CIRANO
Allier savoir et décision

Artificial Intelligence and Market Manipulations:

Ex-ante Evaluation in the Regulator's Arsenal

NATHALIE **DE MARCELLIS-WARIN**
FRÉDÉRIC **MARTY**
EVA **THELISSON**
THIERRY **WARIN**

2020S-64
CAHIER SCIENTIFIQUE
CS

CIRANO
Knowledge into action

Center for Interuniversity Research and Analysis on Organizations

The purpose of the **Working Papers** is to disseminate the results of research conducted by CIRANO research members in order to solicit exchanges and comments. These reports are written in the style of scientific publications. The ideas and opinions expressed in these documents are solely those of the authors.

Les cahiers de la série scientifique visent à rendre accessibles les résultats des recherches effectuées par des chercheurs membres du CIRANO afin de susciter échanges et commentaires. Ces cahiers sont rédigés dans le style des publications scientifiques et n'engagent que leurs auteurs.

CIRANO is a private non-profit organization incorporated under the Quebec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the government of Quebec, and grants and research mandates obtained by its research teams.

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du gouvernement du Québec, de même que des subventions et mandats obtenus par ses équipes de recherche.

ISSN 2292-0838 (online version)

Artificial Intelligence and Market Manipulations: Ex-ante Evaluation in the Regulator's Arsenal

Nathalie de Marcellis-Warin [*], *Frédéric Marty* [†], *Eva Thelisson* [‡], *Thierry Warin* [§]

Abstract/Résumé

The digital economy's development poses questions unprecedented in their magnitude in potential market manipulations and manipulations of consumer choices. 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 sanction's dissuasive effect and its capacity to prevent possibly irreversible damage. To this end, this article considers the available supervision tools for the authorities in charge of market surveillance, the consumers or the stakeholders of the companies concerned.

Keywords/Mots-clés: Algorithmic Manipulation, Deceptive Practices, Unfair Practices, Algorithmic Surveillance

JEL Codes/Codes JEL: D18, K21, L86

[*] Ph.D. Polytechnique Montréal, CIRANO and OBVIA (Canada)
[†] Ph.D. CNRS – Université Côte d'Azur (France) and CIRANO (Canada)
[‡] Ph.D. AI Transparency Institute and Massachusetts Institute of Technologies (USA)
[§] Ph.D. HEC Montréal, CIRANO and OBVIA (Canada)

1. DIGITAL ECONOMY AND THE CONSUMER's RISKS - THE ZOOM AND APPLE CASES

On November 18, 2020, California's Attorney General Xavier Becerra announced a US$113 million agreement between the State of California and Apple regarding the battery management algorithm and performance of iOS-enabled devices (Becerra 2020). The reasons are a misrepresentation of battery life and the algorithmic correction mechanism that has been put in place to manage battery failures (which can lead to the unexpected shutdown of terminals). Nevertheless, the latter was altering the performance of the devices. Therefore, the argument is one of misrepresentation of quality and a correction that masks the initial problem rather than correcting it. Apple had to commit to providing information to its customers to enable them to make informed purchasing decisions and to transparently inform them about the problems affecting their devices and the consequences of its corrective actions on their performance.

This same concern about the transparency and veracity of the information provided to consumers is reflected in the negotiated procedure that ended the procedure initiated by the Federal Trade Commission (from now on, FTC) against Zoom. Zoom is the fastest-growing video conferencing tool during the covid-19 crisis. Created in 2011, the company had 10 million users per day in 2019, mainly self-employed and SMEs. Playing on its freemium model and extending the free conditions to schools and universities reached the threshold of 300 million daily users in April 2020. Its revenues, which were US$622.7 million in 2019, reached US$328.2 million in the first quarter of 2020, US$663.5 million in the second quarter of 2020 and US$1.5 million in the third quarter of 2010[1] and is expected to be between US$685 million and US$690 million in the third quarter.

Because of its activity, Zoom collects data on its users and meeting participants. It also stores the files corresponding to audio and video exchanges, chatting, etc. The case brought by the FTC against Zoom did not include episodes of intrusions into videoconferences (Zoom bombing) but rather dealt with communication that was misleading to users as to the degree of security attached to the service combined with the absence of preventive measures against "commonly known and reasonably foreseeable" threats. In other words, the communication delivered to users was both misleading and concealed the lack of implementation of a reasonable standard of precaution.

[1] https://investors.zoom.us/news-releases/news-release-details/zoom-reports-second-quarter-results-fiscal-year-2021

The FTC denounced four practices: (1) First, Zoom implements misleading communication regarding the encryption of videoconferences. Point-to-point encryption is intended to prevent a third party from accessing the exchanges. However, only part of the exchanges (those via Zoom connect) complies with this claim. (2) Secondly, the company announces 256-byte encryption, whereas the latter only corresponds to a 128-byte key. (3) The third grievance relates to misleading communication about the security level of online video storage. While the company announces encrypted storage as soon as the meeting ends, the encryption can take up to 60 days, the time it takes Zoom to repatriate the files to its storage infrastructure. (4) The fourth grievance relates to an unfair strategy of circumventing security and user privacy and confidentiality measures put in place by third-party companies. This involves installing a web server on the users' computers, tablets and smartphones, allowing a one-click connection (with automatic activation of the camera without an explicit request for consent). This automaticity creates a significant vulnerability in malicious intrusion (especially remote-control execution (RCE) attacks). If an Apple patch is applied through an update to remove this webserver from the concerned devices remotely, the latter can reinstall itself even if the user has uninstalled the Zoom application.

The procedure initiated by the FTC has found an end in a Consent Decree issued on November 9, 2020. According to the terms of the agreement negotiated between the FTC and Zoom, the company commits to implementing a program to secure its service and end deceptive and unfair practices that have compromised users' security and solidity in which they make their decisions. False encryption commitments may have given a false sense of security and altered the terms of consumer trade-offs between competing services. The fact that the proceeding was resolved through a compliance program led to two dissenting opinions within the five-member FTC college, Rebecca Slaughter and Rohit Chopra. These two opinions are particularly interesting because they question the firm's growth strategy and its possible impact on its customers regarding security and privacy.

In her dissenting opinion, Rebecca Slaughter examines the impact of this strategy on consumer protection. Rohit Chopra emphasizes the consequences of the latter in competitive terms. Rebecca Slaughter insists that users' misleading information concealed a trade-off between rapid scale-up (aiming at critical size) and user protection. The quality of the "visible" user experience (the one-click connection) was favoured over the less readily apparent quality. This is not merely a matter of deception but of deliberate attempts to circumvent and thus weaken the protection measures put in place by other companies for their users' benefit. Her dissenting opinion also shows that security alone does not solve all the problems related to privacy protection. In other words, security is a

necessary but not sufficient condition for the protection of privacy. Hence her regret about the absence of a privacy program in Zoom's commitments: « 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[2] . »

Commissioner Rohit Chopra's dissenting opinion disassociates itself from the FTC agreement by insisting on the competitive dimensions (FTC 2020). For him, "deception distorts competition" (Chopra 2020). The idea is that the race for critical size (and thus the market share at which a digital service switches to the degree of dominance that makes it the default application for consumers) involved disclosing misleading information regarding its security level. In other words, "when companies need to act quickly to exploit an opportunity, deploying deception to steal users or sales from competing players is tantalizing."

The conquest of dominance is based on unfair competition vis-à-vis other operators. This dominance is based on the dissemination of misleading information to consumers. As Rohit Chopra points out: "when companies deploy deception, this harms customers and honest competitors, and it distorts the marketplace. Rebecca Slaughter linked data security and user protection, Rohit Chopra links consumer protection with fair competition protection (*fair competition).* He highlights that before the pandemic, the majority of Zoom's paying users were SMEs and independents. Their decisions can be significantly biased by misleading communication because they do not have access to large companies' I.T. departments: « that's why they rely on representations made by those they purchase software and services from. »

In the Apple and Zoom cases, the argument is a vertical differentiation argument. Announcing a 256-byte security level, for example, positions the company at a certain level of vertical differentiation in the eyes of consumers. Zoom benefits from its competitors by not bearing the security costs that could have hindered its growth by offering a product of intrinsic quality below this vertical differentiation level. The Zoom example is a particular example of misinformation and lack of incentives offered to technology companies. Here, it is about the lack of incentives to move in respect to product quality.

Given these two recent examples, positioning in vertical differentiation is extremely important for the tech industry. Let us recall that these companies make architectural innovations and contribute

[2] The highlighted example is Facebook's commitments to the FTC in July 2019.

to implementing the new global technological infrastructure. The technologies and business models based on these technologies also make everything go very fast. Therefore, if the diagnosis that misinformation distorts competition is accurate, the fact remains that ex-post control can show limits in deterring practices that do not meet the standards of a level playing field and consumer protection.

In this article, we will focus on the more general case of market manipulation by these firms. We will discuss the current dilemma: tech companies offer products and services and may misrepresent them. The risk illustrated by these two cases is that ex post-detection (negotiated procedures and possibly repairs) are insufficient to dissuade, ex-ante, these practices. Not to be misunderstood here, we do believe A.I. is a formidable revolution and brings opportunities that we have barely started to see in all domains, from more efficient markets to government policies (de Marcellis-Warin and Warin 2020). However, when tech companies rely both on business models fed by massive data and decision models based on artificial intelligence techniques, everything is moving exponentially faster, including the development and marketing strategies of these companies' products and services. In this context, we believe that the current system needs to develop new tools to avoid anticompetitive or even unethical behaviours to allow markets and societies to benefit from A.I.'s extraordinary promises (Marty and Warin 2020d).

We will propose an argument favouring an ex-ante mechanism: a transparency index for tech companies to ex-post regulation mechanisms. An ex-ante approach seems to us to be an essential tool to add to the regulator's arsenal for several reasons. It is a question of guaranteeing responsible behaviour on the market, avoiding the occurrence of significant and irreversible risks. The latter may result in irreparable damage to competition (irreversible dominance, tacit collusion that cannot be sanctioned...) or significant consequences for consumers linked to discrimination or severe violations of privacy). This point highlights the notion of high-stakes decisions, which describes the consequences of algorithmic choices that can have significant consequences on competition or consumers. This finally leads to the consideration of ex-ante measures allowing market regulators to detect some of these practices before their effects are irreversible and companies themselves to prevent such outcomes, both in managing their legal risks and implementing their strategy social responsibility.

This article aims to analyze companies' practices in online markets that can manipulate consumer choice and distort competition. These practices, which may be based on algorithmic manipulations, deliberately misleading choice presentation architectures or disseminating false information on the quality of services, present two particularly problematic characteristics. Firstly,

they are not visible ex-ante by consumers facing incomplete and asymmetrical information and are not perceptible ex-posts after consuming the service. They are, therefore, much more "trust goods" than "experience goods." Secondly, market supervisors' detection is challenging for two reasons; one is due to the difficulties of detecting algorithmic strategies, and the other is because some of its practices lie at the confluence of data protection, consumer protection, and competition.

Authorities such as the FTC in the United States or the CMA in the United Kingdom, whose remit is vast, may be better placed than other authorities specializing in a particular area. As such, the use of negotiated procedures as a means of settling proceedings or in the context of market investigations along the lines of the British market investigations model introduced by the 2002 Enterprise Act may be interesting avenues to consider. Indeed, the question arises as to the capacity of ex-post sanctions alone to respond to these risks. The two examples cited above can illustrate some of the possible ways to prevent these practices before they create damage or enable the stakeholders to modify their behaviour to limit the consequences:

The first of these is sunshine regulation. Revealing a given company's practices to consumers can make it possible to modify their behaviour and lead to a sanction by the market. This path is at the source of one of the divergences between the majority of FTC commissioners and Rohit Chopra. The latter would have liked Zoom to be obliged to inform all its users about its past actions. For the majority, « 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.

The second path is that of control of the firm by its various stakeholders, notably investors. Failure to comply ex-post by the market supervisory authorities may lead to additional sanctions through individual financiers' withdrawal committed to ethical and responsible policies. However, the latter may legitimately wish to ensure that the firm's strategy complies with their values outside of public procedures. This presupposes the development of indicators enabling all stakeholders to detect such risks. Firms' social responsibility presupposes that they equip themselves with instruments for self-evaluating their practices, and those tools for evaluating firms' algorithmic responsibility be developed by independent institutions.

These tools are all the more important since the development of algorithmic decisions poses unprecedented problems of understanding the predictions that the latter make (notably because of the diffusion of artificial intelligence tools). The issue at stake is the explicability of choices and the detection of manipulative practices. We focus in this paper on algorithmic manipulations and the consequences in terms of privacy and competition protection.

The article is structured as follows. Section 2 presents the risks that increasing the use of Artificial Intelligence (A.I.) in algorithmic recommendation systems may induce consumers to reduce their freedom of choice, behavioural manipulations, or even through personalized but unbalanced contractual conditions. A third section shows that consumer damage can also occur indirectly through an infringement of competition, through the consolidation of individual dominant positions or the increase of algorithmic collusion risks. A fourth section focuses on solutions for regulating algorithms by other algorithms or monitoring procedures, allowing accountability for the actual functioning of the algorithms implemented. The first avenue uses algorithmic tools by the market supervisors to detect possible abnormal patterns that could lead to procedures. A second way may involve the use of algorithmic countermeasures by consumers. A fifth section considers how "highly consequential decisions" can be the subject of special attention by the firm's different stakeholders. Taking these decisions into account can lead to devices' design, allowing firms to guarantee their algorithms' integrity and ethics through scoring methods. We present the confidence index, developed by the *A.I. Transparency Institute*, adapting it to the issues related to consumer and competition law.

2. ARTIFICIAL INTELLIGENCE AS A POTENTIAL VECTOR OF RISKS FOR CONSUMERS

Three types of damage to the consumer can be considered. The first type of damage is the reduction of the choices open to him. The second damage lies in the possibility of manipulation of choices. The range of available solutions is not artificially closed, but the consumer's behaviour is altered by producing stimuli intended to bias their decision. The third type of damage corresponds to what we call an abuse of exploitation. The ability to predict the characteristics of the consumer (technical expertise, ability to pay, etc....) makes it possible to make offers leading to the extraction of the consumer's entire surplus (which would not be possible with uniform prices or imperfectly differentiated prices) or leading to discriminatory offers, whether in terms of price or the quality of the products and services offered.

2.1 Reducing the consumers' choice space

A.I. allows changes that break the usual business paradigms. A.I. can be used within the framework of strategies that can lead, through increasingly finely targeted recommendations or recommendations to limit consumers' freedom of choice. The latter may see their range of choices reduced according to their past consumption or according to the customer segment to which the algorithm links them; A.I. is a prediction tool based on machine learning. (Agrawal, Gans, and Goldfarb, 2018). In other words, they can be enclosed in the equivalent of a filter bubble. Such an

effect may be aggravated by shifting some platforms from a shopping-then-shipping logic to a shipping-then-shopping logic (Agrawal, Gans, and Goldfarb 2017). The customer may indeed incur a cost to reship the product even if the return would be free.

A.I. can also facilitate practices that can lead to manipulating consumer choices through a precise understanding of their behaviour or an accurate estimate of their maximum payment capacity. Indeed, as noted by Ezrachi and Stucke (2020): « […] in a data-driven economy, personal data on user behaviour, preferences, weaknesses, and habits is the new currency for the advertising - and marketing dependent – business models ». These capabilities require the monitoring of massive, diversified and continuously updated data and the mastery of analytical tools to customize offers concerning consumers' decision-making frameworks or better predict other operators' strategies (Marty and Warin 2020b, 2020c). The takeover of Onavo by Facebook or that of Looker by Google testifies the importance of the ability to master the competitive environment through the technical possibility of better and better forecasting the present.

Several examples of the limitation of consumers' capacity for choice could be added to the filter bubble and shipping-then-shopping models presented above. Some are related to cost factors. They are often linked to complementarities between types of equipment that play as many factors aggravating the costs of change between ecosystems. Biases can then arise from how choices are proposed. The case of personal assistants shows how the options proposed can be reduced to a minimal range. The last example, developed by Ezrachi and Stucke (2020), is based on the possibility of controlling the dissemination of innovations in ecosystems by the pivotal operator of each of them.

To illustrate their thinking, Ezrachi and Stucke (2020, p. 42) are based on Rodgers' innovation diffusion model (2003). The adoption of an innovation by a given individual is described as a five-phase process. The first is the knowledge phase. The individual must be informed of the availability of an innovative solution and its functions. The second is that of persuasion. It is through it that the individual forms favourable or unfavourable anticipations towards it. The third is that of the decision of adoption or non-adoption. The fourth is that of implementation. The fifth, finally, is that of the confirmation of the adoption. It can be confirmed by observing the choice of third parties or, on the contrary, be negatively affected by negative messages.

The strategic action of platforms - if they can detect their users' behaviour - can play in favour of the adoption of an innovation developed internally (or by a preferred complement-maker) and unfavourably for an innovation developed by an independent firm. The first variant may explain

why digital ecosystems promote early and massive adoption of innovations. The second may explain why innovations fail to spread.

Steps in the Dissemination process	Favourable pivot Strategy	Unfavourable Pivot Strategy
Knowledge	Ability to propose, to put forward	Reduce the possibilities of information about a potentially available innovation or access to information about how it works (by algorithmic manipulation of the search engine, for example, by de-referencing sites ...)
Persuasion	Ability to target, to demonstrate suitability for personalized needs; attention strategies; identification of possible early adopters and dissemination of personalized information to potential followers	Production of negative opinions or the creation of frictions makes it more difficult to download or interoperate with the ecosystem's various services[3].
Decision	Personalized marketing; free trials; play on friends' recommendations.	Friction blocking: play on status quo behavioural bias - default settings are rarely changed by agents, regardless of their preferences[4]
Implementation	Facilitation of adaptations, bug fixes	Users can be continuously redirected to less efficient options but dependent on the ecosystem.
Confirmation	Redirections by support tools towards the innovation	The pivot firm may degrade the performance of complementary services provided by the competitor to redirect consumers towards better-controlled service.

Table 1: The control of consumers' innovation adoption behaviour in digital ecosystems

2.2 The manipulation of consumers' behaviour

Markets' DNA is ultimately the price mechanism. The latter plays an important function: it informs market participants before they make a decision. There is abundant literature on the concept of "value of information," and some authors have looked at the value of information in the context of price strategies on digital markets (Warin and Leiter 2012; Warin and Troadec 2016) as well as from a regulatory perspective (Marciano, Nicita, and Ramello 2020). With A.I., now some firms

[3] This is the notion of bad sludges that we will detail below.

[4] « As behavioral economics littérature shows, the setting of the default can often determine the outcome (even when transaction costs are minimal) » (Ezrachi and Stucke, 2020, p.48). As Ezrachi and Stucke point out, inertia (the stau quo bias) is not the only reason why consumers may keep default options that are not the best for them. If they have standard skills, they may consider the default choices to be the most favorable in terms of service quality and performance.

have access to the aggregated information and the customer's value of information, through notably recommender systems. With this rich access to the value of information, A.I. can be used to model consumer behaviour and create an incentive leading to purchase (emotional pitch, dark nudge...) at the right time. These problems go beyond the scope of A.I. alone in that they can be observed within the framework of traditional algorithms. For example, many merchant sites can implement drip pricing strategies (Rasch, Thöne, and Wenzel 2020) or price partitioning. The customer can be engaged in a purchasing process by an attractive call price and only "discover" the full price later. The time spent to complete the subsequent pages will make him forget the competitors' price consulted at the beginning of his search, or she will be reluctant to start his search process from scratch (Marty, 2019).

The notion of dark patterns illustrates these practices, which can be aggravated by A.I. performance (Stigler Center, 2019). It covers all the profiling methods, algorithmic proposals or user interfaces that can restrict the ability to make a free and informed choice on the consumer's part. Dark patterns are also called dark nudges or bad sludges. Therefore, they cover strategies that increase the opacity of consumers' choices, making it more difficult for them to express their preferences freely, or that leads them to make decisions that they would not have made spontaneously.

Dark patterns can be produced to lead consumers to make decisions that are not in line with their preferences. While a (positive) nudge is theoretically part of a logic of liberal paternalism - leading the individual to behave in a way that is in line with his interest and the general interest - such dark nudges aim to lead them to act in a way that is not in line with their interests (C. Sunstein 2019; Thaler 2018). Therefore, it is a question of manipulating consumer choices by voluntarily altering their preferences or even creating them through cognitive biases (framing effect, sunk cost fallacy, anchoring, etc.). Dark sludges can therefore be defined as « an evil nudge [...] that can exploit [online consumers] cognitive biases to persuade them to do something that is undesirable, typically by introducing excessive friction into choice architecture » (C. R. Sunstein 2020).

It is essential to distinguish within dark patterns the nuance between bad nudges and bad sludges. A nudge can be defined as an encouragement, a small nudge that leads the agent to act in a given direction. It is a push towards action. It is often presented as positive (the agent is encouraged to act in his interest as long as he does not spontaneously). It can, however, be harmful. For example, a bad nudge is used as part of an emotional sales pitch: a banner appears through an untimely window leading to a click to access a given service for which the consumer is known to have

developed an addiction. Therefore, it is a push - in the sense of a stimulus - to make the consumer "fall" to the side we know she tends to lean to.

The concept of nudge in the behavioural economics literature was mainly based on reflecting on actors' decision-making environment to make better-informed choices without restricting their freedom. Therefore, the aim was to promote a choice architecture that reconciles autonomy and "signalling" the best options for the agent himself (Thaler and Sunstein 2009).

However, this thrust - through the design of the architecture of choice - can also be exercised in a much less benevolent way. As noted above, it can be used in the firm's interest and to the consumer's detriment. It is no longer a question of inciting the consumer to make the right decision for her but pushing him to decide following the firm's interests. Therefore, a nudge can be positive and negative and can be part of a dark pattern.

Conversely, the term sludge evokes friction. It is more in the primary sense of getting bogged down, of losing mobility. It is about creating artificial difficulties to prevent consumers from exercising their freedom of choice, from identifying the most favourable options or, conversely, from exercising them. Thaler (2018) gives a simple but evocative example. The consumer decides to purchase a good or subscribe to an online service by considering the offer of a deferred refund. However, this benefit is conditional on sending proof of purchase by mail within a given period (or the tedious creation of an online account). Many consumers who have based their decisions on this rebate will not claim it. As Thaler notes (2018, 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.'»

A sludge is defined as a « kind of friction, large or small, that people face when they want to go in one or another direction » (C. Sunstein 2019). It can also help prevent a harmful attitude on the part of the consumer himself (curbing a shopping frenzy, ensuring his eligibility conditions or characteristics, forcing him to benefit from a cooling-off period) and, on the contrary, hindering access to legitimate rights.[5]Like a nudge, a sludge is based on the exploitation - through the strategic establishment of the architecture of choice - of economic agents' behavioural biases. What biases can a friction mechanism take advantage of? These can be, for example, inertia biases (Madrian and Shea 2001), of procrastination (Akerlof 1991) and preferably for the present (O'Donoghue and Rabin 2015).

[5] The (bad) sludges can be both private and public. In other words, administrative constraints can act as barriers to entry for some people to access rights or to activate certain procedures in their favor.

The notion of bad sludge is also found in disputes between companies at the heart of major digital ecosystems and their suppliers. For example, as noted in the report on competition in the digital sector published in October 2020 by the Antitrust Subcommittee of the U.S. House of Representatives, the notion of "bad sludge" can be found in disputes between companies at the heart of large digital ecosystems and their suppliers (Judiciary Committee House of representatives, 2020, p.218), one of the arguments raised by EPIC Games in its lawsuit against Google in the U.S.[6] Game developer Fortnite insists that bypassing the Play Store - which is technically possible and easy - is made more difficult and stressful for the consumer to deter them from downloading the game directly :

« 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 ».

Friction is not limited to the initial download but also updates. Still, according to the complaint filed last August by 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 O.S. 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

[6] Complaint for injunctive relief, Epic Games v Google LLC, n°3:20-cv-05671 - ND Cal., Aug., 13, 2020, §96

further drives consumers away from direct downloading and toward Google's monopolized app store » (Ibid, §98).

What is here the potential usefulness of these dark nudges? Maintaining the online application store as a lock on access to the ecosystem (gatekeeper site downloading) ensures the core business's private regulatory power (structuring power).

Therefore, controlling the ecosystem involves imposing technical and psychological frictions to counter the threat of loss of control through direct application downloading.[7] :

« 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. »

The notion of dark pattern goes beyond manipulating choices and behaviours; it can also relate to Internet users' information disclosure (or excessive privacy reduction). In such a case of a nudge, the design of the site or the modes of presentation of the choices make the user go beyond what is necessary or what he would have accepted if his choice had responded to the rationality of type 2 (Acquisti, Brandimarte, and Loewenstein 2020). It is not just a matter of exploiting consumers' or users' vulnerabilities to induce them to make choices that correspond to trends that they could rationally try to curb but in the extreme of eliciting (i.e., constructing) these preferences (Mulligan, Regan, and King 2020). The dark pattern can, therefore, result from the design of a site. This is the case of clickwraps that lead the consumer to make choices in blocks for questions of very different kinds and importance (Obar and Oeldorf-Hirsch 2018). There are, therefore, "manipulative by design" devices. Their analysis is not new in the online world (Calo 2014) or even in the offline world (Hanson and Kysar 1999). However, the nature of the online journey and

[7] It may be interesting for our purposes to compare the terms of the litigation between Epic Games and Google and those between Epic Games and Apple. The complaint for injunctive relief filed in the US District Court for the Northern District of California on August 13, 2020, also describes an impossibility of access to Apple's customers outside the firm's application store. However, the practices at issue do not involve the imposition of friction on users who would like to bypass the application store, as is the case with Google, but rather the impossibility of doing so. This impossibility is due to technical and contractual restrictions. From a technical point of view, customers cannot install an alternative application store to the App Store on their terminals (pt.58). Beyond this lock, the iOS operating system prevents, through technical restrictions, the direct downloading of applications from websites by bypassing the application store (pt.66). Finally, the lock is also contractual. Developers may only be present on the App Store if they agree not to make downloads possible through alternative channels: "[...] to access the iOS user base, app developers 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).

the ability to capture, process, infer and create targeted stimuli make the effects far more far-reaching.

The modalities for implementing these practices can be fundamental. An Internet user, who is asked several times for his personal data protection preferences after each refusal, will accept, either inadvertently or to stop the requests (Luguri and Strahilevitz 2019).[8] The same applies to the techniques described above in which the price is only revealed at the end of the purchase process or transactions in which options that the consumer would not have wished to subscribe to are "preselected" in a non-obvious way.[9]

In the same way, the notion of a dark pattern can cover emotional pitch practices. The idea is to produce a stimulus that would make the consumer switch to an act of purchase. Such stimuli can be the announcement (except in the form of an untimely window) of a price reduction or the announcement of a limited number of remaining products coupled with messages indicating that another Internet user has just bought one.[10] It is then a question of creating a sense of urgency, which will push the consumer to rush his purchase for fear of running out of available stock. (Mathur et al., 2019). In such situations, it is a question of calling upon the system 1 of our brain that of the fast, instinctive guided by routines, emotional choice... and no longer upon system 2, that of rational choice[11] (Kahneman 2011). Whether it is a question of "pushes or frictions," the architecture of the choices and the exact path of the Internet user (and of the consumer in general) exerts a determining influence and calls for questioning the firm's responsibility that sets them up.

A.I.'s development could make these strategies more useful by allowing a better understanding of consumer behaviour after closely linking it to a given segment based on observed and inferred characteristics. In other words, A.I. can promote the personalization of prices and the personalization of manipulations. Indeed, as the Stigler Center notes (2019, p. 238), the use of sludges will have multiplier effects through the use of A.I.: « *Dark patterns are often used to*

[8] Luguri and Strahilevitz (2019) on a representative sample of 1762 American Internet users show the influence of dark patterns on individual online privacy choices. The use of mild dark patterns increases the rate of choice of a proposed data protection regime from 11% to 26% (+228%). The use of aggressive dark patterns increases the choice of a proposed data protection regime from 11% to 42% (+371%). These practices leading an Internet user to unnecessarily waive the protection of his personal data correspond to the notion of *privacy zuckering*.

[9] These practices can also lead to the involuntary addition of a product in a basket in the case of a marketplace or the choice of unnecessary insurance when booking air travel.

[10] Here, the supplier plays on a behavioral bias related to loss aversion.

[11] It should be noted that in the same perspective, a sludge can be favorable to the consumer in that it gives him time to make a "cold" decision and prevents him from making an impulsive commitment.

direct users towards outcomes that involve 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[12] » (Gray et al. 2018).

For example, A.I. can make it possible to determine which stimulus to present to a consumer and when to do so based on an increasingly refined prediction of its characteristics and, therefore, also of its inferred weaknesses. Luguri and Strahilevitz (2019) propose a detailed typology of the different mechanisms linked to dark patterns. We reproduce it in part in table 2 presented below, adding elements developed by Gray et al. (2018).

Practice categories	**Variant**	**Description**
Nagging	The same option, although declined, is presented multiple times in different forms	Repeated requests to make a choice or formatting of a choice that is not final (not versus not now)
Social proof	Messages about the activity of third parties	False announcements that third parties are buying or posting comments
	Contributors' Notices	False Notices
Obstruction	A model of friction (sludge) designed to impede an action or choice.	
	Roach motel	Easy entry into a much more complex and time-consuming benefit/unsubscription or waiver process
	Obstacles to price comparison	The consumer is prevented (through an interface that prevents copying and pasting the selected product's characteristics, for example) from comparing the price on a competitor's site.
	Blur on currencies	Options or services are displayed in different currencies, or the site relies on the use of a virtual currency (a token) that hinders price comparison.
	Immortal account	It is impossible to delete your account definitively.
Sneaking	Sneak into basket	An unrequested item is added by default to the order

[12] Gray et al. (2018) définissent les modèles sombres comme "des cas où les concepteurs utilisent leur connaissance du comportement humain (par exemple la psychologie) et les désirs des utilisateurs finaux pour mettre en œuvre des fonctionnalités trompeuses qui ne sont pas dans le meilleur intérêt de l'utilisateur".

	Hidden costs	Real costs are hidden or revealed very late in the purchasing process (taxes, foreign exchange commission for purchase abroad or prohibitive shipping costs).
	Hidden Subscription/Forced Continuity	Unannounced Tacit Renewal
	Bait & switch	Purchase does not correspond to what was initially presented
	Unsubscription trap	Unsubscribing is coupled with the acceptance of the reuse or resale of personal data.
Interface handling	Information concealment/"aesthetic" manipulation	The most important information is the least highlighted information on the user interface, or the presentation makes the options the least interesting, the most visible or attractive.
	.	The position of the options can also vary over time to lead the customer to click on an unwanted option
	Preselection	Options (unfavourable to the consumer) are checked by default
	Emotions game	The presentation of the pages and options manipulates the user's emotions (colour, vocabulary, illustrations...)
	False hierarchy, pressure to buy	The behaviour is manipulated to lead the user to choose the most expensive option or the most extended commitment.
	Trapping issues	Deliberate ambiguity - the user thinks he is answering a simple question, while the implications of his choice are more complex Variant: questions based on double negatives or unnecessarily complex and ambiguous vocabulary
	Disguised advertising	The Internet user is led to click on a link that does not appear clearly as an advertisement.
	C*onfirmshaming*	The presentation of the option is such that a refusal is reported as stupid.
	Cuteness	A digital assistant is presented in order to play on the feelings of the Internet user.
Constraint action	Spam on the actions of third parties, social pyramid, exploitation (Address book leeching)	Manipulative extraction of information related to the user's contacts

	P*rivacy zuckering*	Consumers are led to make personal information public without them being aware of it.
	Gamification	Returning to the site or selecting specific options on the site is likely to win something for the user.
	Forced identification	The identification on the site is presented as indispensable
	Be directed in the wrong direction	The user's attention is focused on one element to the detriment of a second, more decisive element.
Feeling of rarity	A message announcing a low number of available products	Creating a sense of urgency about the act of buying (loss aversion gambling)
	A message indicating that the demand for the product is strong	The consumer is warned that other Internet users (as personalized as possible) are in the process of making purchases.
Sense of urgency	Countdown	An indicator highlighted on the page shows that the availability time of the offer is steadily reducing
	Untimely window informing that the offer is no longer available for a very long time	The purchase option will disappear in a few moments.

Table 2: A typology of Dark Patterns

The existence of manipulative strategies explains the difference between agents' preferences, such as protecting their data and their actual behaviour. Acquisti et al. (2013) have shown that agents may agree to pay to protect their data. Similarly, their online behaviour (multiple email addresses, specialization of different social networks...) also attests to this strategy (Acquisiti et al., 2020). However, the case of the dark patterns described above shows that it is necessary to distinguish between willingness and opportunity. Psychological biases can thwart the expression of preferences and open up the possibility of manipulations such as those we have just detailed. Table 3 below reproduces the elements presented by Acquisti et al. (2020).

Psychological bias	**Description**	**Possible consequences**	**Ability to manipulate for the firm**
Information asymmetries	Users cannot anticipate the use that will be made of their data.	Impossible to hedge against a risk that cannot	It is in the firm's interest not to make its practices transparent.

		be anticipated or measured	
Limited Rationality	The user cannot integrate all the choice parameters in his decision.	Few can understand the general rules of use (and are willing to read them).	The more complex and technical the rules are, the more obscure they will be, and the less the user will evaluate them.
Presenteeism bias	Overestimation of short-term gains compared to long-term costs	A small incentive to enter into a contract is enough to change the consumer's arbitration even if the future risks are high.	Provide a tangible and immediate incentive to push for data sharing
Biais in valuation of intangibles	Intangible parameters are difficult to isolate, quantify and therefore to take into account.	Low personal data protection's negative consequences are diffuse and difficult to relate to a specific decision (distant in time).	The opacity of default clauses (relating to data use) reduces the reputational risk associated with the sale of data or its strategic use (for example, through discriminatory pricing).
Built preferences	We reason on heuristics that do not take into account the objective costs and benefits.	Individuals do not change the default rules that are proposed to them.	Default rules are more advantageous for the platform than for the customer.
The illusion of control	The agent's perception of real control leads him to take excessive risks.	The illusion can be all the more robust, the more refined the granularity of the options is.	Multiple questions with multiple options play like a negative nudge
Panurgism	Calibrate one's behaviour to that of third parties	Information-sharing decisions are linked to those disclosed by third parties.	Push for disclosures presented as a standard
-adaptation	The agent does not adjust his behaviour in the face of changing circumstances.	Even if the personal data protection policy deteriorates, agents do not revise their initial choices	Gradually degrade conditions for users
Incentives to disclose	Architecture "pushes" individuals to share personal information	Risk of disseminating (e.g., on social networks) self-incriminating information	Behavioural manipulations to share content online ("XX people have seen your photos")

Table 3: Consumers' psychological biases

Some of the effects are related to behavioural biases; others are created and amplified by practices that play on these biases. Errors in risk estimation do not need to be supported by the actions of the sites concerned. A loss of confidentiality is potentially very high but is assessed as low

probability and is distant. Therefore, as in other areas that can induce systemic risks, they will be minimized in decision-making (Kunreuther and Ginsberg 1978).

For other dimensions, it is possible to find the notion of dark patterns presented here. According to Acquisti et al. (2020), it is possible to illustrate how these biases can be instrumentalized. The consumer will overvalue a tangible and immediate gain compared to a deferred and intangible risk. Moreover, as we have seen for bad sludges, it is not even a given to meet the conditions for activating these offers ex-post. The same sensitivity to bad sludges or bad nudges may stem from agents' uncertainty about their preferences. The presentation of options can, therefore, guide agents' decisions... even against their interests. As presented in the table above, apparently "transparent" devices can intuitively work against consumers' interests (Acquisti, John, and Loewenstein 2013). Indeed, the illusion of control - through broad possibilities for personalizing choices - can lead users to accept excessive disclosure of their data (Brandimarte, Acquisti, and Loewenstein 2013). Similarly, gradual but continuous degradation of data protection may be perceived by users but not be subject to a revision of initial choices: "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, Brandimarte, and Loewenstein 2020).

2.3 The imposition of unbalanced contractual conditions

A.I. can lead to an excellent segmentation of customers allowing them to propose almost personalized prices. The problem with the latter is that placed at the level of the consumer's maximum propensity to pay, a perfectly discriminating price makes it possible to confiscate the totality of the consumer's surplus. There is no damage in terms of efficiency in economic terms, but an undue transfer of welfare compared to the distribution that would prevail in perfect competition.

Moreover, it should be noted that discrimination between customers may also prevail in uniform pricing. If it is possible to determine the customer's needs and the level of technical expertise, it is conceivable to offer him a product with less attractive characteristics or with degraded performance. The vendor can take advantage of the informational advantage he knows he has over his customer and the production flexibility that will increasingly allow him to use Industry 4.0 models. These are the so-called versioning practices. The less expert consumer may be offered products or services that are ultimately more expensive than the personalized offer's intrinsic value. (Marty, 2019).

The notion of "augmented dark pattern" can, therefore, cover several modalities. The one we have just seen corresponds to manipulation by transaction costs. The consumer's well-being is degraded through exploitative abuses in the form of a confiscated personalized price (of his surplus), an offer with a degraded price/quality ratio, or in the form of barriers to exit. The notion of dark pattern also covers consumer behaviour manipulations, based on the acceptable identification of their characteristics and, more precisely, their weaknesses.

3. ARTIFICIAL INTELLIGENCE AND THE POTENTIAL DAMAGE TO THE COMPETITIVE PROCESS

Even if this advantage is based on merits and should not be sanctioned based on competition rules, it is an obstacle to a level playing field and is likely to tip the market towards a situation of overwhelming and possibly perennial dominance. Indeed, this advantage may mean that a new entrant could not immediately be as efficient as the dominant firm. The market would then no longer be contestable in the economic sense of the term. The advantage in terms of A.I. and computational capacity would likely become an impassable barrier to entry. The damage would perhaps not be damage in terms of efficiency (both for the economy and for the consumer) or innovation, but damage to the competition process itself (3.1). Simultaneously, the use of A.I. in an oligopolistic market structure can lead to the faster emergence of collusive equilibria, which are more stable, insofar as competing firms gain a better understanding of the market and can predict their common reactions more and more precisely (3.2).

3.1 The locking of a dominant position

In such a case, the use of A.I. is aimed less at exploiting an advantage vis-à-vis the consumer than at acquiring, consolidating or extending a dominant position to the detriment of its current competitors or potential competitors, whether in the same relevant market or related markets (of complementary goods and services, for example). The practices in question may be favourable to the consumer in terms of welfare, at least in the short term. They may nevertheless be detrimental to the continuation of free, undistorted competition on the merits.

Dominant Internet firms may have a competitive advantage over their current or future competitors and their trading partners, in other words, firms that act as complementors in their respective ecosystems (Marty and Warin 2020c, 2020a). Abundant literature in economics and management sciences is developing on kill zones and the notion of killer mergers or consolidating acquisitions (Marty and Warin 2020e). The possession of massive, continuously renewed, and diversified data and the design of A.I. algorithms coupled with the development of computational resources can enable dominant companies (the pivotal firms of each ecosystem) to identify competitive threats,

promising technologies or potentially disruptive developments at a very early stage. Therefore, they can eliminate or clone the service or even buy out the company concerned well before entering the market.

The consumer may not be harmed in the short term. A possible innovation will not be eliminated; it can be integrated into the dominant firm's offer and may be more efficient and attractive. However, the advanced detection capability perpetuates dominance by acting as a barrier to market entry. (Anticompetitive) foreclosure arises from a capacity to detect weak signals in the market. Different algorithmic tools facilitate these nowcasting practices. Sentiment analysis is one of the main ones.

3.2 The emergence and consolidation of algorithmic collusions

The practices described above correspond to unilateral practices, i.e. practices implemented by a dominant firm independently of its competitors. Algorithms in general and A.I., in particular, can also facilitate development, if not the emergence of collusive equilibria. Once again, abundant literature has developed on the question of the ability of A.I. to promote and stabilize tacit collusion equilibria (Calvano et al., 2019). These are situations in which algorithms capable of autonomous machine learning by understanding the functioning of the market and the reactions of competitors spontaneously converge towards a cooperative equilibrium (i.e., armed peace) insofar as this is the one that maximizes everyone's profits over the long term.

Human players would potentially arrive at the same result, but under much more restrictive assumptions (in terms of the number of participants, the complexity of the environment, etc.) and over a much more extended period. Moreover, such an equilibrium would be much more stable with A.I. than with humans insofar as the former could present less cognitive bias, leading them to misinterpret their competitors' strategies or overreact in case of observed deviation from the tacit collusion equilibrium. Moreover, the demonstration of an anticompetitive intention would be much more challenging to make, which would likely significantly reduce the probability of being sanctioned by competition rules (Marty 2017).

Finally, despite the efficiency gains that A.I. will bring and the development of computational capabilities, the academic literature insists on the associated risks. Thus, the race towards A.I. would not only be the solution to the search for efficiency in our algorithm-driven economies; it could also prove to be a future problem. Our fourth section shows that while A.I. may be a problem, it may also be the solution to control these possible risks. However, it is a question of considering these avenues in their practical, legal and ethical dimensions.

4. AVENUES FOR ALGORITHM REGULATION BY ALGORITHMS

The tools provided by A.I. can help to correct their possible anticompetitive or harmful effects on consumers. They can be implemented by the competition regulator (4.1) or by consumers themselves (4.2).

4.1 The use of artificial intelligence by market supervisors to prevent manipulative strategies

The resources provided by the A.I. can be used ex-ante as part of algorithm validation procedures (in a logic of requiring conformity by design) or as part of sectoral surveys. The algorithms are then used based on market data transmitted by companies to observe possible biases. In the case of collusion by algorithms, these checks could function as stress tests. It would be a matter of seeing under what conditions and with what speed the competing firms' algorithms could converge towards such equilibria through algorithmic collusion incubators. It would be up to the regulator and the firms to define the conditions (speed or frequency of price changes, for example) to limit competitive risks.

Then, the algorithms can be used ex-post in market surveillance. As is already the case in financial market regulation for high-frequency transactions, it is possible to analyze market patterns to detect strategies that would not make economic sense outside of abusive practice. It is then up to the company concerned to prove that its decisions were not part of such a strategy (the logic of complying or explain). Note that competition law is not the only legal tool that can make algorithms accountable. For example, when it comes to offers to final consumers (in terms of price or quality of the products offered), practices of a discriminatory nature could be subject to consumer protection measures how so many unfair or misleading practices. However, it should be noted that the finer the analysis produced by the algorithm of the customer's needs and characteristics, the more difficult it will be to detect manipulation and detect damage.

The Stigler Centre (2019, p.254) proposes some avenues discussed in its paper. A concealed dark pattern that increases consumers' exit costs or is likely to exploit weaker consumers' weaknesses must be subject to a presumption of consumer harm. In other words, manipulative intent would be presumed if the algorithm's design appears to obscure the developer's intent and its effects intentionally. The practices referred to by the Stigler Center (2019) are worth discussing, particularly concerning the development of A.I.

4.2 The use of artificial intelligence by consumers

Consumers themselves can use algorithmic tools to detect or counteract possible manipulation by firms. This can be done through distributed monitoring devices set up by non-profit institutions or through over-the-counter algorithms by consumers themselves. The latter can play with firms' algorithmic strategies or even deceive them by emitting false signals.

It should be noted that all consumers' segments are not equivalent to algorithmic manipulations and the responses they are likely to make to them. First, the probability of being manipulated even by a mild dark pattern depends on the consumer's level of knowledge (i.e., often their level of education). Thus, the ability to perceive the manipulation itself also depends on the consumer's level of knowledge (i.e., often their level of education). Second, access even to such countermeasures is related to the same characteristics. In other words, possible algorithmic manipulations will not affect different categories of consumers in the same way, which raises both ethical (Marty and Warin 2020d) and distributional issues.

Consumers may misperceive the use of A.I. by platforms. The personalization of prices and offers can be analyzed as discrimination aimed at "exploitative abuse." The "guiding" of choice behaviour, especially if perceived as part of a dark pattern, can be interpreted as an unfair and prejudicial manipulative practice (deceptive practices). In doing so, a phenomenon of consumer backlash can be observed (Stigler Center, 2019). This may result in measured but still damaging strategies on the part of firms. The use of less obvious dark nudges (mild dark patterns) may make it possible to obtain the desired effects in the most naïve consumers without alienating the most informed consumers who could react negatively towards the operator if aggressive dark patterns are used. Paradoxically, the consumers are least exposed to dark patterns that are likely to pose a credible threat against the firms that would use them. The more exposed consumers are, the less likely to respond to them, the less they can identify them. Moreover, targeting the vulnerabilities of specific segments of consumers increases the effectiveness of manipulation to the detriment of the most vulnerable consumers.

It is, therefore, a question of forcing firms to be accountable for algorithms despite their opacity (black box logic) through regulation by the spotlight that can raise reputational issues (encouraging the development of responsible A.I.) and of encouraging the emergence of algorithmic combats that allow consumers to exercise countervailing competitive power.

However, three issues remain to be considered. The first is how to make algorithms accountable in a model where opacification is voluntary but is consubstantial with the technology (e.g. in the case of deep learning). The second relates to the asymmetry of performance between algorithms

and counter-algorithms concerning the capacities accumulated by the dominant operators in terms of A.I. and information processing capacities, especially from the perspective of quantum computing development. The third relates to the ethical dimensions of using algorithms to detect and sanction practices on the one hand and the use of counter-algorithms on the other. On the first aspect, one of the relevant issues is considering an underlying model of what the market should be to characterize a market strategy as abnormal. On the second aspect, one of the dimensions to be considered is the unequal access of consumers to countermeasures and the potential reinforcement of inequalities that may result.

This last dimension may open up to a broader questioning of the differences that may increase from one consumer to another due to increasing recourse to algorithmic decisions on the part of both firms and consumers themselves. The digital economy and the discrimination strategies it facilitates may increase inequalities between captive consumers and consumers opting for multi-homing strategies, between "informed" and "naïve" consumers, and finally between consumers capable of implementing technical devices that allow them to exercise not compensatory market power, but a technical capacity to correct the algorithmic strategies of firms and others.

Finally, in the case of interfaces and manipulative algorithms by design (dark patterns), it should be emphasized that consumers may be led to attribute the fault to their behaviour and not to a strategy implemented by the firm if they even manage to identify ex-post a "damage" and make the link between their past choices and this damage. However, as we show in our fifth section, the damage caused by algorithms in the field of consumer and competition protection can be particularly significant when they involve "high stake decisions.").

5. WHICH FRAMEWORK FOR HIGH-STAKE DECISIONS?

Algorithms are having an unprecedented impact on the lives of businesses and consumers. The computer science literature pays particular attention to the issues of highly consequential decisions (Buolamwini 2018; Citron and Pasquale 2014; Kleinberg et al. 2018; O'Neil 2016; Parkes, Vohra et al., 2019). We propose considering them concerning competition rules or rules applicable to financial markets' supervision (5.1) and then about its various stakeholders (5.2).

5.1 Competition law, capital market law and high-stake decisions

In competition law, markets often involve very far-reaching decisions that are mainly taken by human decision-makers. The keyword here is "high-stake decisions." The characteristic is the pace of these decisions. Competition law has been built around high-stake decisions and the pace of decisions taken on markets by humans. The objective is to guarantee a level playing field in a competitive configuration in which firms' decisions are based on data analysis and are increasingly automated without human intervention in the decision-making circuit via machine learning algorithms.

The increasing availability of data combined with strong computation power offers an unprecedented opportunity to develop machine-learning models (ML). These ML models can be used in two ways. First, A.I. can be used in an "augmented intelligence" perspective (de Marcellis-Warin, Munoz, and Warin 2020). These ML models can help human decision-makers make better decisions. Second, A.I. can be used as automated intelligence: these ML models make decisions that humans have traditionally made. (Citron and Pasquale, 2014; O'Neil, 2016). Concisely, ML models can assist or even replace human beings in the decision-making process, depending on their autonomy level. These algorithms can play the market dynamics at very high speed with a succession of consequent low impacts, except that the long-term impact of these high-frequency market algorithmic manipulations can be very consequential.

However, the applicability of A.I. to the above parameters is limited by some fundamental challenges. First, the above parameters require the design of models that consider fairness and interpretability. However, most existing money laundering models are primarily optimized for forecast accuracy and are not inherently fair or interpretable. Second, the data available in these contexts are often subject to various selection biases. These frameworks are subject to missing counterfactuals, i.e., the data capture only the decisions made by human decision-makers and not the counterfactuals. However, these tools, already used in the supervision of financial activities, can be used in the field of competition when the challenge is to control the pricing policy of certain market actors with personalized and dynamic prices, i.e. abundant and apparently (but only apparently) erratic. Also, ML algorithms can be used for criminal activities. In a paper dated May 2020, Mizuta demonstrated that an ML algorithm could discover market manipulation through learning via an A.I. market simulation, despite no bad intention of the ML algorithm designer. The A.I. discovered autonomously that market manipulation is an optimal investment strategy. This result suggests the need to regulate A.I. in financial markets to prevent artificial intelligence from performing market manipulation. This also questions the liability schema to be applied in such a case because of the potential harm, loss or injury resulting from the use of such emerging

technology. The risks posed by ML algorithms justify a strict liability schema due to the increased risk of harm. The one benefiting from the operation and who is in control of the operation shall be liable. The increased risk of damage, loss or injury to individuals resulting from the use of ML algorithms justifies supervision mechanisms carried out internally on behalf of the firm and its stakeholders (5.2).

5.2 Supervision carried out internally on behalf of the firm itself and its stakeholders

Supervision of corporate governance can be public or private. In many jurisdictions, complaints to the regulator are considered the most effective enforcement mechanism. Public oversight takes time to be implemented for emerging technologies.

Therefore, we propose to consider private supervision to complement public supervision, considering the firms' responsibility and moral consciousness towards its stakeholders.

This is important because self-learning agents and the design of the choices' architecture open to users can harm its users and raise ethical issues. The interface itself can format and channel behaviour and alter, if not construct, users' choices. For Fogg (2002), the design of the algorithm can include seven types of persuasion strategies (reduction, channelling, fine adaptation [tayloring], suggestion, self-censorship, monitoring and conditioning) (Fogg 2002). As we have seen for both sludges and nudges, these persuasive architectures can play in a direction that is favourable to the user or unfavourable (Berdichevsky and Neuenschwander 1999). This result may depend on two phenomena. The first may relate to the desire to manipulate the consumer (Nodder 2013). The second may correspond to a lack of consideration and consideration of the issues themselves. This is an antipattern situation in which the user's damage is involuntary and can be explained by the lack of care in coding (Koenig 1995). Moreover, according to the Internet user himself (his skills, attention, experience, etc.), the same architecture can be manipulative or acceptable. Does the solution involve technical devices such as privacy-enhancing technologies (PETs - techniques that improve users' confidentiality and protect their data), for example, through data encryption that reconciles respect for privacy and collective gains linked to data processing, or through algorithm certification devices?

One of the concrete internal assessment tools available has been developed by the *AI Transparency Institute*. This is the one we adapted to the context of the digital platforms for illustration purposes. The objective of this tool is to create indexes available to companies. According to a set of predefined criteria, these indexes can inform companies offering services or goods via a digital

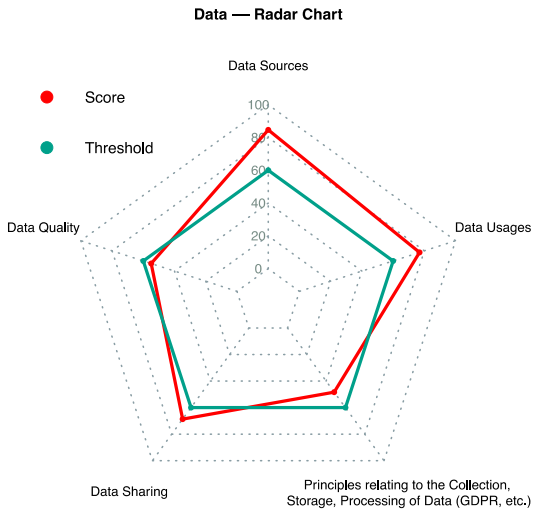
platform of their organization's level of maturity. These criteria are related to data, I.T. security, transparency, explainability of algorithms, ethical values and compliance with legal standards. They take into account the interests of all stakeholders, employees, customers and the community at large (including investors).

These indexes can facilitate obtaining certification or a quality label. They are based on a mathematical model and a list of questions, which are weighted. A trust index and a set of thematic graphical charters emerge from them. The graphical representations presented below give examples based on the online version.

The first concerns the data used and the risks related to them (see Table 4). It reflects the quality of the control implemented by the company concerning: 1) Data sources (data collected with the user's consent, observed from his online behaviour, inferred, observed in the ecosystem, purchased from data brokers, etc.), 2) Data quality (particularly in terms of bias), 3) Data usages, 4) Data sharing with third parties or compliance with specific protection rules such as the RGPD and 5) Principles related to the collection storage. A scalar graph is obtained from the scores obtained on each of the risk criteria (which group together several questions that are subject to a weighted score) and make it possible to compare the firm to a reference chosen as a benchmark or to the average of the firms participating in the program).

The Data – Index: 66.70

Strength: Acceptable



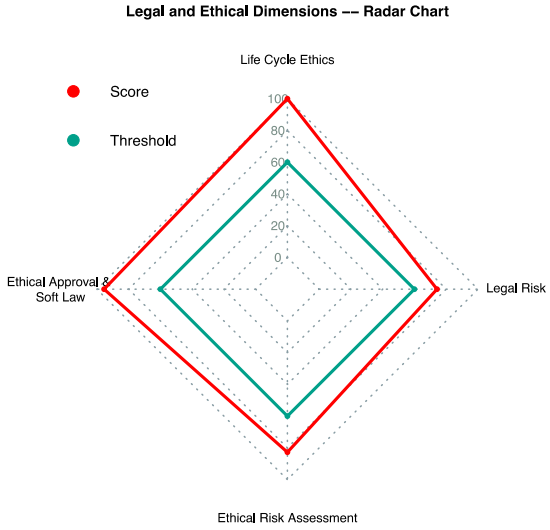
Risk Criteria	Score	Required	Strength
Data Sources	84.67	60.0	Amazing
Data Quality	54.94	60.0	Acceptable
Data Sharing	68.67	60.0	Good
Principles relating to the Collection, Storage, Processing of Data (GDPR etc.)	48.34	60.0	Acceptable
Data Usages	76.88	60.0	Good

Table 4: Scoring based on data sources and treatments

The scores can be broken down into numerous dimensions, as shown in the graphs below, which reflect compliance with ethical commitments or legal requirements, I.T. system security or the training and control of the algorithms used.

Legal & Ethical – Index: 88.11

Strength: Amazing



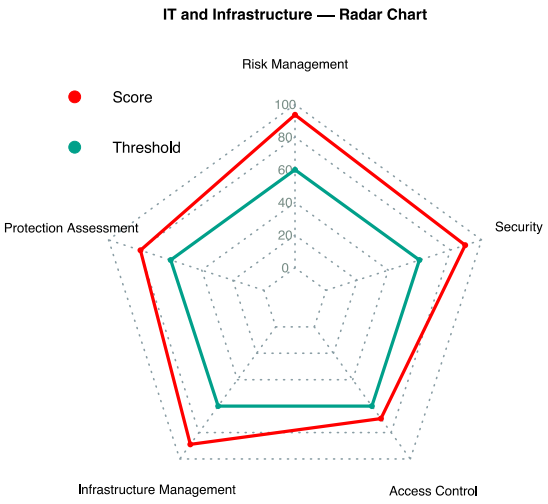
Risk Criteria	Score	Required	Strength
Life Cycle Ethics	56.38	60.0	Acceptable
Ethics Approval & Soft Law	63.97	60.0	Good
Ethical Risk Assessment	77.67	60.0	Good
Legal Risks	89.24	60.0	Amazing

Table 5: Scoring system on Legal and Ethical dimensions

Table 5 illustrates the maturity level of an organization related to its legal and ethical aspects. It considers policies and processes in place during the whole data life cycle related to data protection and ethics, consumer protection, and competition policy. It verifies, in particular, the risk of algorithmic collision and the risk of abuse of a dominant position. In this example, we can deduce from the results that ethics approval and soft Law mechanisms are in place, as well as Ethical risk assessment. Also, legal risks look to be handled appropriately. This result will be confirmed by a review of the legal documentation by an independent expert.

IT & Infrastructure – Index: 84.20

Strength: Amazing



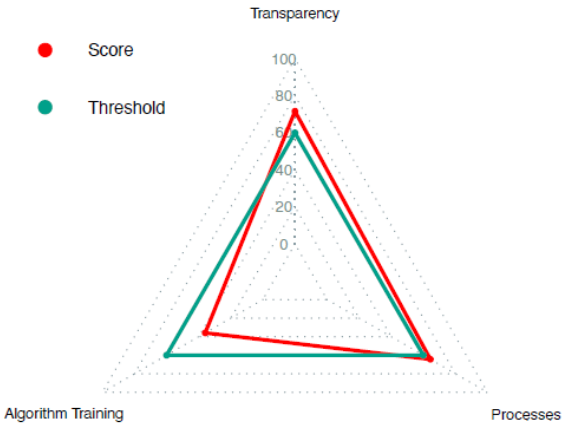
Risk Criteria	Score	Required	Strength
Risk Management	93.67	60.0	Amazing
Protection Assessment	79.44	60.0	Good
Infrastructure Management	88.97	60.0	Amazing
Access Control	69.43	60.0	Good
Security	89.47	60.0	Amazing

Table 6: Scoring system on I.T. & Infrastructure

Table 6 illustrates the risks related to I.T. and Infrastructure as they result from ISO norms. This matrix considers, in particular, infrastructure management, policies in risk management, security and access management. In this example, Risk management policy, Infrastructure management and Security appear as being adequately handled. A comprehensive review of the corporate documentation will confirm this result.

Data Management – Index: 57.33

Strength: Acceptable



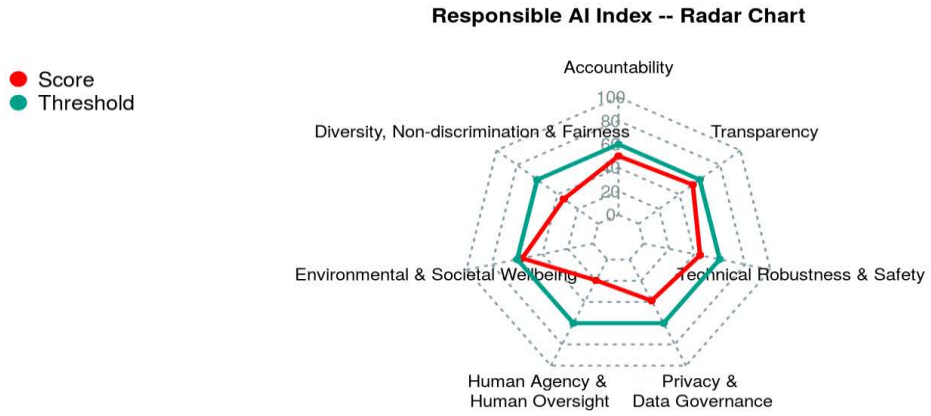
Risk Criteria	Score	Required	Strength
Transparency	71.67	60.0	Good
Algorithm Training	35.93	60.0	Bad
Processes	64.38	60.0	Good

Table 7: Data management score

Table 8 illustrates a scoring system related to Transparency, Algorithm training and Processes. In this example, the company has to use its resources to improve the algorithm training. On the contrary, no resources need to be invested in transparency or traditional Processes.

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

Table 8: Scoring system on Responsible A.I.

In this illustration, the company shall make a global review of its strategy and policy on the Responsible use of data and A.I. internally. It should mainly invest resources in Human Agency and Oversight and improve its policy on diversity, non-discrimination and fairness. Technical robustness and Safety can also be strongly improved.

These trust indexes can have an impact on both business and consumer behaviour. Indeed, the company could identify its shortcomings on this basis and invest its resources in essential issues for its shareholder value and the common good. It will gain efficiency by improving its internal processes and demonstrating its responsible management to all its stakeholders. The company increases confidence in its corporate governance by indicating the scores obtained in its annual report and on its website.

These trust indexes encourage companies to behave diligently since they can be used as an element in their governance's financial rating (e.g. through the Ethos Foundation). These indexes will be accessible to shareholders and have the potential to have a significant influence on the shareholder value of companies by rewarding transparent and diligent behaviour. They will show how companies position themselves vis-à-vis their competitors. They will encourage adopting similar behaviours in the same market and thus improve their score for responsible and transparent behaviour towards consumers and the community. In other words, we believe that if we make visible what the companies achieve, we create an effective incentive for socially responsible corporate behaviour.

Conversely, the consumer has the potential to obtain important information about the degree of trust he or she can place in an economic actor such as an online platform through the publication of these indices on the company's website or in its annual report. These indices will increase the power of control of consumers. While they bring undeniable value to consumers and companies, they also offer increased transparency to shareholders regarding corporate governance. Indeed, they make visible parameters that are difficult to access. While they are based on a voluntary ex-ante audit of organizations, they can also be used by independent administrative authorities in competition law for ex-post evaluation and auditing purposes. Finally, it remains to be determined whether this trust index solution is sufficient to give consumers full confidence in online platforms and services and products incorporating artificial intelligence. Supported by the regulator and investment funds, this self-regulatory approach appears promising, even more so as companies' awareness of social and environmental responsibility is increasing (Townsend, 2020).

Legislative intervention in competition law and the law of obligations are likely to be indispensable to change online platforms' behaviour sustainably. The stakes are immense (Parkes, Vohra, and participants 2019). It is not excluded that companies' general obligation to demonstrate dynamic behaviour towards the climate, customers, investors and the wider community may need to be pronounced. Online platforms contribute to sustainable development issues and should voluntarily

integrate social, environmental, and economic concerns into their stakeholders' activities and interactions. The ISO 26000 standard creates a reference standard in this sense.

In France, Articles L 225-102-1 of the Commercial Code require companies to publish a performance statement that replaces the corporate social responsibility report. This is a tool for steering corporate strategy, supplemented by an index to quantify artificial intelligence algorithms' confidence. These articles are supplemented by the law of March 27, 2017, relating to parent companies' duty of vigilance. This law aims to ensure the respect of human rights by multinationals. A vigilance plan must be published to prevent environmental and human rights risks and corruption in their activities and those of their subsidiaries, subcontractors and suppliers, in France and abroad. These initiatives are commendable but limited to the national territory.

The trust indexes have a global reach. They are intended to be applied outside the territory of the nation-state and can assist transnational players in initiating a beneficial and sustainable digital transformation in society's interest. They are likely to play a crucial role in spreading a responsible culture for online platforms. Used by regulators, these tools have the potential to foster international cooperation, which is essential to strengthen international standards, where necessary, and to ensure their uniform application in order to protect against negative cross-border, regional, and global externalities that affect the digital economy. A general and global obligation to publish an annual report presenting the company's effective measures to behave as a responsible actor appears to be an indispensable step. In the absence of multilateral sanctions, the financial markets' pressure appears to be an interesting lever to be examined thanks to the trust and digital responsibility indexes.

6. CONCLUSION

This article considered the available supervision tools for the authorities in charge of market surveillance, the consumers, or the companies' stakeholders. This is a crucial topic because the increasing autonomy of algorithms can lead to the damageable behaviour of algorithms powered by Artificial Intelligence. In particular, learning models can provoke algorithmic collusion and abuse of dominant position. In November and December 2020, the E.U. Commission has published several regulations related to data management and to digital markets: the DGA (data governance act, November 25, 2020[13]) and the DGA (Digital Services Act) and the DMA (Digital Markets Act) that are expected December 15, 2020. As the British initiative of November 27,

[13] https://ec.europa.eu/digital-single-market/en/news/data-governance-act

2020, these regulations lead to specific online platforms' ex-ante regulation. We consider that the E.U. Commission could use the scoring tool presented in this paper as a possible new competition tool. It could also be used ex-ante by corporations to assess their maturity level on critical criteria. This proposal appears meaningful because ex-post detection and prosecution seem rather challenging. Criteria of this alternative solution will have to evolve with the development of technology. Its main benefit comes from its possible implementation to transnational players in initiating a beneficial and sustainable digital transformation in all stakeholders' interests. Adopting this approach by the regulators and by investment funds could act as a catalyst and accelerate A.I. actors' supervision.

References

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. 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, Alessandro, Leslie K. John, and George Loewenstein. 2013. « What Is Privacy Worth? » *The Journal of Legal Studies* 42(2): 249-74.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 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 (November 25, 2020).

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

Becerra, Xavier. 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 (November 18, 2020).

Berdichevsky, Daniel and Erik Neuenschwander. 1999. « Toward an ethics of persuasive technology .» *Communications of the ACM* 42(5): 51–58.

Brandimarte, Laura, Alessandro Acquisti, and George Loewenstein. 2013. « Misplaced Confidences: Privacy and the Control Paradox .» *Social Psychological and Personality Science* 4(3): 340-47.

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

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

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello. 2019. « Algorithmic Pricing What Implications for Competition Policy? » *Review of Industrial Organization* 55(1): 155-71.

Chopra, Rohit. 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 (November 18, 2020).

Citron, Danielle Keats, and Frank Pasquale. 2014. « The Scored Society: Due Process for Automated Predictions .» *Washington Law Review* 89: 1.

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

Fogg, B. J. 2002. *Persuasive Technology: Using Computers to Change What We Think and Do*. $ {nombre}er édition. Amsterdam ; Boston: Morgan Kaufmann.

FTC. 2020. « Zoom Video Communications, Inc., In the Matter Of .» *Federal Trade Commission*. https://www.ftc.gov/enforcement/cases-proceedings/192-3167/zoom-video-communications-inc-matter (4 décembre 2020).

Gray, Colin 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, Jon D., and Douglas A. Kysar. 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).

Judiciary Committee. 2020. *Investigation Of Competition In Digital Markets: Majority Staff Report And Recommendations*. S.l.: Nimble Books.

Kahneman, Daniel. 2011. « Thinking, Fast and Slow .» *PenguinRandomhouse.com*. https://www.penguinrandomhouse.com/books/89308/thinking-fast-and-slow-by-daniel-kahneman/ (4 décembre 2020).

Kleinberg, Jon 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.*

Kunreuther, Howard, and Ralph Ginsberg. 1978. *Disaster Insurance Protection: Public Policy Lessons*. Wiley.

Luguri, Jamie, and Lior Strahilevitz. 2019. « Shining a Light on Dark Patterns .» *SSRN Electronic Journal*. https://www.ssrn.com/abstract=3431205 (4 décembre 2020).

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

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

de Marcellis-Warin, Nathalie, and Thierry Warin. 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, Alain, Antonio Nicita, and Giovanni Battista Ramello. 2020. « Big Data and Big Techs: Understanding the Value of Information in Platform Capitalism .» *European Journal of Law and Economics*. https://doi.org/10.1007/s10657-020-09675-1 (4 décembre 2020).

Marty, Frédéric. 2017. « Algorithmes de prix, intelligence artificielle et équilibres collusifs ». *Revue internationale de droit economique* t. XXXI(2): 83-116.

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

Marty, Frédéric, and Thierry Warin. 2020a. « Concurrence et Innovation Dans Les Écosystèmes Numériques à l'ère de l'intelligence Artificielle ». *Concurrences / Competition Law Review* 1: 36-41.

———. 2020b. « Innovation in Digital Ecosystems: Challenges and Questions for Competition Policy .» *CIRANO Working Paper Series*. https://cirano.qc.ca/fr/sommaires/2020s-10.

———. 2020c. « Keystone Players and Complementors: An Innovation Perspective .» *CIRANO Working Paper Series* 2020s-61. https://cirano.qc.ca/fr/sommaires/2020s-61.

———. 2020d. « The Use of A.I. by Online Intermediation Platforms .» *Delphi - Interdisciplinary Review of Emerging Technologies* 2(4): 217-25.

———. 2020e. « Visa Acquiring Plaid: A Tartan over a Killer Acquisition? Reflections on the risks of harming competition through the acquisition of startups within digital ecosystems ». *CIRANO Working Paper Series* 2020s-62. https://cirano.qc.ca/en/summaries/2020s-62.

Mathur, Arunesh 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, Deirdre K., Priscilla M. Regan, and Jennifer King. 2020. « The Fertile Dark Matter of Privacy Takes on the Dark Patterns of Surveillance .» *Journal of Consumer Psychology* 30(4): 767-73.

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

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

O'Donoghue, Ted, and Matthew Rabin. 2015. « Present Bias: Lessons Learned and to Be Learned .» *American Economic Review* 105(5): 273-79.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

Parkes, David C., Rakesh V. Vohra, and other workshop participants. 2019. « Algorithmic and Economic Perspectives on Fairness .» *arXiv:1909.05282 [cs]*. http://arxiv.org/abs/1909.05282 (September 21, 2020).

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

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

Sunstein, Cass. 2019. « Sludge and Ordeals .» *Duke Law Journal* 68(8): 1843-83.

Sunstein, Cass R., 2020. « Sludge Audits .» *Behavioural Public Policy*: 1-20.

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

Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth, and Happiness.* Updated édition. New York: Penguin Books.

Townsend, Blaine. 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, Thierry, and Daniel Leiter. 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, Thierry, and Antoine Troadec. 2016. « Price Strategies in a Big Data World .» *Encyclopedia of E-Commerce Development, Implementation, and Management*: 625-38.