CATÓLICA LAW REVIEW

DIREITO PRIVADO

Caroline Jonnaert

Federico Galli

Judith Martins-Costa

Ljupcho Grozdanovski

Manuel Carneiro da Frada

Pedro Oliveira

Vibe Ulfbeck

Paulo Olavo Cunha

António Agostinho Guedes

AI and consumer manipulation: What is the role of EU fair marketing law?*

Federico Galli**

Ph.D. researcher in Law, Science & Technology at Law Department (CIRSFID), University of Bologna; Ph.D. researcher in Computer Science at Computer Science and Communication Department (CSC), University of Luxembourg

SUMMARY

---

1. Introduction

In 2014, American legal scholar Ryan Calo warned against the perils of "digital market manipulation", defined as "the specific set of emerging technologies and techniques that will empower corporations to discover and exploit the limits of each consumer's ability to pursue his or her own self-interest" (Calo, 2014). Five years later, Calo's prophecy is seemingly coming true. Reports and studies in different fields of research show how big tech companies can employ different Machine Learning techniques to predict psychological traits and emotional states of their users and use this knowledge to shape the virtual and legal environment where they make purchase decisions (Burr *et al*., 2018; Burr and Cristianini, 2019). Such developments represent a threat to the autonomy and freedom of individuals (Mik, 2016) and constitute one of the challenges for EU consumer law in the algorithmic era (Jabłonowska *et al*., 2018).

Against this background, this paper conceptualises two ways in which AI fundamentally increases the risk of manipulation for consumers and considers how these developments question established EU fair marketing law.

The work is divided into two parts. The first part clarifies the concept of 'AI' and 'consumer manipulation' and describes two ways in which AI is reshaping our traditional comprehension of manipulative marketing. One relates to the new pool of personal, fine-grained, marketing-sensitive information that businesses can extract from consumers' data thanks to ever more refined predictive techniques. The other relies on the power of AI systems to instantly and continuously learn from consumers' feedback in order to optimise marketing tasks. The second part questions the role of current EU fair marketing law in protecting consumers against 'AI-powered' manipulative commercial practices. Here, an account will be given of the virtues and the vices of the Unfair Commercial Practice Directive, the prime European legal tool dealing with distortive and manipulative commercial practices. Subsequently, the recent US legislative proposal of the Deceptive Experience To Online Users Reduction Act will be reviewed, as it may represent an example of how fair marketing law can regulate the new sophisticated tools of technological influence. Conclusive takeaways will follow.

2. "Artificial Intelligence" and "consumer manipulation"

There can be several ways to start a discourse on Artificial Intelligence and consumer manipulation. The topic is extensive and confining it to an article is not an easy task. Besides, the two concepts alone are already ambiguous and vague by themselves and do not certainly facilitate defining tasks. However,

since this paper tackles the problem with a regulatory focus in mind, the best way to proceed is to have at least a clear understanding of the subject of discussion – without any ambition of completeness.

As always, one would be tempted to give a general definition of AI. As intriguing as it sounds, the author believes this exercise is fundamentally wrong. Broad definitions can lead to generalisation and loss of nuance. Moreover, one-size-fits-all definition of AI is probably not needed, at least for purposes of legal research and regulation: AI is not the same thing when it comes to self-driving cars or autonomous weapons or personalised advertising. Instead, it is far more critical to have a clear understanding of what we mean when we refer to AI, in our specific problem-driven context. This work follows the idea that AI should not be merely seen as a technology but as a complex socio-technical phenomenon (Balkin, 2015, p. 48). Rather than focusing on AI as such, one should consider how people interact with AI and how people interact with other people using AI. Hence, in this paper, we will focus on how companies employing AI interact with consumers, thereby increasing the risk of manipulative outcomes. In such a context, 'AI' is understood simply as *algorithms that learn from experience to optimise specific tasks that are specifically related to the management of marketing and consumers relationships*.

As a further point, it is well known that the issue of digital manipulation has a wide echo. The infamous scandal of Cambridge Analytica has put digital intermediaries into the public spotlight – for the first time in such a resounding way – revealing their ability to shake the very democratic foundations of our societies by subtly and hiddenly influencing our beliefs and political opinions. Researchers, including legal scholars, are increasingly paying attention to this phenomenon inspired by protecting and enhancing citizens' cognitive freedom vis-à-vis the new gatekeepers of information (Tucker *et al*., 2018; Harambam *et al*., 2018). Nevertheless, in this paper, we look at the effect of AI only with regard to the 'commercial side' of manipulation, that is, the somehow unfair distortion of purely economic behaviours of consumers. For this reason, we will just show how AI can affect the deliberative process of digital consumers in connection with contract-related decisions, which have an intrinsic economic value (e.g., clicking on online ads, consenting to the processing of personal data, purchasing a product or subscribing to a service, deactivating an user account, etc.). This choice is deliberate. As we will see in the next sections, EU fair marketing law, which is the benchmark of the analysis, has always had a particular interest in protecting consumers' economic interests against undue forms of commercial manipulation.

When does a commercial practice manipulate consumers? In marketing, there is always a thin line between *persuasion*, typically designated as the

original mission of advertising, and *manipulation*, which is intuitively perceived as unfair. Consequently, it is not clear when the distortion of consumers' economic behaviour – undoubtedly present also in non-manipulative persuasion – becomes a problem for the law. This issue has always received scarce attention from lawyers because of the elusive meaning of 'manipulation' and its difficult digestion in a legal order that defends economic freedom. Today, however, manipulation discourse is back in the forefront. In many fields of social sciences, private and public research has progressively proved the easy-to-manipulate nature of human beings and these findings are today within reach of companies and governments.

According to one of the most prominent jurists who has recently dedicated much attention to the definition of manipulation, "a statement or action is manipulative to the extent that it does not sufficiently engage or appeal to people's capacity for reflective and deliberative choice" (Sunstein, 2016). In fact, it is now widely believed that human beings are guided by two 'cognitive systems' when making decisions. System 1 is fast, automatic, intuitive; System 2 is slower, reflective and deliberative (Kahneman, 2011). System 1 does not necessarily work badly. We often use intuition and heuristics with good results (for example, we can positively use educated guess when looking for the toilette during the night or when driving under uncertain circumstances), while reflection and deliberation can lead to mistakes (using deliberation in case of multiplication of high figures may often lead to errors). Nevertheless, the problem with System 1 is that it is prone to several types of common biases, e.g., optimism bias (the tendency to be too optimistic), present bias (the tendency to give more weight to payoffs that are closer in time), anchoring (the tendency to rely heavily on a specific trait or piece of information when making decisions), framing effect (drawing different conclusions from the same information depending on how that information is presented). These biases can be easily anticipated, and the use of proper choice architecture (so-called 'nudge') can profoundly influence the course of individuals' decision-making (Thaler and Sunstein, 2009). Following Sunstein's definition of manipulation, manipulators target System 1 (exploiting bias) and prevent the recipient of manipulation from engaging with System 2. For example, graphic warnings on packets of cigarettes aim to exploit people's attentional bias, i.e., the tendency of perception to be influenced by recurring thoughts (the more you look to graphic warnings, the more you think about quitting smoking); an insurance company that presents its advertising by describing a horrible situation in which insurance would be very useful leverages our framing effect bias, etc.

Sunstein notes, however, that not all cases of manipulation are ethically reprehensible; this will depend on the perspective adopted. If we take a deontological

stance, manipulation always fails to respect people's dignity and autonomy. From a welfarist point of view, to which only those who choose know what is best for them, manipulation is not necessarily evil. Some people may benefit from being manipulated (e.g., a smoker who desperately wants to quit). However, even for welfare reasons, manipulation becomes ethically reprehensible and may justify an appropriate normative response when (a) the manipulator's motives become more self-interested or venal and (b) their efforts to circumvent people's deliberative capacities become more successful. Such a delineation of ethically (and possibly legally) reprehensible manipulation is not black-and-white, but incremental and nuanced (Sunstein speaks of 'fifty shades of manipulation'): the more the manipulator's motive is self-serving and lucrative, and the more significant the effectiveness of the manipulative attempt, the more reprehensible the manipulation will be, and legal intervention will be justified.

Thus understood, we can see the relationship between the concepts of 'AI' and 'consumer manipulation', namely how AI increases the risks of manipulation, thereby justifying a legal intervention.

Following Sustein's definition, first of all, the manipulator's interest must be lucrative and self-serving. We can hardly think of private companies that use AI for the sole benefit of consumers. They are companies, and they are driven by profit. Their motive is lucrative and self-serving by definition. Nevertheless, this does not necessarily imply that the motive is also hateful and obnoxious. There may be situations where the interests of the company and the consumer converge. Think, for example, of the company's algorithm that targets the consumer with a personalised ad that reflects his or her true preferences, by appealing to their attentional bias (e.g., by using a particular colour, font, by leveraging a certain claim or nudge). In that case, the system might manipulate the consumer, but it would do so for the consumer's own well-being. (However, this would be hardly sustainable from a welfarist point of view, because only the consumer knows what is best for him or her, but it certainly makes manipulation less reprehensible.) However, the opposite case can also be envisaged, where the interests of the company and those of the consumer diverge. Imagine the case where the company's algorithm has to choose between a free video and a paid video while calculating the expected revenue that could come from different advertisements associated with different videos. In such a scenario, if the system shows the paid video to the consumer, the companies act in their own interest and against the consumer's interest. There may be different scenarios in which the company's interest and the consumer's interest concur or compete, which would give different degrees of meaning to the 'lucrative and self-interest motive' requirement. However, one can generalise by saying that a company has a lucrative and self-interested motive when AI systems can choose between the

interests of the company and those of the consumer and go for the former and not the latter.

At this point, the reader can understand how AI increases the effectiveness of companies' attempts to circumvent consumers' deliberative capabilities.

3. Predictive powers

A prediction is an estimate of the probability that something will happen in the future based on experience. Prediction is at the core of managing customers relationships; companies continuously have to predict events or values such as return on investments in marketing, risk assessment, customers' preferences, customer budget constraint, customer retention, customer conversion, etc. Today, many of these tasks can be optimised through predictive analytics.

Predictive analytics is a field at the intersection of data mining, statistical modelling and machine learning that aims to build models that make predictions based on patterns extracted from historical data (Kelleher *et al*., 2015). While traditional business analytics has been able to build predictive models that consider only a static and finite amount of data, the employment of machine learning algorithms for data and predictive analytics allows to respond to new data and improve the granularity and accuracy of prediction over time. The peculiarity of learning algorithms is that they are able to infer rules of future behaviour directly from data without being explicitly programmed. There can be different machine learning models to perform this prediction (e.g., decision tree, regression, classification, etc.), either characterised by the need of a phase of previous training by a human (so-called supervised learning) or not (so-called unsupervised learning). The choice of the technique will depend on the task to be performed – for predictive analytics, the most frequently used are decision tree, logistic and linear regression and neural networks.

A rather well-established practice in business intelligence is to use predictive analytics to estimate customers' preferences in order to anticipate their desires and needs: for example, sequential purchase of customers (Goel *et al*., 2010), improving targeted advertising (Richardson *et al*., 2007), understanding brand perception (Culotta and Cutler, 2016). In recent years, however, studies have shown how machine learning techniques can also be employed to predict more 'sensitive' aspects of consumers' behaviour. For example, researchers have demonstrated the ability to use social media profiles (Quercia *et al*., 2011; Kosinski *et al*., 2013), blogs (Yarkoni, 2010) or language use (Schwartz *et al*., 2016) to accurately infer users' psychological traits, such as the 'Big 5', customer's values, degree of deliberation, intelligence and risk aversion, etc. Other studies

have proved that it is possible to predict users' psychological transient states, such as mood and emotions, attention and stress, by analysing real-time spoken and written language (D'mello and Kory, 2015; Hu and Flaxman, 2018) and video (Teixeira *et al.*, 2012). There is a whole emerging field of research and business applications called 'affective computing' whose purpose is to create software that recognises and processes human emotions (Cambria, 2016).

Although expressed in the form of forecasts and not facts, this new wealth of knowledge is of extreme value to businesses. Predictions can be used as input for different kinds of decision-making tasks: from determining what piece of advertising to display to single consumers (Roffo and Vinciarelli, 2016; Matz *et al.*, 2017) to optimising entertainment or content relevance through the use of recommendation systems (Ning *et al.*, 2019); from personalising commercial offers and the content of contracts (Hacker, 2017) to customising website and apps graphics, so-called 'morphing' (Hauser *et al.*, 2009). Companies are able to take advantage of this sensitive information to customise any aspect of the commercial interaction with consumers, both 'virtual' (such as the 'look and feel' of the web page, the layout, the page rank) and 'real' (advertising specific products, offering tailor-made services, setting customised prices, personalising contract terms of service).

It is not only a common intuition but also a scientific fact that there is a strong relationship between successful marketing and the ability of businesses to leverage the personality and emotions of consumers (Odekerken-Schröder *et al.*, 2003; Hirsh *et al.*, 2012; Bagozzi *et al.*, 1999). However, effective marketing strategies do not imply that consumers are being manipulated. In fact, some experiments show how personalising the website according to consumers' personality traits increases their happiness and engagement (Matz *et al.*, 2016). While this is true, it is also true that there may be cases where certain psychological traits expose consumers to increased vulnerability to marketing activities (e.g., agreeableness, neuroticism, risk aversion, subjective well-being, etc.). Moreover, when it comes to temporary moods or emotional states, advocates of a deontological account of manipulation would argue that consumers must be treated with dignity and that exploiting their non-rational being for commercial purposes is inherently unfair. This thesis is deeply rooted in our conception of modern legality, in which individuals are autonomous and rational beings and must be treated in this way. Even if we were to take a welfarist stance, we would end up saying that exploiting psychological weaknesses or emotions *without revealing it* (as it would generally happen if exploitation takes place through an algorithm) diminishes the ability of consumers to deliberate autonomously and freely on what is best for them, thus increasing the likelihood of manipulation.

Increased AI-driven predictive powers enable businesses to gain a precious, marketing-sensitive knowledge on consumers that has an inherent potential to bypass consumers' deliberative system, thereby magnifying the risks of manipulation.

4. Loops of feedback and learning

So far, we have seen how AI *quantitatively* increases the risk of consumer manipulation, i.e., how it broadens and deepens the pool of information from which companies can draw from to persuade consumers to make certain choices. Now, we will explain how AI also *qualitatively*, by its nature, strengthens the possibility for businesses to distort consumers' behaviour. For this purpose, we must think of the relationship between a company and a consumer as a dynamic learning process mediated by an algorithm (Burr *et al*., 2018).

Let us imagine the situation where a company wants a consumer to buy its products and therefore needs to establish what kind of advertising the consumer is more likely to click on. The company's algorithm will choose from a set of available actions (i.e., a different set of ads) by leveraging what it knows about the consumer. A statistical model is created that correlates each possible action with the consumer's click probability. The model can be multifactorial and take into account the price (how much the consumer is likely to spend), the context, the layout, the colour and the font of the ad, and possibly even the timing of the advertising. The algorithm can also be designed to consider different strategies to present the choice to the consumer using most popular nudges (e.g., defaults, framing, simplification, rankings, etc.).

Once the algorithm chooses one of the possible actions, the consumer reacts by clicking or not. This feedback enables the algorithm to determine and learn from the consequences of its action and serves to update the function to achieve its goal. The process is repeated iteratively, in a trial-and-error exercise (i.e., each time the consumer is targeted with a different piece of advertising) until the consumer clicks and the algorithm learns the setting that led to the desired result.

This process, called 'feedback loops' in behavioural theory, is basically what lies at the basis of many ML tools of large companies applying 'reinforcement learning' (RL) techniques. A system based on RL decides according to experience the sequence of actions to perform in an uncertain and complex environment in order to achieve some goals (François-Lavet *et al*., 2018). The company must set the goal and, using rewards and punishments as proxies of a desired or an undesirable action, indicates how the algorithm will learn to achieve it. This

suggests that the company does not necessarily require a complete *a priori* knowledge about consumers in order to choose the best marketing strategy, but it can also learn *ex-post*, from individual interactions with consumers.

Given also the increasing use of RL-based marketing applications, AI is going to influence the way of doing business profoundly.[1] By relying on pattern recognition and prediction, digital companies will take advantage of the new machine-based way of experimenting with consumers. Unlike data analysis, which only expresses probability, experiments allow businesses to understand the causal relationships between company marketing decisions and consumers' behaviour, which can then be modelled to achieve the desired result (Varian, 2010, 2018). This logic ('set the goal and let the machine do the rest') has the potential to alter our understanding of consumer manipulation dramatically. Despite everything, we are accustomed to think about manipulation tactics just as an *attempt* to circumvent people deliberative capacities.

The more refined the tactic, the more plausible its success, but there will always be a degree of uncertainty. Many people are sensible and cunning to commercial manipulation, and those that are not can be trained to become more careful and suspicious. To a certain extent, educating people is what can also be done in cases of vigorous attempts at manipulation through AI, where companies can exploit the knowledge of our deepest selves. However, with AI applications mediating business-consumers and learning instantly and continuously from experience, manipulation becomes an *act of behavioural engineering* where companies only need to master the art of reinforcements and punishments that can reliably produce the specific behaviour that the company selects. This implies that manipulation becomes a structural feature of a commercial practice (such as marketing) mediated by AI that no longer alters consumer behaviour but creates, shapes, and precisely engineers wanted behaviours through skilful coding. Shoshana Zuboff calls this 'the reality business', a commercial logic centred on the "knowledge about real-time behaviour that creates opportunities to intervene in and modify behaviour for-profit" (Zuboff, 2015).

Predictive powers and loops of feedback and learning are two sides of the same phenomenon: the ability of businesses to increasingly automate tasks for the management of customer relations – both in the observing-predicting phase and in the decisional phase – through a continuous process of learning.

---

1   George Karapalidis, *Three examples of how reinforcement learning could revolutionise digital marketing*, accessible at <https://econsultancy.com/reinforcement-learning-revolutionise-digital-marketing-case-studies/>; Sourish Dey, *Leveraging Power of Reinforcement learning in Digital Marketing*, accessible at <https://towardsdatascience.com/leveraging-power-of-reinforcement-learning-in-digital-marketing-d373c88a39ab>.

This paradigmatic shift is changing the very essence of marketing. The latter is defined by American Marketing Association as "the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large". In the era of AI, marketing is increasingly becoming *the activity of preventive design of future customer interactions through the use of AI processes*.[2] As a result, manipulation becomes *the ability of companies to design AI-driven processes to guide customers towards a predetermined outcome mechanically*.

5. The role of EU fair marketing law?

The capacity of intelligent systems to increase manipulative outcomes (not only) for consumers is accounted as one of the ethical/legal challenges to establish good governance of AI (Ebers, 2019; Mazzini, 2019). To this latter end, many countries and international organisations have recently developed strategies or guidelines on the development, the deployment and the use of AI systems and their impact on society.

In April 2018, the European Commission presented the Communication on 'Artificial Intelligence for Europe', the first comprehensive policy position specifically focused on Artificial Intelligence.[3] One of the building blocks of the European strategy is to ensure an appropriate ethical and legal framework for the development and use of AI. Among others, a key aspect is to avoid that AI harms citizens and consumers and ensure that "*large-scale use of AI-enabled tools in business-to-consumer transactions ... (are) fair, transparent and compliant with consumer legislation*".

The Communication also established a group of experts to work on ethical guidelines for AI. Last December, the High-Level Expert Group on Artificial Intelligence published the 'Ethical Guidelines for a Trustworthy AI'.[4] According to the latter, AI systems should be developed by treating individuals "*as moral subjects, rather than merely as objects to be sifted, sorted, scored, herded, conditioned or manipulated*". Additionally, government and non-governmental organisations must also mitigate the capability of AI systems to generate "*(in)direct illegitimate*

---

2   Floridi (2019) argues that "marketing sees and uses people as interfaces. Its goal is to identify the most efficient and effective ways to use the human interfaces so as to obtain what the interfaces give access to which, in politics, is represented, in logical order, by attention, consent, and vote". In our context, we could easily replace 'politics' with 'commercial relationships' and 'vote' with 'money'.

3   European Commission, Communication 'Artificial Intelligence for Europe', COM (2018) 237 final.

4   The European Commission's High-Level Expert Group on Trustworthy AI, Draft 'Ethics Guidelines for Trustworthy AI', Working Document for stakeholders' consultation, Brussels, December 18, 2018, https://ec.europa.eu/digital-single-market/en/news/ draft-ethics-guidelines-trustworthy-ai.

coercion, threats to mental autonomy and mental health, unjustified surveillance, deception and unfair manipulation".

At intergovernmental level, the Council of Europe adopted last February a declaration on "the manipulative capabilities of algorithmic processes".

> «Contemporary machine learning tools have the growing capacity not only to predict choices but also to influence emotions and thoughts and alter an anticipated course of action, sometimes subliminally. The dangers for democratic societies that emanate from the possibility to employ such capacity to manipulate and control not only economic choices but also social and political behaviours, have only recently become apparent. In this context, particular attention should be paid to the significant power that technological advancement confers to those – be they public entities or private actors – who may use such algorithmic tools without adequate democratic oversight or control».[5]

Among the recommendations, the Council encourages Member States to initiate a discussion with a view to "providing guidance on where to draw the line between forms of permissible persuasion and unacceptable manipulation [...] and take appropriate and proportionate measures to ensure that effective legal guarantees are in place against such forms of illegitimate interference."[6]

It is also worth mentioning the position paper published in 2018 by the European Consumer Organisation (BEUC) on 'Automated Decision Making and Artificial Intelligence'[7] which recognises the increased vulnerability of consumers to automated and AI decision-making and the increased risk of being manipulated by businesses into a specific purchasing choice.

From a methodological point of view, all these documents agree on the fact that, before introducing new legislation, policy-makers may need to review the adequacy of existing legislative bodies to explore whether they are out of step with ethical and legal challenges that AI is presenting.

---

5   Council of Europe, Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes, Decl (13/02/2019) 1, 13 February 2019 https://search.coe.int/cm/pages/result_details.aspx?ObjectId=090000168092dd4b.

6   The recommendation specifies that unacceptable manipulation "*may take the form of influence that is subliminal, exploits existing vulnerabilities or cognitive biases, and/or encroaches on the independence and authenticity of individual decision-making*".

7   BEUC, Automated Decision Making and Artificial Intelligence, BEUC Position Paper, 20/06/2018, https://www.beuc.eu/publications/beuc-x-2018-058_automated_decision_making_and_artificial_intelligence.pdf.

In this regard, protecting the freedom of individuals against unfair commercial manipulation is certainly not new to the law. Historically, every national system of the Western legal tradition has developed, to a greater or lesser extent, its own legal framework to ensure fair relations between traders and consumers.[8] In the European Union, since the last decades of the 20th century, this task has been taken up to a substantial part by the Community institutions with the objective to ensure a high level of consumer protection (Howells *et al*., 2006; Keirsbilck, 2011). After a period of negative harmonisation carried out by the European Court of Justice, since the 1980s, the Community has issued several instruments to harmonise different activities related to advertising and marketing[9]: the Misleading Advertising Directive[10] later extended to cover also comparative advertising[11], the Unfair Contract Terms Directive[12], the Data Protection Directive[13], the Price Indication Directive[14], the ECommerce Directive[15]. The culmination of European imprinting on the law of fair marketing trading law was achieved in 2005 with the adoption of the Unfair Commercial Practice Directive,[16] a maximum harmonisation directive providing a general framework for unfair business-to-consumer practices.

The regulatory options have varied over time (substantial versus information-based intervention), especially in recent decades the EU has put at the centre of its agenda the need to empower consumers to protect their interests by making autonomous and informed choices (Weatherill, 2013). This ambition is

---

8   Generally, national fair marketing law protects three sets of interests: (1) the non-economic (e.g., health and safety) and economic (i.e., to some extent misleading and deceptive practices) interests of consumers); (2) the interest of competitors for a fair competition (The interest of competitors is conceptually different from the interest of general competition law that seeks to prevent the abuse of market power; it is more connected to intellectual property law, especially trademarks law and passing off actions), and (3) the interest of the public at large to promote taste and decency and appropriate standard of conducts.

9   Consumer protection law includes a wide and diverse set of regulatory tools and policies that can be classified in distinct (and also largely autonomous) legal areas or subfields. Here, we refer to '*fair marketing law*' as the part of consumer protection law that directly regulates fairness in advertising and marketing.

10   Directive 84/450/EC on the approximation of the laws, regulations and administrative provisions of the Member States concerning misleading advertising, OJ 1984 L250/17.

11   Directive 97/55/EC amending Directive 84/450/EEC concerning misleading advertising so as to include comparative advertising, OJ 1997 L290/18.

12   Directive 93/13/EEC on unfair terms in consumer contracts, OJ 1993 L95/29.

13   Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data OJ 1995 L281/31.

14   Directive 98/6/EC on consumer protection in the indication of the prices of products offered to consumers, OJ 1998 L80/27.

15   Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, OJ 2000 L178/1.

16   Directive 2005/29/EC concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC and Regulation N. 2006/2004, OJ 2005 L149/22.

deeply linked to the project of a competitive internal market. Competition can only be guaranteed if consumers are informed and autonomous in making purchasing decisions because traders who act unfairly will no longer be able to win business away from competitors who play by the rules (Stuyck *et al.*, 2006). In order to have a competitive single market, there must also be a level playing field with regard to national fair market laws, so that both traders and consumers can be more confident in trading and shopping across borders without having to face different regulations. In this context, it is useful to understand that companies using manipulative tactics not only invade the autonomy of consumers (thus causing a possible loss of welfare if consumers would have chosen otherwise) but also contributes to distorting the competitive dynamics of the market.

Therefore, questions arise on the role of existing EU fair marketing law in the development of AI that does not interfere with and distort consumers' autonomous and informed decision-making on their purchase choices.

Since the early days of the Internet, EU policy-makers have always stressed the importance of applying the same consumer protection rules online and offline. This has generally led to a somewhat neutral (or 'non-exceptionalist') approach to EU fair marketing and general consumer protection law in relation to digital phenomena (Brownsword, 2016). For some years, however, there have been significant indications that this trend has reversed. Latest developments (the Geo-Blocking Regulation[17], the Digital Content Directive[18], the Sale of Goods Directive[19], the so-called "Omnibus Consumer Directive"[20]) show an increased focus on digital, data-driven trends and the need to update new rules or provide special rules to regulate them (Twigg-Flesner, 2018).

In the context of this current debate, a serious discussion is needed to determine whether existing EU fair trade legislation, in particular rules preventing distortion of consumers' economic behaviour, can withstand the impact of new forms of AI-based manipulation or whether it should be amended to provide new solutions for consumers.

---

17 Regulation (EU) 2018/302 on addressing unjustified geo-blocking and other forms of discrimination based on customers' nationality, place of residence or place of establishment within the internal market and amending Regulations (EC) No 2006/2004 and (EU) 2017/2394 and Directive 2009/22/EC, OJ 2018 L60I/1.

18 Directive (EU) 2019/770 on certain aspects concerning contracts for the supply of digital content and digital services, OJ 2019 L 136/1.

19 Directive (EU) 2019/771 on certain aspects concerning contracts for the sale of goods, amending Regulation (EU) 2017/2394 and Directive 2009/22/EC, and repealing Directive 1999/44/EC, OJ 2019 L136/28.

20 Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules, *OJ L 328/7*.

6. The Unfair Commercial Practices Directive: virtues and vices

Recently, some contributions (Mik, 2016; Helberger, 2016; Sax *et al*., 2018) have designated the Unfair Commercial Practice Directive as a suitable regulatory candidate to address the problems of 'digital market manipulation'[21].

As mentioned in the previous section, the Directive on unfair commercial practices is the highest achievement of EU fair marketing law. The Directive seeks to provide a European conception of fairness in business-consumer relationships by introducing a general clause on unfair commercial practices (Howells *et al*., 2006, 2017). The UCPD was adopted fourteen years ago, but still today remains a very relevant and debated legal instrument, not only because of its many ambiguities, but also for its implications for business and consumers, and Member States' legislation. The Directive has an impact on national legislation on fair marketing, intersects national contract law, not to mention the complex and multi-level enforcement system, which deeply involves each Member State.

Exactly the vagueness of the text makes the UCPD a very flexible tool that could also accommodate current AI-powered marketing practices. Article 5 broadly defines unfair business-to-consumers practices as those that *(a)* contrary to the requirement of professional diligence *(b)* are able to materially distort the economic behaviour with regard to the product of the average consumer whom it reaches or is addressed.

On the basis of this general clause, the Directive prohibits misleading practices and aggressive practices.[22] More specifically, Article 8 defines as 'aggressive' those practices that "*by harassment, coercion, including the use of physical force, or undue influence, (it) significantly impair(s) or (is/) are likely to significantly impair the average consumer's freedom of choice or conduct with regard to the product and thereby causes him or (is/) are likely to cause him to take a transactional decision that he would not have taken otherwise*". Ultimately, the Directive also provides for an annex containing specific practices considered unfair in all circumstance.

---

21   These works build on the seminal paper by Calo (2014) who defines digital market manipulation as "*the specific set of emerging technologies and techniques that will empower corporations to discover and exploit the limits of each consumer's ability to pursue his or her own self-interest*".

22   A commercial practice is *misleading* if it contains false or untrue information or is likely to deceive the average consumer, even though the information may be correct, and cause them to take a transactional decision they would not have otherwise taken. Examples of such actions include false or deceptive information regarding the existence or nature of the product; the main characteristics of the product (its availability, benefits, risks, composition, geographical origin, results to be expected from its use, etc.); the extent of the trader's commitments (in codes of conduct by which the trader has agreed to be bound), etc. Cf. Article 6 UCPD.

From the debate around the UCPD and its aptitude to meet the challenges of digital manipulation, it appears that both the prohibition of aggressive practices and – alternatively and secondarily – the general clause present interesting starting points for further elaboration.

As far as aggressive practices are concerned, the concept of 'undue influence' seems to fit the discussion (Sax *et al*., 2018). This is described as the exploitation of a position of power to apply pressure and limit the consumers' ability to make an informed decision (Article 2, lit. j). Considering the increased information asymmetry and the unprecedented power of digital business to shape the virtual and legal substance of commercial relationships, it could be argued that companies using data-driven AI techniques for marketing are in a position of power over consumers so that any exploitation of such position could lead to undue influence. As a further element, it is appreciated that the Directive does not require a causal link between the practice and the actual distortion of behaviour («is likely to significantly impair [...] are likely to cause»). This characteristic is of great importance considering AI-powered manipulative schemes that rely on prediction (probability), rather than certainty (causality) (Mik, 2016). More generally, commentators notice that the Directive notes that the UCPD does not exclusively protect the interests of the individual consumer, but also "broader conceptual questions about the kind of information economy we would like to live in, and the values that should shape it", (Helberger, 2016). This note is particularly relevant to the issue of AI-based manipulation, where it is difficult to completely separate the impact of manipulation on individual consumers from its repercussion from a broader societal perspective.

That said, it is also recognised that the same policy that has traditionally characterised the interpretation and application of the UCPD would represent – to the *statu quo* – an obstacle to provide consumers with meaningful protections. These observations are often associated with the poor implementation in consumer policy and law-making (including in the UCPD) of the findings coming from behavioural economics (Trzaskowski, 2018).[23]

The paradigm of the rational consumer.

The UCPD offers flexible protection depending on the addressee of the practice. As a general rule, the unfairness of a practice is assessed from the perspective of the 'average consumer'. With some exceptions, this notion is steadily interpreted by the European Court of Justice as the "reasonably well-informed

---

23   As noted by the Author, there is indeed a close connection between the uptake of behavioural research, both in private and public spaces, technological innovation and AI advancement and the need for a 'behavioural turn' in the policy-making, interpretation and enforcement of the law, including consumer protection law.

and reasonably observant and circumspect". At the same time, however, it has been repeatedly observed that this traditional conception of consumer clashes with the basic findings of behavioural economics, according to which consumers have bounded rationality and limited willpower, are guided by heuristics and their judgement is often biased (Incardona and Poncibo, 2007; Sibony, 2014). This is even truer in the AI landscape, where companies can easily estimate the tendency of each consumer to deviate from rational decision-making, detect and react to emotion.

A strict notion of vulnerability.

Article 5 of the UCPD provides for an alternative benchmark when the commercial practice has the potential to distort the behaviour of consumers considered *vulnerable*. In such cases, however, the Directive shows a narrow understanding of 'vulnerability'. This is demonstrated by the fact that only certain specific categories of consumers who are intrinsically exposed to marketing and contracting are granted with higher protection (mentally or physically ill, minors, and credulous people). This conception has been criticised for being too severe, primarily because vulnerability can be 'situational' as there are many factors that are external to consumers themselves than can cause vulnerability (e.g., lack of knowledge of the language, lack of education to deal with highly-complex commercial sectors, or the use of technologies with which the consumer is not familiar). Secondly, the concept of vulnerability can also be 'relational' and may depend on some consumers' increased susceptibility to specific techniques and marketing practices implemented by businesses (Waddington, 2013).

A limited elaboration of professional diligence.

According to the general clause of Article 5, a commercial practice is unfair only if it is contrary to the requirement of professional diligence. This is defined as the "the standard of special skill and care which a trader may reasonably be expected to exercise towards consumers, commensurate with honest market practice and/or the general principle of good faith in the trader's field of activity" (Article 2, lit. h). Since the early days of the adoption of the UCPD, the notion of professional diligence has had little elaboration, which has reduced its legal relevance (Howells *et al*., 2006). This has led commentators to focus on the second tier of the general clause ('the material distortion requirement'). Arguably, however, a serious discussion on new technology-mediated manipulation and its legitimacy should also reconsider the role of a European concept of professional diligence in the field of AI (Trzaskowski, 2018).

The incompatibility between the new marketing trends and the UCPD is even clearer if we consider two other developments that are – so to speak

– 'AI-driven' and that are incoherent not simply with the substantive policy of the UCPD, but also with the very architecture of the Directive.

Personalisation of harm.

As mentioned in Section 3, AI techniques allow today for the personalisation of any aspect of business-consumer interactions. While it is possible to custom-ise websites and advertising to increase consumers' engagement and welfare, companies can also adopt the persuasion strategy for each individual consumer that best fits their psychographic or emotional profile (Kaptein *et al*., 2015). It follows that what is harmful to one specific consumer can be harmful to another. In this context of immanent 'personalised harm', a particular consumer may not be vulnerable and be perfectly circumspect and well-informed, but in the face of such invasive and powerful business practices, he or she may become vulnera-ble at any time during the day. This implies that the dichotomy between rational consumers and vulnerable consumers no longer reflects the reality of marketing, and that vulnerability does not necessarily pre-exist in the commercial relation-ship, but often becomes a by-product of the interaction itself (Calo, 2014). That is why fair marketing law should intervene not only when a consumer is already vulnerable and firms are taking advantage, but also – and indeed even more so – when the company leverages what it knows about the consumer to purposefully make him or her vulnerable.

Opacity and oversight.

Together with the intrinsic opaque essence of manipulation (If manipulation circumvents our deliberative system, it may be the case that we do not realise we are being manipulated), the fact that attempts at manipulation are made through the use of AI generates additional layers of opacity, namely (1) the cor-porate secrecy on the algorithms; (2) the technical illiteracy of consumers over AI processes; (3) the intrinsic lack of transparency of some ML algorithms (Burrell, 2016). This multi-faceted opacity causes serious problems to the current gov-ernance system of the UCPD. The latter was designed with legislators taking into account 'offline' commercial practices. The control over traditional marketing practices has always relied on more or less extensive degrees of publicity (think of advertising made through street billboards, press, radio and TV). Hence the prevalence in many Member States of a public system of oversight and enforce-ment in case of manipulative business conducts. In the AI era, however, where manipulation is tailor-made for every single consumer, only the latter is able to recognise the unfairness of the practice. Yet, the single consumer may not even realise that he or she has been manipulated or when he or she realises, he or she may fail to undertake action because often the cost of acting is simply much

higher than the cost of ignoring the manipulative trick. For this reason, it might be that for every individual consumer the damage is small, but on the societal level, the most basic principles we agreed upon as a community are being openly disregarded (Jabłonowska *et al*., 2018). A kind of public monitoring is therefore needed, and to this end, fair marketing law will have to explore new avenues.

The UCPD sees 'manipulation' as the material distortion of the average consumer's economic behaviour. This design creates a strong link between the manipulative practice and the ideal consumer to whom it is addressed. The effect is that the analysis on the unlawfulness of the practice is often reduced to simply establishing who the average consumer is. The author believes that this architecture is fundamentally inconsistent with the reality of marketing practices. As said in Section 4, AI is conceptually transforming manipulation into the ability of companies to manage and shape automatic processes to guide customers towards predetermined outcomes. In this scenario, where decisions are made behind the scenes and embedded in automated decision-making processes, the likelihood of a manipulative outcome will depend less on whether the average consumer is rational or vulnerable or otherwise. In personalised commercial practices, there is nothing as the 'average consumer': depending on the profile, for each individual consumer there might exist a specific strategy of persuasion, thus a potential "personalised harm". It will be difficult to abstract away a standardised notion of the consumer from each specific situation. Second, if manipulation becomes a structural feature of AI-mediated business-to-consumer interactions, the ability of consumers to remove themselves from the business's influence will be reduced and impractical, unless they decide not to interact with the business in the first place.

It seems that against this development the only way to ensure fair commercial relations is to 'open the box' and find original solutions that can better balance the opposite interests of market players.

An attempt in this direction was recently made by the EU Directive 2019/2161, issued with the goal of a "better enforcement and modernisation of EU consumer protection rules", the so-called "Omnibus Directive". The Directive amends four EU consumer law directives: the Unfair Contract Terms Directive, the Price Indication Directive, the Contract Rights Directive and the Unfair Commercial Practice Directive. As expressed in the title, the legislative act has been driven by the need to improve awareness among consumers, traders and legal practitioners about consumer rights and to improve enforcement of consumer rights and consumer redress, as well as update or introduce new rules tailored to online market transactions.

In this latter respect, the reform brings some important novelties that address two developments directly related to the issue of consumer manipulation through means of technological influence. First, the operator of an online marketplace

must provide consumers with clear information regarding the criteria used to rank offers and, more specifically, whether the higher ranking of a certain product is the consequence of a direct or indirect payment.[24] Ranking refers to the prominence of offers of traders and the relevance given to search results as presented, organised and communicated to consumers (resulting, for example, from algorithmic sequencing, visual highlights and rating or review mechanisms). By its nature, ranking is considered one of the most effective nudges of consumer choice (Mik, 2016): it does not hide all relevant options, but the positioning of a search result heavily conditions the likelihood of a choice being made. Secondly, consumers must be informed when a price presented to them online is based on algorithms taking into account their personal behaviour.[25] Price personalisation, too, has elements that make it problematic for the consumer to take free decisions. As noted by Ezrachi and Stucke, through price personalisation techniques, firms can nudge consumers closer to their reservation price by the way they present options online (Ezrachi and Stucke, 2016). For example, the practice of price steering (e.g., showing different products to customers in different demographic groups) can lead certain consumers to make sub-optimal decisions and more expensive purchase choices (Hannak *et al.*, 2014).

These legislative developments represent an attempt to prevent the use of technology in online transactions from taking place at the expense of consumers. Nonetheless, the author believes that they can only provide partial protection to consumers. Firstly, they are largely based on information requirements. There is no ban on using algorithmic systems to influence consumers in ranking or in personalising prices, as long as this is made clear to the user. Moreover, ranking and personalisation are only two of the nudges currently deployed by online platforms operators. Emotional tricks, framing effect, image motivation, social norms are other very effective triggers that influence choice relevantly and might need to be taken into account in a more comprehensive legislative reform (Mirsch *et al.*, 2017). Finally, the recent legislative developments do not take a serious stance on the problems for consumers' decision-making connected to the deployment of AI tools in marketing – and not simply of automated decision-making systems. The latter are capable – and will be even more so in the future – of outperforming human beings in intellective and reactive capacity to understand consumers, their desire, needs, fears and emotions. This will likely result in the risk of widespread manipulative outcomes and the systematic failure of consumers' sovereignty in purchase decisions, as long as serious reform-oriented debate is not undertaken.

---

24   Article 3(4)(b) and Article 3(7)(a) of the Omnibus Directive.

25   Article 4(4)(a)(ii) of the Omnibus Directive. See also Recital 45.

7. New regulation: The example of the US Detour Act

In addition to the aforementioned EU legislative development, it seems appropriate for our analysis to consider the recent US legislative proposal of the DETOUR Act which, under the umbrella of US fair market law, explores new ways to protect consumers against manipulative practices of digital business.

The Federal Trade Commission Act is the most essential US fair trade and consumer protection law. It establishes the Federal Trade Commission (FTC), the authority responsible for enforcing its provisions and promoting consumer protection. The FTC has considerable powers to prosecute any "unfair or deceptive acts or practices in or affecting commerce" (Section 5 FTC Act).

Through this broad mandate, the FTC has become the key player in the suppression of unfair and manipulative marketing practices over the last century. In particular, the vague definition of 'unfair or deceptive act' has enabled the Commission to play a critical role in limiting unforeseen new business practices made possible by technological advancements (Hoofnagle, 2016). Over the years, however, the Commission's power has been progressively reduced through policy rules, additional regulation, and case law. Most recently, the US Congress has limited the authority of the FTC by specifying that it can consider only established public policies as evidence of consumer harm. This change significantly reduced the ability of the Commission to develop autonomously new enforcement policies to cope with digital transformations.[26]

That is why it is material that, on 7th April 2019, two US Senators issued a draft bill (called DETOUR Act, as an acronym for 'Deception Experience To Online User Reduction Act') whose purpose is to "prohibit the usage of exploitative and deceptive practices by large online operators and to promote consumer welfare in the use of behavioural research by such providers."[27]

The whole bill revolves around Section 3, which is relevant for three reasons.

A) First, it proclaims certain practices that are carried out by operators of large online platforms[28] and related to the manipulation of user interface, as "*unfair and deceptive acts and practices*". In this, the proposal immediately suggests that the new law should be regarded as a *lex specialis* of the FTC Act.

---

26   Section 45, lit. n, FTC Act.

27   S. 1084/2017, info at https://www.congress.gov/bill/116th-congress/senate-bill/1084/ all-info, text at https://it.scribd.com/document/405606873/Detour-Act-Final.

28   These are defined as any person that a) provides an online service ("a website or a service, other than an internet access service, that is made available to the public over the internet, including a social network, a search engine, or email service"); b) has more than one million of authenticated users in any 30-day period; and c) is subject to the jurisdiction of the FTC.

Accordingly, under Section 3(d), violations of the provisions will be treated as violations of the rules defining an unfair or deceptive act or practice under Section 18(a)(1)(B) FTC Act. The FTC will thus be formally invested by the Congress with powers to police new deceptive practices. More specifically, the conducts considered unfair and thus prohibited are:

1. "to design, modify, or manipulate a user interface with the purpose or substantial effect of obscuring, subverting, or impairing user autonomy, decision-making, or choice to obtain user data";
2. "to subdivide or segment consumers of online services into groups for the purposes of behavioural or psychological experiments or studies, except with the informed consent of each user involved";
3. "to design, modify, or manipulate a user interface on a website or online service, or portion thereof, that is directed to an individual under the age of 13, with the purpose or substantial effect of cultivating compulsive usage, inclusive of video auto-play functions initiated without the consent of a user".

B) Secondly, regardless of the specific practice, if the online platform engages in any form of behavioural or psychological research based on the activity of its user, it shall routinely, but no less than once each 90 days, disclose (1) to its users, "*any experiments or studies that user was subjected to or enrolled in with the purpose of promoting engagement or product conversion*"; (2) to the public, "*any experiments or studies with the purposes of promoting engagement or product conversion being currently undertaken, or concluded since the prior disclosure*". In this regard, the platforms shall also establish an internal Independent Review Board (IRB) that has to review and approve any behavioural or psychological experiments or research conducted by the online platform operator.

C) Finally, the draft establishes the option for an association of online platforms to register as Professional Standard Bodies (PSB). These will be qualified to develop guidelines for the development and design of technology products consistent with the provisions of the Act. The PSB will also be competent to define conducts that "*do not have the purpose or substantial effect of subverting or impairing user autonomy, decision-making, or choice, or cultivating compulsive usage for children*" against which the FTC cannot bring an enforcement action. These are, for example, a) *de minimis* user interface changes derived from testing consumer preferences including styles, layouts, text, except when they are aimed at obtaining user consent or user data; b) algorithms out of the control

of the platform; c) establishing default settings that provide enhanced users' autonomy and decision-making.

We will not go much into details of the proposal because, although it is promising, it remains a proposal. Given the clear impact of the new provisions on the commercial activity of platforms, we will probably have to wait a long time before the proposal becomes, if ever, a law. Nevertheless, without advancing any conclusive judgements, I would like to draw attention to two main points which, in my view, suggest where fair marketing law is heading.

Behavioural and psychological experiments and research with the purpose of promoting engagement or product conversion.
As we said, Section 3 places a general prohibition on the practice of subdividing or segmenting consumers into groups for behavioural and psychological experiments and research, unless valid consent is obtained from each user involved. For its validity, consent must be given before the user is included in any experiments or study and must be informed. According to Section 2, it should allow for an informed decision on the *voluntary participation* in the experiment or study and ensure an *understanding of any benefits, risks or consequences of participation*. In addition, regardless of whether it subdivides or segments consumers into groups, when the platform performs behavioural and psychological experiments and research, it must comply with the disclosure requirements set out in Section 3(b)(1). The latter requires the communication to the user and the public of any experiments or studies of which the user has been a part of "*with the purpose of promoting engagement or product conversion*".
Both provisions focus on the notion of 'behavioural and psychological experiments and research'. This is described as any "*study, including through human experimentation, of overt or observable actions and mental phenomena inferred from behaviour, including interactions between and among individuals and the activities of social groups*". The definition is very broad and seems to encompass any possible analysis of user data. While 'overt and observable action' refers to simple behavioural data (e.g., purchase records, transactional data, search data, app usage data, etc.), 'mental phenomena inferred from behaviour' seem to be exactly the use of analytics techniques to predict psychological traits or emotional state (both are mental phenomena). Data on the private and public interaction between individuals and within a group are also covered (e.g., likes, post, images on social media, etc.). In other words, any kind of data-analysis activity is likely to be a 'behavioural and psychological experiment and research'.[29]

---

29   The prohibition in Section(3)(a) applies only when the platform carries out behavioural and psychological experiment and research *upon having subdivided or segmented consumers into groups*. However, it is

As an additional element, the definition also includes experiments, especially 'human experiments'. Although the term is not clear and full of ambiguities, one might expect that the notion of 'experiments' here refers to the activity of repeatedly testing in a digital environment the actual behaviour of consumers (hence the 'human' component) under different conditions.[30] If 'experiments' is so understood, the proposal will certainly put under scrutiny many of the common techniques that are used in web design, such as A/B testing. At the same time, it is not clear whether the concept will also include new, more sophisticated methods of experimentation driven by AI. Reinforcement learning, in particular, is described as "type of sequential experimentation" as it implies a process of trial-and-error coupled with feedback provided by the environment that signals the utility of the result (Varian, 2018). If yes, the platform that uses RL-based marketing tools supposedly should provide meaningful information on how the algorithms works and how it affects consumers' decision-making, so as to ensure that users understand any possible risks, including potential distorting outcomes (e.g., a mapping between consumer action and the algorithm's goal). However, this aspect needs further investigation.

Operators of the platform should disclose behavioural and psychological experiments and research that are aimed at *promoting engagement* and *product conversion*. It is easy to understand why these two terms appear. Out of rhetoric, 'engagement and product conversion' are rather ambivalent concepts: they indicate the 'end' of a business strategy, but do not consider the 'means' of it. Manipulative tactics using 'morphing' may well promote the user website engagements, as much as exploitative advertising techniques could lead to optimisation of conversion. However, according to what we said in the previous sections, they would be hardly considered acceptable from an ethical point of view. It is therefore commendable that the proposal establishes not only safeguards before the study or the experiment (*users' consent*), but also afterwards (*the obligation to periodically disclose the conducted experiments and study*). The latter, in particular, aims to create an open and transparent accountability system that allows both competitors (who may have an interest in consumers not being manipulated) and law enforcement authorities to monitor the operations of platforms on users' data.

---

rare that the research and experiments that the bill is supposed to address do not previously involve some form of subdivision or segmentation of consumers into groups. Large online platforms indeed are generally based on big data and analysing such a large amount of data would be too expensive if some profiling or clustering techniques were used.

30   OFCOM, *Using experiments in consumer research*, 2010, accessible at <https://www.ofcom.org.uk/__data/ assets/pdf_file/0023/31865/experiments.pdf>.

Obscuring, subverting, or impairing user autonomy, decision-making, or choice.

If the bill becomes law, large online platforms will no longer be able to design, modify, or manipulate their website with the purpose or the substantial effect of 'obscuring, subverting, or impairing user autonomy, decision-making, or choice to obtain consent or user data'. This prohibition refers to the infamous phenomenon of 'dark patterns', which describes the practice of online platforms' operators of implementing misleading website design to extort consumers consent, either for the acceptance of terms and conditions or for the processing of their personal data (Gray *et al*., 2018).

In addition to the general ban, the bill also demonstrates a constructive approach for protecting users' autonomy vis-à-vis new AI-mediated forms of manipulation. Section 3(c)(3) states that associations of online platforms' operators (i.e., Professional Standards Body) are responsible for establishing behaviours that do not have the purpose or the substantial effect of subverting or impairing user autonomy, decision-making, or choice, among which 'default settings' that 'enhance their autonomy and decision-making ability'. In the wake of the 'privacy by default' movement (i.e., the adoption of highest standard of privacy protection as the default rule), the bill seems to support the idea of a share, sectorial implementation of an autonomy-enhancing environment (Brownsword, 2011), which empowers consumers to make autonomous and informed choices and avoids undue forms of influence (e.g., ad-blocking tools, personalised filters, customised settings, friendly UX design standards, etc.).[31]

It has been noted that many of the provisions of the proposal are poorly defined, which, if not changed, will leave too much discretion to the FTC.[32] Surely, however, there will be time and space for changes and improvements in the delineation of some concepts. Yet, for the time being, the draft proposal of the Detour Act already foresees interesting pathways that the EU fair marketing law could follow. First, it shows that the institutional secrecy on algorithms and machine-driven experimentation can no longer be sustained if the system aims to ensure fair marketing relationships and prevent manipulating outcomes. Fair marketing law will have to consider new forms of accountability, either through private disclosure requirements or through some form of public control. Secondly, fair marketing law should further explore the opportunity of a co-regulatory

---

31   It would also be interesting to start a discussion on 'consumer protection by design', i.e., AI-powered marketing tools that integrate *a priori* consumers' autonomy protection standards that prevent a manipulating outcome.

32   Will Rinehart, *DETOUR Act Gives Sweeping Powers to FTC*, American Action Forum (April 15, 2019), accessible at <https://www.americanactionforum.org/insight/ detour-act-gives-sweeping-powers-to-ftc/.>.

approach, where the legislator establishes general prohibitions and/or mandates for disclosures and the organisations of platforms' operators – under the supervision of an independent and external authority – provide for technical standards. Lawyers, especially consumer lawyers, should be involved in the future discussion.

8. Conclusion

Artificial Intelligence is affecting business activities and strategies and will do so more and more in the years to come. This transformation is having enormous repercussions on the economic and legal substance of the relationship between businesses and consumers. In particular, the fact that relationships between businesses and consumers are increasingly mediated by the presence of intelligent algorithms that learn how to improve in order to carry out tasks on the basis of continuous consumers' feedback, will not only profoundly strengthen the ability of companies to better allocate their products, but will also diminish the ability of consumers to make their own decisions and pursue their self-interest.

The use of AI by companies in marketing rekindles the debate about commercial manipulation and the opportunity for legal intervention. Since the dawn of business, whenever a new technology is used in marketing, a general wave of 'persuasion panic' is often contrasted with sceptics who claim that new techniques are not as effective as one would think or are indistinguishable from already established practice. Against AI-powered marketing, this argument of 'nothing new under the sun' seems very weak. Not simply because of a matter of scale (AI increases the granularity and the 'sensitivity' of information about consumers to an unprecedented degree, thus increasing the effectiveness of marketing and advertising), but also because AI is leading marketing to a change in nature. The latter is turning into the activity of companies to pre-design interactions with consumers in order to gain an economic advantage by means of skillful coding. This is not to say that every company that uses AI is obnoxious and manipulates consumers, but that manipulation of consumers behaviour might result as a by-product of business-consumer interactions.

In this context, it is not inappropriate to question the role of fair marketing law. Directly or indirectly, the latter has always been concerned to protect consumers' decision-making from any unfair intrusions by commercial actors. If this mission is to be continued in the future, consideration must be given to whether the existing regulation is capable of protecting from AI applications that are spreading in the market. In the European Union, the Unfair Commercial Practice Directive provides a remarkably open and technology-neutral framework which,

at first sight, seems to offer enough room for evolving and novel interpretations. At the same time, however, the legal structure and the normative bedrock of the Directive end up revealing a socioeconomic reality, that of purely 'analogic' and standardised consumption, which is largely already behind us. The recent EU legislative attempt to modernise fair marketing law and to regulate the use of automated decision-making to influence consumers' choices reflects only a partial understanding of the current and future technological developments in the market reality. We are not simply living in the era of online data-driven consumption, but entering the age of "AI-mediated consumption". Here, businesses operations are augmented by AI technologies which offer new opportunities for customising advertising, offerings, and contracts, but at the same time pose serious and unprecedented threats to consumers (Zuboff, 2019).

Fair marketing legislation may need to explore new avenues. Looking especially at some regulatory developments that are taking place overseas, we briefly reviewed two original solutions. The first responds to the need of finding a new balance between the economic freedom of businesses and the autonomy and cognitive freedom of consumers, and allow a diffuse control over algorithmic processes that potentially affect consumer decision-making. The second is aimed at providing, within the framework of general legislation and with 'multi-stakeholder' approach, (consumer protection by) design standards to create technology-mediated business interactions that preserve the autonomy of consumers.

Bibliography

BAGOZZI, R. P., GOPINATH, M., and NYER, P. U. (1999). The role of emotions in marketing. *Journal of the academy of marketing science*, 27(2):184-206.

BALKIN, J. M. (2015). The path of robotics law. *California Law Review, Forthcoming*.

BROWNSWORD, R. (2011). Autonomy, delegation, and responsibility: Agents in autonomic computing environments. In *Law, Human Agency and Autonomic Computing*, pp. 80-100. Routledge.

BROWNSWORD, R. (2016). The E-Commerce Directive, consumer transactions, and the digital single market: questions of regulatory fitness, regulatory disconnection and rule redirection. In *European contract law in the digital age*, pp. 163-204. Intersentia.

BURR, C. and CRISTIANINI, N. (2019). Can machines read our minds? *Minds and Machines*, pp. 1-34.

BURR, C., CRISTIANINI, N., and LADYMAN, J. (2018). An analysis of the interaction between intelligent software agents and human users. *Minds and machines*, 28(4):735-774.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*. https://doi.org/10.1177/2053951715622512

Calo, R. (2014). Digital market manipulation. *Geo. Wash. L. Rev.*, 82:995.

Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intelligent Systems*, 31(2):102-107.

Culotta, A. and Cutler, J. (2016). Mining brand perceptions from twitter social networks. *Marketing science*, 35(3):343-362.

D'mello, S. K. and Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing Surveys (CSUR)*, 47(3):43.

Ebers, M. (2019). Chapter 2: Regulating AI and Robotics: Ethical and Legal Challenges, in Martin Ebers and Susana Navas Navarro (eds.), Algorithms and Law, Cambridge, Cambridge University Press. Available at SSRN: https://ssrn.com/abstract=3392379 or http://dx.doi.org/10.2139/ssrn.3392379

Ezrachi, A., and Stucke, M. (2016). E. *Virtual Competition*. Harvard University Press.

Floridi, L. (2019). Marketing as control of human interfaces and its political exploitation. *Philosophy & Technology*, pp. 1-10.

François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., Pineau, J., *et al*., (2018). An introduction to deep reinforcement learning. *Foundations and Trends in Machine Learning*, 11(3-4):219-354.

Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., and Watts, D. J. (2010). Predicting consumer behavior with web search. *Proceedings of the National academy of sciences*, 107(41):17486-17490.

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., and Toombs, A. L. (2018). The dark (patterns) side of UX design. In Proceedings of the 2018 CHI Conference on Human *Factors in Computing Systems*, p. 534. ACM.

Hacker, P. (2017). Personal data, exploitative contracts, and algorithmic fairness: Autonomous vehicles meet the Internet of Things. *International Data Privacy Law*, 7(4):266–286.

Hannak, A., Soeller, G., Lazer, D., Mislove, A., and Wilson, C. (2014). Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 conference on internet measurement conference* (pp. 305-318). ACM.

Harambam, J., Helberger, N., and van Hoboken, J. (2018). Democratizing algorithmic news recommenders: how to materialize voice in a technologically saturated media ecosystem. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133):20180088.

Hauser, J. R., Urban, G. L., Liberali, G., and Braun, M. (2009). Website morphing. *Marketing Science*, 28(2):202-223.

HELBERGER, N. (2016). Profiling and targeting consumers in the Internet of Things – A new challenge for consumer law. In *Digital revolution: challenges for contract law in practice*, pp. 135-161. Nomos.

HIRSH, J. B., KANG, S. K., and BODENHAUSEN, G. V. (2012). Personalized persuasion: Tailoring persuasive appeals to recipients' personality traits. *Psychological science*, 23(6):578-581.

HOOFNAGLE, C. J. (2016). *Federal Trade Commission privacy law and policy*. Cambridge University Press.

HOWELLS, G., MICKLITZ, H.-W., and WILHELMSSON, T. (2006). *European fair-trading law: The unfair commercial practices directive*. Routledge.

HOWELLS, G., TWIGG-FLESNER, C., and Wilhelmsson, T. (2017). *Rethinking EU consumer law*. Routledge.

HU, A. and FLAXMAN, S. (2018). Multimodal sentiment analysis to explore the structure of emotions. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 350-358. ACM.

INCARDONA, R. and PONCIBO, C. (2007). The average consumer, the unfair commercial practices directive, and the cognitive revolution. *Journal of consumer policy*, 30(1):21-38.

JABŁONOWSKA, A., KUZIEMSKI, M., NOWAK, A. M., MICKLITZ, H.-W., PAŁKA, P., and SARTOR, G. (2018). Consumer law and Artificial Intelligence: Challenges to the EU consumer law and policy stemming from the business' use of Artificial Intelligence – Final report of the ARTSY project.

KAHNEMAN, D. (2011). *Thinking, fast and slow*. Macmillan.

KAPTEIN, M., MARKOPOULOS, P., DE RUYTER, B., and AARTS, E. (2015). Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles. *International Journal of Human-Computer Studies*, 77:38-51.

KEIRSBILCK, B. (2011). *The new European law of unfair commercial practices and competition law*. Hart Oxford.

KELLEHER, J. D., MAC NAMEE, B., and D'ARCY, A. (2015). *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. MIT Press.

KOSINSKI, M., STILLWELL, D., and GRAEPEL, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802-5805.

MATZ, S. C., GLADSTONE, J. J., and STILLWELL, D. J. (2016). Money buys happiness when spending fits our personality. *Psychological science*, 27(5):715-725.

MATZ, S. C., KOSINSKI, M., NAVE, G., and STILLWELL, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the national academy of sciences*, 114(48):12714-12719.

MAZZINI, G. (2019). A System of Governance for Artificial Intelligence through the

Lens of Emerging Intersections between AI and EU Law, in A. De Franceschi – R. Schulze (eds.), *Digital Revolution – New challenges for Law*. Available at SSRN: https://ssrn.com/abstract=3369266

Mik, E. (2016). The erosion of autonomy in online consumer transactions. *Law, Innovation and Technology*, 8(1):1-38.

Mirsch, T., Lehrer, C., and Jung, R. (2017). Digital nudging: Altering user behavior in digital environments. *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)*, 634-648.

Ning, H., Dhelim, S., and Aung, N. (2019). PersoNet: Friend Recommendation System Based on Big-Five Personality Traits and Hybrid Filtering, in *IEEE Transactions on Computational Social Systems*, vol. 6, no. 3, pp. 394-402. doi: 10.1109/TCSS.2019.2903857.

Odekerken-Schröder, G., De Wulf, K., and Schumacher, P. (2003). Strengthening outcomes of retailer-consumer relationships: The dual impact of relationship marketing tactics and consumer personality. *Journal of business research*, 56(3):177-190.

Quercia, D., Kosinski, M., Stillwell, D., and Crowcroft, J. (2011). Our twitter profiles, our selves: Predicting personality with twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, pp. 180-185. IEEE.

Richardson, M., Dominowska, E., and Ragno, R. (2007). Predicting clicks: estimating the click-through rate for new ads. In *Proceedings of the 16th international conference on World Wide Web*, pages 521-530. ACM.

Roffo, G. and Vinciarelli, A. (2016). Personality in computational advertising: A benchmark.

Sax, M., Helberger, N., and Bol, N. (2018). Health as a means towards profitable ends: mHealth apps, user autonomy, and unfair commercial practices. *Journal of consumer policy*, 41(2):103-134.

Schwartz, H. A., Sap, M., Kern, M. L., Eichstaedt, J. C., Kapelner, A., Agrawal, M., Blanco, E., Dziurzynski, L., Park, G., Stillwell, D., *et al*. (2016). Predicting individual well-being through the language of social media. In *Biocomputing 2016: Proceedings of the Pacific Symposium*, pp. 516-527. World Scientific.

Sibony, A.-L. (2014). Can EU consumer law benefit from behavioural insights? An analysis of the unfair practices directive. *European Review of Private Law*, 22(6):901-941.

Stuyck, J., Terryn, E., and Van Dyck, T. (2006). Confidence through fairness? The new Directive on unfair business-to-consumer commercial practices in the internal market. *Common market law review*, 43(1):107-152.

Sunstein, C. R. (2016). *The ethics of influence: Government in the age of behavioral science*. Cambridge University Press.

Teixeira, T., Wedel, M., and Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of marketing research*, 49(2):144--159.

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Trzaskowski, J. (2018). Behavioural innovations in marketing law. In *Research Methods in Consumer Law*. Edward Elgar Publishing.

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., and Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political Polarization, and Political Disinformation: A Review of the Scientific Literature (March 19, 2018)*.

Twigg-Flesner, C. (2018). The EU's Proposals for regulating B2B relationships on online platforms transparency, fairness and beyond. *Journal of European Consumer and Market Law*, 7(6):222–233.

Varian, H. (2018). Artificial intelligence, economics, and industrial organization. Technical report, National Bureau of Economic Research.

Varian, H. R. (2010). Computer mediated transactions. *American Economic Review*, 100(2):1-10.

Waddington, L. (2013). Vulnerable and confused: the protection of "vulnerable" consumers under EU law. *European law review*, 38(6):757-782.

Weatherill, S. (2013). *EU consumer law and policy*. Edward Elgar Publishing.

Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of research in personality*, 44(3):363-373.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1):75-89.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

DIREITO PRIVADO