés à des biais de comportement, d'autres sont créés et amplifiés par des pratiques qui jouent sur ces mêmes biais. Les erreurs d'estimation des risques n'ont pas besoin d'être soutenues par des manœuvres des sites concernés. Les conséquences d'une perte de confidentialité sont potentiellement très élevées, mais appréciées comme étant d'une faible probabilité et s'inscrivent de surcroît dans un futur lointain. Elles seront donc – comme dans d'autres domaines pouvant induire des risques systémiques – minorées dans la prise de décision (Kunreuther *et al.*, 1978).

Pour d'autres dimensions, la notion de *dark pattern* peut à nouveau être mobilisée. Acquisti *et al.* (2020) illustrent de quelle façon ces biais peuvent être instrumentalisés. Le consommateur va survaloriser un gain tangible et immédiat par rapport à un risque différé et intangible. De surcroît, comme nous l'avons vu pour les *bad sludges*, il n'est même pas acquis qu'il remplisse *ex post* les conditions d'activation de ces offres. La même sensibilité aux *bad sludges* ou aux *bad nudges* peut procéder de l'incertitude des agents quant à leurs propres préférences. La présentation des options peut à ce titre orienter les décisions des agents… y compris contre leurs intérêts. Comme cela a déjà été présenté dans le tableau *supra*, des dispositifs pourtant apparemment « transparents » peuvent aller intuitivement dans le sens inverse des intérêts des consommateurs (Acquisti *et al.*, 2013). En effet, l'illusion du contrôle – au travers de larges possibilités de personnalisation des choix – peut conduire les utilisateurs à accepter une divulgation excessive de leurs données (Brandimarte *et al.*, 2013). De la même façon, une dégradation progressive, mais continue, de la protection des données peut être perçue par les utilisateurs, mais ne pas faire l'objet d'une révision des choix initiaux : « *the human brain seems to interpret the persistence of a problem as evidence that the problem is intractable, and hence not worthy of further attention, so it dials down the emotional response* » (Acquisti *et al.*, 2020).

2.3 L'imposition de conditions contractuelles déséquilibrées

L'I.A. peut conduire à une segmentation très fine des clients permettant de proposer des prix quasiment personnalisés. Le problème de ces derniers est que, placé au niveau de la propension maximale à payer du consommateur, un prix parfaitement discriminant permet de confisquer la totalité du surplus du consommateur. En termes économiques, il n'y a pas de dommages en matière d'efficacité, mais un transfert indu de bien-être par rapport à la répartition qui prévaudrait en concurrence parfaite.

Il convient en outre de relever que la discrimination entre les clients peut également prévaloir en cas de prix uniforme. S'il est possible de déterminer les

besoins, mais aussi le niveau d'expertise technique du client, il est envisageable de lui proposer un produit aux caractéristiques moins attractives ou aux performances dégradées. Le vendeur peut profiter de l'avantage informationnel qu'il sait posséder sur son client et de la flexibilité de la production que permettront de façon croissante les modèles d'industrie 4.0. Il s'agit des pratiques dites de *versioning*. Le consommateur moins expert peut se voir proposer des offres *in fine* plus coûteuses que la valeur intrinsèque de l'offre personnalisée qui lui est faite (Marty, 2019).

La notion de *dark pattern augmenté* peut donc recouvrir plusieurs modalités. Celle que nous venons de voir correspond à une *manipulation by transaction costs*. Le consommateur voit son bien-être dégradé au travers d'abus d'exploitation prenant la forme d'un prix personnalisé *confiscateur* (de son surplus), d'une offre dont le rapport qualité-prix est dégradé ou prenant la forme de barrières à la sortie. La notion de *dark pattern* recouvre des manipulations du comportement des consommateurs, basées sur l'identification fine de leurs caractéristiques et plus précisément de leurs faiblesses[24].

3 INTELLIGENCE ARTIFICIELLE ET DOMMAGES AU PROCESSUS DE CONCURRENCE

Le recours à l'I.A. peut conduire à donner à une entreprise un avantage déterminant par rapport à ses concurrentes.

Même si cet avantage repose sur les mérites, et ne devrait pas à ce titre être sanctionné sur la base des règles de concurrence, il n'en demeure pas moins qu'il fait obstacle à une concurrence à égalité des armes, et s'avère susceptible de faire basculer le marché dans une situation de dominance écrasante et éventuellement pérenne. En effet, cet avantage peut faire qu'un nouvel entrant ne saurait être d'emblée aussi efficace que l'entreprise dominante. Le marché ne serait plus dès lors contestable au sens économique du terme. L'avantage en matière d'I.A. serait alors susceptible de se muer en barrière à l'entrée infranchissable. Le dommage ne serait peut-être pas un dommage en matière d'efficacité (tant pour l'économie que pour le consommateur) ni en matière d'innovation, mais un dommage pour le processus de concurrence lui-même (3.1). Dans le même temps, le recours à l'I.A. dans une structure de marché oligopolistique peut conduire à l'émergence plus rapide d'équilibres collusifs,

24. Il convient de noter que les seconds types de *sludges* sont beaucoup plus difficiles à identifier *ex post* dans la mesure où ils sont personnalisés en fonction de chaque consommateur et que ces derniers peuvent avoir des difficultés à identifier, dans le cadre de leur expérience utilisateur, quelles sont les pratiques en causes et dans quelle mesure celles-ci sont personnalisées (Stigler Center, 2019, p. 240).

de surcroît plus stables, dans la mesure où les firmes en concurrence accèdent à une meilleure compréhension du marché et sont en mesure de prédire de plus en plus finement leurs réactions réciproques (3.2).

3.1 Le verrouillage d'une position dominante

Dans un tel cas de figure, le recours à l'I.A. vise moins à exploiter un avantage vis-à-vis du consommateur qu'à acquérir, consolider ou étendre une position dominante au détriment de ses concurrents actuels ou de compétiteurs potentiels, que cela soit sur un même marché pertinent ou sur des marchés connexes (de biens et services complémentaires, par exemple). Les pratiques en cause peuvent être – au moins à court terme – favorables au consommateur, en matière de bien-être. Elles peuvent néanmoins porter préjudice à la pérennité d'une concurrence libre, non faussée et par les mérites.

Les avantages que possèdent les firmes dominantes par rapport à leurs concurrents actuels ou futurs, et vis-à-vis de leurs partenaires commerciaux (en d'autres termes, des firmes qui jouent le rôle de *complémenteurs* dans leurs écosystèmes respectifs), leur permettent de maîtriser leur environnement et donc de limiter la situation d'incertitude radicale qui théoriquement caractérise la concurrence. Une foisonnante littérature de sciences économiques et de sciences de gestion se développe sur la notion de zone mortifère (*kill zone*) et sur celle d'acquisitions tueuses (*killer mergers*) ou encore consolidantes. La détention de données massives, continûment renouvelées et diversifiées, et la conception d'algorithmes d'I.A., couplée à leurs capacités de calcul et donc de traitement de l'information, peuvent permettre aux entreprises dominantes (les firmes pivots de chaque écosystème) d'identifier très en amont les menaces concurrentielles, les technologies prometteuses ou les développements potentiellement disruptifs. Les firmes dominantes peuvent donc éliminer ou cloner le service ou encore racheter l'entreprise concernée bien avant qu'elle accède effectivement au marché[25].

Le consommateur peut ne subir aucun préjudice à court terme. Une éventuelle innovation ne sera pas éliminée, elle pourra être intégrée à l'offre de la firme dominante et sera peut-être ainsi même plus efficace et attractive. Cependant, la capacité de détection avancée pérennise la dominance en jouant comme une barrière à l'entrée sur le marché. Le verrouillage (anticoncurrentiel) procède d'une capacité à détecter les signaux faibles sur le marché. Différents

25. L'opacité des algorithmes de la plateforme dominante rend difficile la démonstration de stratégies d'éviction de sa part (que cela soit par altération des classements dans les résultats de recherche ou par réplication des caractéristiques de l'offre du concurrent visé). Pour une synthèse, voir Patterson (2018).

outils algorithmiques facilitent ces pratiques dites de *nowcasting*. L'analyse de sentiment en constitue l'une des principales.

3.2 L'émergence et la consolidation de collusions algorithmiques

Les pratiques décrites *supra* correspondent à des pratiques unilatérales, c'est-à-dire mises en œuvre par une entreprise dominante de façon indépendante de ses concurrentes. Les algorithmes en général et l'I.A. en particulier peuvent également faciliter le développement si ce n'est l'émergence d'équilibres collusifs. À nouveau, une littérature foisonnante s'est développée sur la question de la capacité de l'I.A. à favoriser et à stabiliser des équilibres de collusion tacite (Calvano *et al.*, 2019). Il s'agit de situations dans lesquelles des algorithmes, capables « d'apprentissage machine non supervisé »[26] en *comprenant* le fonctionnement du marché et les réactions des concurrents, convergent spontanément vers un équilibre coopératif (c'est-à-dire de paix armée) dans la mesure où c'est celui qui maximise les profits de chacun sur le long terme.

Des joueurs *humains* arriveraient potentiellement au même résultat, mais dans le cadre d'hypothèses bien plus restrictives (en matière de nombre de participants, de complexité de l'environnement…) et d'une période bien plus longue. En outre, un tel équilibre serait bien plus stable avec des I.A. qu'avec des humains dans la mesure où les premières pourraient présenter moins de biais cognitifs les conduisant à mal interpréter les stratégies des leurs *concurrents* ou à surréagir en cas d'écart observé vis-à-vis de l'équilibre de collusion tacite. Qui plus est, la démonstration d'une intention anticoncurrentielle serait bien plus difficile à faire, ce qui serait susceptible de réduire significativement la probabilité d'être sanctionné par les règles de concurrence (Marty, 2017).

Pour finir, malgré les gains d'efficacité qu'apportera l'I.A., la littérature académique insiste sur des risques associés[27]. Ainsi, la course vers l'I.A. ne serait pas exclusivement la solution à la recherche d'efficacité dans nos économies guidées par les algorithmes, elle pourrait également être un problème. Notre quatrième section montre que si l'I.A. peut être un problème, elle

26. L'« apprentissage machine non supervisé correspond à un algorithme capable de s'entraîner de façon autonome sur des données non étiquetées, c'est-à-dire des données brutes et non préalablement classées ».

27. La balance entre risques concurrentiels et gains d'efficience est particulièrement difficile à établir dans le cadre d'une concurrence entre écosystèmes numériques caractérisée par de forts investissements en innovation. En conséquence, les effets de l'activation des règles de concurrence, voire de dispositifs de régulation spécifique, sont particulièrement élevés. Pour une discussion, voir Petit (2020).

peut également être la solution pour maîtriser ces risques éventuels. Il s'agit cependant de considérer ces pistes dans leurs dimensions pratiques, légales et éthiques.

4 PISTES POUR UNE RÉGULATION DES ALGORITHMES PAR LES ALGORITHMES

Les outils fournis par l'I.A. peuvent contribuer à corriger leurs éventuels effets anticoncurrentiels ou préjudiciables aux consommateurs. Ils peuvent être mis en œuvre par le régulateur de la concurrence (4.1) ou par les consommateurs eux-mêmes (4.2).

4.1 Le recours à l'intelligence artificielle par les autorités de supervision des marchés pour prévenir des stratégies de manipulation

Les ressources apportées par l'I.A. peuvent être utilisées *ex ante* dans le cadre de procédures de validation des algorithmes (dans une logique d'exigence d'une conformité par conception) ou dans le cadre d'enquêtes sectorielles. Il s'agit alors de faire jouer les algorithmes sur la base des données de marché transmises par les entreprises pour observer d'éventuels biais. En matière de collusion par algorithmes, ces contrôles pourraient fonctionner comme des *stress tests*. Il s'agirait, au travers d'incubateurs de collusion algorithmique, de voir sous quelles conditions et avec quelle rapidité les algorithmes de firmes concurrentes pourraient converger vers de tels équilibres. Il appartiendrait au régulateur et aux firmes de définir les conditions (en matière de vitesse ou de fréquence de changement de prix par exemple) pour limiter les risques concurrentiels.

Ensuite, les algorithmes peuvent être utilisés *ex post* dans une surveillance des marchés. Comme cela est déjà le cas dans le cadre de la régulation des marchés financiers pour les transactions à haute fréquence, il est possible d'analyser les *patterns* de marché pour y détecter des stratégies qui n'auraient pas de sens économiquement en dehors d'une pratique abusive. Il appartient alors à l'entreprise concernée d'apporter la preuve que ses décisions ne participaient pas d'une telle stratégie (logique de *comply or explain* – être conforme ou s'expliquer).

Relevons que le droit de la concurrence n'est pas le seul outil juridique qui peut permettre de faire rendre des comptes aux algorithmes. Par exemple, s'agissant des offres aux consommateurs finaux (en matière de prix ou de

qualité des produits proposés), des pratiques de nature discriminatoire pourraient être soumises aux règles de protection des consommateurs comme autant de pratiques déloyales ou trompeuses. Il convient cependant de noter que, plus fine sera l'analyse produite par l'algorithme des besoins et des caractéristiques du client, plus il sera difficile de détecter une manipulation et de détecter un dommage.

Le Centre Stigler (2019, p. 254) propose dans son rapport quelques pistes qui peuvent être discutées. Un *dark pattern* dissimulé, qui accroît les coûts de sortie des consommateurs ou qui est susceptible d'exploiter les faiblesses des consommateurs les plus vulnérables doit faire l'objet d'une présomption de dommages au consommateur[28]. En d'autres termes, l'intention manipulatrice serait présumée si la conception même de l'algorithme apparaissait comme opacifiant volontairement l'intention du développeur et ses effets. Les pratiques visées par le Centre Stigler (2019) méritent d'être discutées, notamment en regard du développement de l'I.A.

4.2 Le recours à l'intelligence artificielle par les consommateurs ou leurs associations

Les consommateurs eux-mêmes peuvent utiliser des outils algorithmiques pour détecter ou contrecarrer les éventuelles manipulations des firmes. Cela peut passer par des dispositifs de surveillance « distribués » mis en place par des institutions à but non lucratif ou par l'utilisation de contre-algorithmes par les consommateurs eux-mêmes. Ces dispositifs peuvent permettre de se jouer des stratégies algorithmiques mises en place par les firmes, voire de tromper ces dernières, par l'émission de faux signaux.

Il convient de relever que tous les segments de consommateurs ne sont pas en position équivalente vis-à-vis des manipulations algorithmiques et des réponses qu'ils sont susceptibles de leur apporter. Premièrement, la probabilité d'être manipulé, même par un *mild dark pattern*, dépend du niveau de connaissances du consommateur, c'est-à-dire souvent son niveau d'éducation. Il en va ainsi également de la capacité de percevoir la manipulation elle-même[29]. Deuxièmement, l'accès même à de telles contre-mesures est relié aux mêmes

28. Ils firent l'objet d'une proposition d'interdiction par le Sénat américain dans le cadre du *Deceptive Experiences to Online Users Reduction Act* (*Detour Act* – S.1084 – 116th Congress, 2019-2020).

29. La transparence quant à l'existence de biais liés aux paramètres par défaut semble de surcroît avoir des effets ambigus. Certes, les consommateurs révisent leurs perceptions quant au comportement éthique de l'entreprise qui les met en œuvre, mais ils ne modifient pas pour autant significativement leurs choix. Une intervention plus active est encore nécessaire. Il faut leur proposer des paramètres alternatifs pour voir leurs choix changer (Steffel *et al.*, 2016).

caractéristiques. En d'autres termes, les éventuelles manipulations algorithmiques ne toucheront pas de la même façon les différentes catégories de consommateurs, ce qui pose à la fois des questions éthiques (Marty et Warin, 2019) et des questions distributives.

Le recours à l'I.A. par les plateformes peut être mal perçu par les consommateurs. La personnalisation des prix et des offres peut être analysée comme une discrimination visant à s'approprier la totalité du surplus que dégage de la transaction le consommateur. Le « guidage » des comportements de choix, surtout s'il est perçu comme relevant d'un *dark pattern* peut être interprété comme une pratique manipulatrice déloyale et préjudiciable (*deceptive practices*). Ce faisant, un phénomène de rejet de l'entreprise par les consommateurs peut être observé (Stigler Center, 2019). Il peut s'ensuivre des stratégies mesurées, mais toujours dommageables de la part des firmes. L'utilisation de *dark nudges* moins manifestes (*mild dark patterns*) peut permettre d'obtenir les effets recherchés chez les consommateurs les plus naïfs sans pour autant s'aliéner les plus avertis qui pourraient réagir négativement envers l'opérateur en cas de recours à des *aggressive dark patterns*. Paradoxalement, ce sont les consommateurs les moins exposés aux *dark patterns* qui seraient susceptibles d'exercer une menace crédible contre les firmes qui les utiliseraient. Les consommateurs les plus exposés sont d'autant moins susceptibles de pouvoir y répondre qu'ils sont moins bien placés pour les identifier. De surcroît, le ciblage sur les vulnérabilités de certains segments précis de consommateurs accroît d'autant l'efficacité des manipulations, au plus grand détriment des consommateurs les plus vulnérables.

Il s'agit donc de contraindre les firmes à une redevabilité (*accountability*) des algorithmes malgré leur opacité (logique de boîte noire) au travers d'une régulation par coup de projecteur pouvant soulever des enjeux de réputation. Ces outils peuvent contribuer à responsabiliser les entreprises et les encourager à développer une I.A. responsable. Le développement d'outils algorithmiques au profit des consommateurs peut leur permettre d'exercer un contre-pouvoir concurrentiel compensateur au travers d'un *combat algorithmique* (Gal, 2019).

Cependant, trois enjeux restent à considérer. Le premier tient à la capacité de rendre compte des résultats des algorithmes dans un modèle où l'opacification n'est pas seulement volontaire, mais consubstantielle à la technologie (par exemple, en cas d'apprentissage profond). Cette question fait écho à la notion d'intelligence artificielle explicable[30] (XAI). Le deuxième tient à la dissymétrie de la performance entre algorithmes et contre-algorithmes au regard des capacités accumulées par les opérateurs dominants en matière d'I.A. et des capacités de traitement de l'information, surtout dans une perspective de développement

30. Voir notamment sur ce point, Smith (2020).

de l'informatique quantique. Le troisième tient aux dimensions éthiques de l'utilisation des algorithmes en matière de détection et de sanction des pratiques d'une part et d'utilisation de contre-algorithmes d'autre part. Sur le premier aspect, une des questions pertinentes tient à la prise en compte d'un modèle sous-jacent de ce que devrait être une situation normale de marché, de façon à pouvoir caractériser une stratégie comme suspecte. Sur le second aspect, une des dimensions à considérer tient à l'inégal accès des consommateurs aux contre-mesures et au renforcement potentiel des inégalités qui peut en résulter.

Cette dernière dimension peut mener à un questionnement plus ample quant aux différences qui peuvent s'accroître d'un consommateur à l'autre du fait d'un recours croissant aux décisions algorithmiques tant de la part des firmes que des consommateurs eux-mêmes. L'économie numérique et les stratégies de discrimination qu'elle facilite peuvent accroître les inégalités entre consommateurs captifs et consommateurs optant pour des stratégies de multi-hébergement, entre consommateurs « informés » et consommateurs « naïfs », et enfin entre consommateurs aptes à mettre en œuvre des dispositifs techniques leur permettant d'exercer non pas un *pouvoir de marché compensateur,* mais une *capacité technique de correction* des stratégies algorithmiques *des firmes* et les autres.

Il convient enfin, dans le cas des interfaces et des algorithmes manipulatoires par conception (*dark patterns*), de souligner le fait que le consommateur peut être conduit à attribuer la faute à son propre comportement et non à une stratégie mise en œuvre par la firme… et ce dernier peut même parvenir à identifier *ex post* un « dommage » et à faire le lien entre ses choix passés et celui-ci. Or, comme nous le montrons dans notre cinquième section, les dommages causés par les algorithmes en matière de protection du consommateur et de la concurrence peuvent être particulièrement importants, en ce qu'ils mettent en jeu des « décisions hautement conséquentes – *high stake decisions*) ».

5 QUEL ENCADREMENT POUR LES DÉCISIONS HAUTEMENT CONSÉQUENTES ?

Les algorithmes ont ici un impact sans précédent sur un pan de la vie des entreprises et des consommateurs. La littérature en informatique accorde une attention particulière aux questions des décisions hautement conséquentes (Buolamwini, 2018 ; Citron et Pasquale, 2014 ; Kleinberg *et al.*, 2018 ; O'Neil, 2016 ; Parkes, Vohra et participants, 2019). Nous nous proposons de les envisager au regard des règles de concurrence ou des règles applicables à la supervision des marchés financiers (5.1), puis par rapport aux différentes parties prenantes de l'entreprise (5.2).

5.1 Droit de la concurrence et droit des marchés financiers et décisions hautement conséquentes

Dans le domaine du droit de la concurrence, les marchés impliquent souvent des décisions très lourdes de conséquences qui sont principalement prises par des décideurs humains. Le mot-clé est ici « décisions hautement conséquentes ». La caractéristique est le rythme de ces décisions. Le droit de la concurrence s'est construit autour de la question des décisions hautement conséquentes et du rythme des décisions prises sur les marchés par les humains. L'objectif est de garantir une concurrence à égalité des armes dans une configuration concurrentielle dans laquelle les décisions des firmes sont fondées sur l'analyse des données et sont de façon croissante automatisées sans intervention humaine dans le circuit décisionnel.

La disponibilité croissante des données relatives à ces décisions offre une opportunité sans précédent de développer des modèles d'apprentissage automatique (ci-après AA, voir Marcellis-Warin, Munoz et Warin, 2020). Ces modèles d'AA peuvent être utilisés de deux manières. Premièrement, l'I.A. peut être utilisée dans une perspective d'intelligence augmentée. Ces modèles d'AA peuvent aider les décideurs humains à prendre de meilleures décisions. Deuxièmement, l'I.A. peut être utilisée comme une intelligence automatisée : ces modèles d'AA prennent des décisions qui sont traditionnellement exécutées par des humains (Citron et Pasquale, 2014 ; O'Neil, 2016). Ces algorithmes peuvent jouer la dynamique du marché à très grande vitesse avec une succession de décisions à impacts certes faibles, mais « conséquents », qui peuvent à terme avoir des effets très significatifs comme le montre le cas des manipulations algorithmiques sur les marchés à haute fréquence.

Toutefois, l'applicabilité de l'AA aux paramètres susmentionnés est limitée par certains défis fondamentaux. Premièrement, les paramètres mentionnés *supra* exigent la conception de modèles qui tiennent compte de l'équité et de la possibilité d'interprétation. Cependant, la plupart des modèles de blanchiment d'argent existants sont principalement optimisés pour la précision des prévisions et ne sont pas intrinsèquement équitables ou interprétables. Deuxièmement, les données disponibles dans ces contextes sont souvent soumises à divers biais de sélection.

Ces cadres sont exposés au problème des contre-factuels manquants, c'est-à-dire que les données ne saisissent que les résultats des décisions prises par les décideurs humains et non les contre-factuels. Cependant, ces outils, déjà utilisés en matière de supervision des activités financières, peuvent être utilisés dans le domaine de la concurrence dès lors que l'enjeu est de contrôler la politique de prix de certains acteurs du marché avec des prix personnalisés et dynamiques, c'est-à-dire foisonnants et apparemment (mais en apparence seulement) erratiques.

5.2 Une supervision opérée en interne pour le compte de la firme elle-même et de ses parties prenantes

Nous avons vu que des exigences de transparence et des outils de révélation (dans une logique d'*explainable artificial intelligence* ou XAI) pouvaient s'avérer efficaces pour responsabiliser les plateformes. Nous avons également examiné le rôle positif d'une régulation favorisant les « coups de projecteurs », c'est-à-dire pouvant nuire à la réputation de l'entreprise, faire baisser sa valeur actionnariale et réduire ses parts de marché.

Nous allons maintenant examiner l'impact d'un « paternalisme libéral » sur la responsabilisation des plateformes. Il s'agit d'examiner les incitations économiques visant à engager les entreprises dans des modèles d'affaires bénéfiques pour l'ensemble des parties prenantes en alignant les intérêts de chacun. En effet, une plateforme tournée vers les parties prenantes et vers la firme elle-même sera à même de gagner des parts de marché et d'accroître sa valeur actionnariale, tout en préservant les intérêts des générations futures.

Dans la mesure où les choix et les comportements des acteurs peuvent être altérés par le fonctionnement des algorithmes et l'architecture des choix qui se présentent à eux (tant au travers de poussées que de frictions stratégiques), leur conception même peut faire l'objet d'évaluations. Sunstein (2019) considère, par exemple, que certaines *sludges* peuvent avoir des effets opposés, tenant à la fois à des outils de manipulation et à des dispositifs d'auto-sélection ou de protection. La proposition serait alors celle d'un *sludge audit*, basé sur une analyse coût-avantage (« *to weight their benefits against their costs and a careful assessment of their distributional effects* »).

Cependant, d'autres dispositifs peuvent être mis en place, allant au-delà d'une logique de type coût-avantage (Sunstein, 2018) et reposant sur une notion de confiance vis-à-vis de l'algorithme. Comment, par exemple, s'assurer de l'absence de stratégies visant à altérer les déterminants des décisions des agents au travers d'entraves ou au travers de manipulations algorithmiques ? Comment apporter des garanties quant à la collecte, au traitement ou à l'utilisation des données ? Par exemple, comment garantir que les dispositifs mis en œuvre par les algorithmes eux-mêmes contribuent à un alignement des choix avec les préférences *ex ante* des internautes (si celles-ci préexistent au choix d'ailleurs…) dès lors que la transparence ou la granularité des choix peuvent jouer dans certaines situations à la fois comme des *bad nudges* ou des *bad sludges* (illusion du contrôle) et que, symétriquement, les agents peuvent être enclins à accepter sans discussion des paramètres par défaut qui peuvent leur être défavorables (Acquisti, 2009 ; Acquisti *et al.*, 2017) ?

La responsabilité des firmes peut être considérée. La conception même de l'architecture des choix qui s'ouvrent aux utilisateurs peut être porteuse d'enjeux

éthiques. L'interface même peut formater et canaliser le comportement, et altérer, sinon construire, les choix des utilisateurs. Pour Fogg (2002), la conception de l'algorithme peut porter sept types de stratégies de persuasion (réduction, canalisation, adaptation fine [*tayloring*], suggestion, autocensure, surveillance et conditionnement). Comme nous l'avons vu tant pour les *sludges* que pour les *nudges*, ces architectures persuasives peuvent jouer dans un sens soit favorable à l'utilisateur, soit défavorable (Berdichevsky and Neuenschwander, 1999). Ce résultat peut dépendre de deux phénomènes. Le premier peut tenir à la volonté de manipuler le consommateur (Nodder, 2013). Le second peut correspondre à une absence de considération et de prise en compte des enjeux eux-mêmes. Il s'agit d'une situation d'*anti-pattern* dans laquelle le dommage causé à l'utilisateur est involontaire, explicable par l'absence de soin dans le codage (Kœnig, 1995). De surcroît, selon l'internaute lui-même (ses compétences, son attention, son expérience…), une même architecture peut être manipulatrice ou acceptable.

La solution passe-t-elle par des dispositifs techniques, tels des *privacy enhancing technologies* (PET – techniques permettant d'améliorer la confidentialité des utilisateurs et la protection de leurs données personnelles), passant par exemple par le cryptage des données permettant de concilier respect de la vie privée et gains collectifs liés au traitement des données ou par des dispositifs de certification des algorithmes ?

Un des dispositifs concrets à disposition a été développé par l'AI Transparency Institute. C'est celui-ci que nous présentons à titre d'illustration. L'objectif de ses concepteurs est de créer des indices de confiance à la disposition des entreprises. Ces indices ont le potentiel d'informer les entreprises proposant des services ou des biens via une plateforme digitale du niveau de maturité de leur organisation en fonction d'un ensemble de critères prédéfinis. Ces critères concernent les aspects de sécurité, de gestion des données, de transparence, du caractère explicable des algorithmes, les valeurs éthiques, le respect de normes juridiques et l'État de droit. Ils tiennent compte de l'intérêt de toutes les parties prenantes, les employés, les clients et la communauté au sens large (y compris les investisseurs).

Ces indices de confiance peuvent faciliter l'obtention d'une certification ou d'un label de qualité. Ils se basent sur un modèle mathématique et sur une liste de questions pondérées. Il en ressort des indices de confiance et un ensemble de chartes graphiques thématiques. Les restitutions graphiques présentées *infra* donnent des exemples reposant sur les versions en ligne.

Le premier d'entre eux porte sur les données utilisées. Il traduit la qualité du contrôle mis en œuvre par l'entreprise considérée en matière d'origine des données (données collectées après consentement de l'utilisateur, observées à partir de son comportement en ligne, inférées, observées dans l'écosystème,

achetées auprès de courtiers en données, etc.), de vérification de leur qualité (notamment en termes de biais), de leur usage, de leur partage avec des tiers ou encore de leur conformité avec des règles de protections spécifiques comme le RGPD. Les graphes scalaires sont obtenus à partir des scores obtenus sur chacun des points (qui regroupent plusieurs questions qui font l'objet d'une note pondérée) et permettent de comparer la firme à une référence choisie comme *benchmark* ou à la moyenne des firmes participant au programme.

The Data – Index: 66.70

Strength: Acceptable

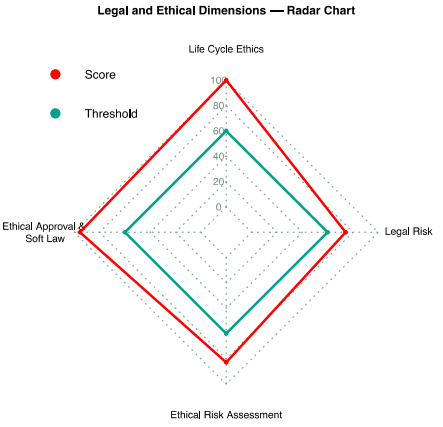
Risk Criteria	Score	Required	Strength
Data Sources	84.67	60.0	Amazing
Data Quality	54.94	60.0	Acceptable
Data Sharing	68.67	60.0	Good
Principles relating to the Collection, Storage, Processing of Data (GDPR etc.)	48.34	60.0	Acceptable
Data Usages	76.88	60.0	Good

Les scores peuvent se décliner sur de nombreuses dimensions comme le montrent les graphes présentés *infra* qui traduisent la conformité aux engagements éthiques ou aux exigences légales, la sécurité des systèmes IT ou encore l'entraînement et le contrôle des algorithmes utilisés.

Ces indices de confiance peuvent avoir un impact sur le comportement de l'entreprise et sur celui du consommateur. En effet, sur cette base, l'entreprise a la possibilité d'identifier ses lacunes, et d'investir ses ressources dans des enjeux essentiels tant pour sa valeur actionnariale que pour le bien commun. Elle pourra gagner en efficacité en améliorant ses processus internes et démontrer sa gestion responsable envers l'ensemble des parties prenantes. En indiquant les scores obtenus dans son rapport annuel et sur son site Internet, l'entreprise accroît la confiance en sa gouvernance d'entreprise.

Legal & Ethical – Index: 88.11

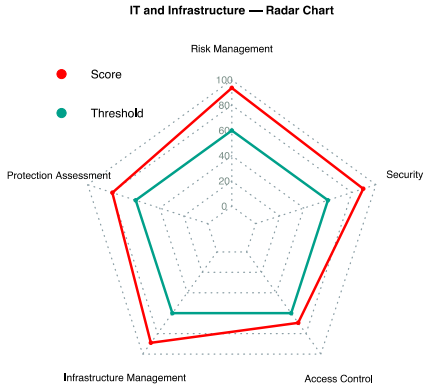
Strength: Amazing



Risk Criteria	Score	Required	Strength
Life Cycle Ethics	56.38	60.0	Acceptable
Ethics Approval & Soft Law	63.97	60.0	Good
Ethical Risk Assessment	77.67	60.0	Good
Legal Risks	89.24	60.0	Amazing

IT & Infrastructure – Index: 84.20

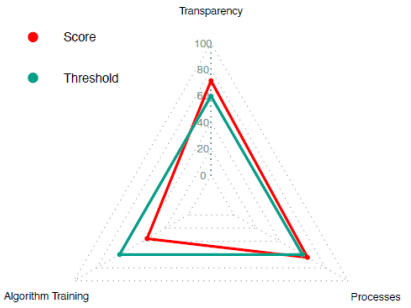
Strength: Amazing



Risk Criteria	Score	Required	Strength
Risk Management	93.67	60.0	Amazing
Protection Assessment	79.44	60.0	Good
Infrastructure Management	88.97	60.0	Amazing
Access Control	69.43	60.0	Good
Security	89.47	60.0	Amazing

Data Management – Index: 57.33

Strength: Acceptable

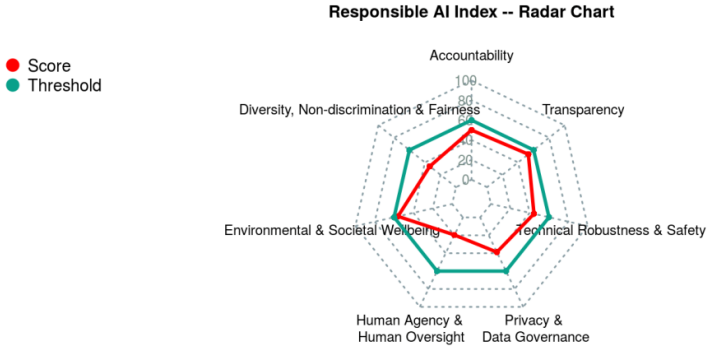


Risk Criteria	Score	Required	Strength
Transparency	71.67	60.0	Good
Algorithm Training	35.93	60.0	Bad
Processes	64.38	60.0	Good

Analysis & Visualization

Responsible AI Index: 42.26

Strength: Acceptable



Risk Criteria	Score	Required	Strength
Accountability	50.00	60.00	Acceptable
Diversity, Non-discrimination & Fairness	33.75	60.00	Bad
Environmental & Societal Wellbeing	56.00	60.00	Acceptable
Human Agency & Human Oversight	19.61	60.00	Poor
Privacy & Data Governance	38.78	60.00	Bad
Technical Robustness & Safety	44.49	60.00	Acceptable
Transparency	53.19	60.00	Acceptable

Ces indices de confiance incitent les entreprises à un comportement diligent puisqu'ils pourront constituer un élément de la notation financière de leur gouvernance (par exemple, par le biais de la Fondation Ethos). Ces indices seront accessibles aux actionnaires et ont le potentiel pour avoir une influence significative sur la valeur actionnariale des entreprises, en récompensant les comportements transparents et diligents. Ils montreront comment les entreprises se positionnent vis-à-vis de leurs concurrents. Ils inciteront à l'adoption d'un comportement similaire par les acteurs économiques d'un même marché et à une amélioration de leur note, donc de leur comportement responsable et transparent envers les consommateurs et la communauté. En rendant visible ce que les acteurs économiques réalisent, cela incite les entreprises à adopter un comportement socialement responsable.

Réciproquement, le consommateur a la capacité d'obtenir une information importante concernant le degré de confiance qu'il peut consentir à un acteur économique, comme une plateforme en ligne, grâce à la publication de ces indices sur son site Internet ou dans son rapport annuel. Ces indices augmenteront le pouvoir de contrôle des consommateurs.

S'ils apportent une valeur indéniable aux consommateurs et entreprises, ils offrent également une transparence accrue aux actionnaires concernant la gouvernance d'entreprise. Ils rendent en effet visibles des paramètres difficilement accessibles. Certes, ils reposent sur un audit *ex-ante* volontaire des organisations, mais pourront également être utilisés par des autorités administratives indépendantes dans le domaine du droit de la concurrence à des fins d'évaluation et d'audit *ex post*.

Il reste enfin à déterminer si cette solution d'un indice de confiance est suffisante pour donner pleinement confiance aux consommateurs dans les plateformes en ligne et les services et produits intégrant de l'intelligence artificielle. Soutenue par le régulateur et les fonds d'investissement, cette démarche d'autorégulation paraît prometteuse, d'autant que la prise de conscience des entreprises dans le domaine de la responsabilité sociale et environnementale augmente (Townsend, 2020).

Une intervention du législateur dans le domaine du droit de la concurrence et du droit des obligations sera vraisemblablement indispensable pour modifier le comportement des plateformes en ligne de manière durable. Les enjeux sont immenses (Parkes, Vohra et participants, 2019). Il n'est pas exclu qu'une obligation générale pour les entreprises de démontrer un comportement diligent envers le climat, les clients, les investisseurs et la communauté au sens large doive être imposée. Les plateformes en ligne contribuent aux enjeux du développement durable et devraient intégrer les préoccupations sociales, environnementales et économiques dans leurs activités et dans leurs interactions avec leurs parties prenantes sur une base volontaire. La norme ISO 26000 crée en ce sens un standard de référence.

En France, les articles L. 225-102-1 et suivants du Code du commerce obligent les entreprises à publier une déclaration de performance qui remplace le rapport de responsabilité sociétale. Il s'agit d'un outil de pilotage de la stratégie de l'entreprise, qui peut être complété par un indice pour quantifier la confiance dans les algorithmes d'intelligence artificielle. Ces articles sont complétés par la loi n° 2017-399 du 27 mars 2017 relative au devoir de vigilance des sociétés mères. Cette loi a pour ambition de veiller au respect des droits humains par les multinationales. Un plan de vigilance doit être publié pour prévenir les risques en matière d'environnement, de droits humains, de corruption sur leurs propres activités, mais aussi celles de leurs filiales, sous-traitants et fournisseurs, en France comme à l'étranger. Ces initiatives sont louables, mais limitées au territoire national.

Les indices de confiance ont une portée mondiale. Ils ont vocation à s'appliquer en dehors du territoire des États-nations et ont le potentiel pour assister les acteurs transnationaux dans une transformation numérique bénéfique et durable dans l'intérêt de la société. Ils pourront vraisemblablement jouer un rôle clef dans la diffusion d'une culture responsable pour les plateformes en ligne. Utilisés par les régulateurs, ces outils sont en mesure de favoriser la coopération internationale indispensable afin de renforcer les normes internationales là où cela est nécessaire, et de veiller à leur application uniforme afin d'assurer une protection contre les externalités négatives transfrontalières, régionales et mondiales qui affectent la confiance dans l'économie numérique. Une obligation générale et mondiale de publier un rapport annuel présentant

les mesures effectives prises par l'entreprise pour se comporter en acteur responsable apparaît comme une étape indispensable. En l'absence de sanctions multilatérales, la pression des marchés financiers apparaît un levier intéressant à examiner grâce à des indices de confiance et de responsabilité numérique.

CONCLUSION

Tant en matière de concurrence que de protection du consommateur, le développement des risques de manipulation algorithmique peut faire craindre la survenance de dommages qui peuvent être détectables avec une très faible probabilité et auxquels il serait particulièrement difficile de remédier au moyen de sanctions traditionnelles (injonctions comportementales et amendes). Les exemples, que nous avons présentés en introduction, des procédures négociées avec Apple et Zoom aux États-Unis montrent que les mesures prises peuvent ne pas réparer les éventuels dommages causés à la concurrence et aux consommateurs.

Dans ce contexte, l'intervention régulatoire tend à s'inscrire dans une temporalité plus large avec la mise en œuvre de mesures *ex ante*. Il en est ainsi de la *Digital Market Unit* dont la création a été annoncée en novembre 2020 au Royaume-Uni dans la continuité des préconisations du Rapport Furman (2019) et de l'enquête sectorielle de la *Competition and Markets Authority* dont les conclusions ont été rendues publiques en juillet 2020[31]. Cette même logique de supervision est présente dans les propositions du *Digital Services Act* et du *Digital Markets Act*, rendues publiques par la Commission européenne le 15 décembre 2020. Cependant, cette complémentarité entre interventions *ex post* et encadrement *ex ante* peut être renforcée par des mesures visant à prévenir les dommages algorithmiques (manipulation des consommateurs, création ou aggravation de biais…) tenant au contrôle des données et des algorithmes par les parties prenantes elles-mêmes, au premier rang desquelles les firmes elles-mêmes.

Celles-ci doivent dans le cadre de leur responsabilité sociétale rendre compte, notamment à leurs financeurs, des contrôles internes et des dispositifs mis en place pour garantir non seulement un fonctionnement des algorithmes qui corresponde aux règles de protection des données personnelles des consommateurs et de la concurrence, mais qui réponde également à des enjeux éthiques[32]. Veiller à la prévention et à la correction d'éventuels biais algorithmiques (liés

31. https://www.gov.uk/government/publications/government-response-to-the-cma-digital-advertising-market-study.
32. Pour une analyse des différents modèles de régulation de l'intelligence artificielle, se reporter à Petit et De Cooman (2020).

aux données ou aux algorithmes eux-mêmes), prévenir l'exploitation des biais de comportements en ligne et garantir le caractère explicable et contestable de décisions algorithmiques qui peuvent avoir des effets majeurs sur le consommateur et/ou le marché, s'inscrit dans cette logique.

SUMMARY: ARTIFICIAL INTELLIGENCE AND MANIPULATION OF MARKET BEHAVIOR: EX-ANTE *EVALUATION IN THE REGULATOR'S ARSENAL*

The development of the digital economy poses questions unprecedented in their magnitude concerning potential market manipulations and manipulations of consumer choice. Deceptive and unfair strategies in consumer law may coexist and mutually reinforce each other with infringements in the field of competition, whether it be algorithmic collusion or abuse of a dominant position. Faced with the difficulty of detecting and sanctioning these practices ex-post, questions are raised about the dissuasive effect of sanctions and their capacity to prevent possibly irreversible damage. To this end, this article considers the supervision tools available to the authorities in charge of market surveillance, to consumers, and to the stakeholders of the companies concerned.

Mots clés : manipulation algorithmique, pratiques trompeuses, pratiques déloyales, surveillance algorithmique
Codes JEL : D18, K21, L86

Keywords: algorithmic manipulation, deceptive practices, unfair practices, algorithmic supervision
JEL Codes: D18, K21, L86

Références

Acquisti A., Brandimarte L. and Loewenstein G., 2020, « Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving it in the Digital Age », *Journal of Consumer Psychology* 30(4): 736-58.

Acquisti A., John L.K. and Loewenstein G., 2013, « What Is Privacy Worth? », *The Journal of Legal Studies* 42(2): 249-74.

Agrawal A., Gans J. and Goldfarb A., 2017, « How AI Will Change Strategy: A Thought Experiment », *Harvard Business Review*, https://hbr.org/2017/10/how-ai-will-change-strategy-a-thought-experiment (25 novembre 2020).

———. 2018, *Prediction Machines. The Simple Economics of Artificial Intelligence*, Boston, Massachusetts, Harvard Business Review Press.

Akerlof G.A., 1991, « Procrastination and Obedience », *American Economic Review* 81(2): 1-19.

Becerra X., 2020, « Attorney General Becerra Announces $113 Million Multistate Settlement Against Apple for Misrepresenting IPhone Batteries and Performance Throttling », *State of California – Department of Justice – Office of the Attorney General*, https://oag.ca.gov/news/press-releases/attorney-general-becerra-announces-113-million-multistate-settlement-against (18 November 2020).

Berdichevsky D. and Neuenschwander E., 1999, « Toward an Ethics of Persuasive Technology », *Communications of the ACM* 42(5): 51-58.

Brandimarte L., Acquisti A. and Loewenstein G., 2013, « Misplaced Confidences: Privacy and the Control Paradox », S*ocial Psychological and Personality Science* 4(3): 340-47.

Buolamwini J., 2018, « Gender Shades », *MIT Media Lab*, https://www.media.mit.edu/publications/full-gender-shades-thesis-17/ (18 September 2020).

Calo R., 2014, « Digital Market Manipulation », *The George Washington Law Review* 82(4): 995-1051.

Calvano E., Calzolari G., Denicolò V. and Pastorello S., 2019, « Algorithmic Pricing What Implications for Competition Policy? », R*eview of Industrial Organization* 55(1): 155-71.

Chopra R., 2020, « Dissenting Statement of Commissioner Rohit Chopra Regarding Zoom Video Communications, Inc. », *Federal Trade Commission*, https://www.ftc.gov/public-statements/2020/11/dissenting-statement-commissioner-rohit-chopra-regarding-zoom-video (18 novembre 2020).

Citron D.K. and Pasquale F., 2014, « The Scored Society: Due Process for Automated Predictions », *Washington Law Review* 89: 1.

Ezrachi A. and Stucke M.E., 2020, *Digitalisation and its impact on innovation*. *R&I Paper Series*, Working Paper 2020/07.

Fogg B.J., 2002, *Persuasive Technology: Using Computers to Change What We Think and Do*. Amsterdam/Boston, Morgan Kaufmann.

FTC, 2020, « FTC Requires Zoom to Enhance Its Security Practices as Part of Settlement », *Federal Trade Commission*, https://www.ftc.gov/news-events/press-releases/2020/11/ftc-requires-zoom-enhance-its-security-practices-part-settlement (18 novembre 2020).

Furman J. et al., (2019), Unlocking Digital Competition, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf.

Gal M., 2019, « Algorithms as Illegal Agreements », *Berkeley Technology Law Journal,* vol. 34, pp.67-118.

Gray C.M. *et al.*, 2018, « The Dark (Patterns) Side of UX Design », *in* Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI' 18, New York, NY, USA: Association for Computing Machinery, 1-14, https://doi.org/10.1145/3173574.3174108 (4 décembre 2020).

Hanson J.D. and Kysar D.A., 1999, « Taking Behavioralism Seriously: The Problem of Market Manipulation », *NYU Law Review,* 74(3), https://www.nyulawreview.org/issues/volume-74-number-3/taking-behavioralism-seriously-the-problem-of-market-manipulation/ (4 décembre 2020).

Kahneman D., 2011, « Thinking, Fast and Slow », Penguin Randomhouse.com.

Kunreuther H. and Ginsberg R., 1978, *Disaster Insurance Protection: Public Policy Lessons*. New York, Wiley.

Karpik L., *Traité de sociologie économique*, Paris, PUF, 2013.

Kenney M. and Zysman J., « The Rise of the Platform Economy », *Issues in science and technology*, 2016, vol. 32, no 3, p. 61.

Kleinberg J. *et al.*, 2018, « Human Decisions and Machine Predictions », *The Quarterly Journal of Economics* 133(1): 237-93.

Koenig, A., 1995, « Patterns and Antipatterns », *J. Object Oriented Program.* 8(1), pp. 45-48.

Judiciary Committee, 2020, *Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations*, US House of Representatives.

Luguri J. and Strahilevitz L., 2019, *Shining a Light on Dark Patterns*, Rochester, NY: Social Science Research Network. SSRN Scholarly Paper, https://papers.ssrn.com/abstract=3431205 (16 septembre 2020).

Madrian B.C. and Shea D.F., 2001, « The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior », *The Quarterly Journal of Economics* 116(4): 1149-87.

Marcellis-Warin (de) N., Munoz J.M. and Warin Th., 2020, « A.I. in Business: Seeing through the Fog of War », *California Management Review*, https://cmr.berkeley.edu/2020/02/ai-fog-of-war/.

Marcellis-Warin (de) N. and Warin Th., 2020, « Government 4.0 and Evidence-Based Policies: A.I. and Data Analytics to the Rescue », *in Handbook of Artificial Intelligence and Robotic Process Automation: Policy and Government Applications*, Anthem Press, 31.

Marciano A., Nicita A. and Ramello G.B., 2020, « Big Data and Big Techs: Understanding the Value of Information in Platform Capitalism », E*uropean Journal of Law and Economics*, https://doi.org/10.1007/s10657-020-09675-1.

Marty F., « Algorithmes de prix, intelligence artificielle et équilibres collusifs », *RIDE* 2017, t. XXXI(2), pp. 83-116.

———. 2019, « Plateformes Numériques, Algorithmes et Discrimination », *Revue de l'OFCE* 164, pp. 91-118.

Marty F. and Warin T., 2019."The Use of AI by Online Intermediation Platforms: Conciliating Economic Efficiency and Ethical Issues", Delphi, 4/2019, pp.217-2250.

Marty F. et Warin Th., 2020a, « Concurrence et innovation dans les écosystèmes numériques à l'ère de l'intelligence artificielle », *Concurrences* 1-2020.

———. 2020b, « Keystone Players and Complementors: An Innovation Perspective », *CIRANO Scientific Series,* 2020-61s, November.

Mathur A. *et al.*, 2019, « Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites », *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1-32.

Mulligan D.K., Regan P.M. and King J., 2020, « The Fertile Dark Matter of Privacy Takes on the Dark Patterns of Surveillance », *Journal of Consumer Psychology* 30(4): 767-73.

Nodder Ch., 2013, *Evil by Design. Interaction Design to Lead Us into Temptation*, 1st edition, Indianapolis, Wiley.

Obar J.A. and Oeldorf-Hirsch A., 2018, « The Clickwrap: A Political Economic Mechanism for Manufacturing Consent on Social Media », *Social Media + Society* 4(3): 2056305118784770.

O'Donoghue T. and Rabin M., 2015, « Present Bias: Lessons Learned and to Be Learned », *American Economic Review*, 105(5): 273-79.

O'Neil C., 2016, *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*, New York, Crown Publishers.

Parkes D.C., R.V. Vohra and other workshop participants, 2019, « Algorithmic and Economic Perspectives on Fairness », *arXiv:1909.05282 [cs]*, http://arxiv.org/abs/1909.05282 (21 September).

Patterson R, 2018, « Algorithmic Opacity and Exclusion in Antitrust Law », *Italian Antitrust Review*, 5(1), pp.23-31, DOI:10.12870/iar-12870.

Petit N., 2020, *Big Tech and the Digital Economy. The Molygopoly Scenario,* Oxford, Oxford University Press.

Petit N. and De Cooman J., 2020, « Models of Law and Regulation for AI », *Robert Schuman Centre for Advanced Studies Research Paper,* EUI Department of Law Research Paper n° RSCAS 2020/63, October.

Rasch A., Thöne M. and Wenzel T., 2020, « Drip Pricing and Its Regulation: Experimental Evidence », *Journal of Economic Behavior & Organization* 176: 353-70.

Smith A, 2020, « Using Artificial Intelligence and Algorithms », US Federal Trade Commission, FTC Business Blog, April, https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms.

Steffel M., Williams E.F. and Pogacar R. 2016. Ethically deployed defaults. Transparency and consumer protection through disclosure and preference articulation. Journal of Marketing Research; https://doi.org/10.1509/jmr.14.0421

Stigler Center, 2019, *Report of the Committee for the Study of Digital Platforms*. University of Chicago.

Sunstein C.R., 2018, *The Cost-benefit Revolution,* MIT Press.

———. 2019, « Sludge and Ordeals », *Duke Law Journa*l 68(8): 1843-83.

———. 2020. « Sludge Audits », *Behavioural Public Policy*: 1-20.

Thaler R.H., 2018, « Nudge, Not Sludge », *Science* (New York, N.Y.) 361(6401): 431.

Thaler R.H.and Sunstein C.R., 2009, N*udge: Improving Decisions About Health, Wealth, and Happiness*, Updated edition, New York, Penguin Books.

Townsend B., 2020, « From SRI to ESG: The Origins of Socially Responsible and Sustainable Investing », *The Journal of Impact and ESG Investing* 1(1): 10-25.

Warin Th. and Leiter D., 2012, « Homogenous Goods Markets: an Empirical Study of Price Dispersion on the Internet », *International Journal of Economics and Business Research* 4(5): 514-29.

Warin Th. and Troadec A., 2016, « Price Strategies in a Big Data World », *Encyclopedia of E-Commerce Development, Implementation, and Management*: 625-38.