European Review of Digital Administration & Law – Erdal
2022, Volume 3, Issue 1, pp. 31-44
ISSN 2724-5969 – ISBN 979-12-218-0078-4 DOI 10.53136/9791221800784 03

*Law, Digital Nudging and Manipulation: Dark Patterns, Artificial Intelligence and the Right to Good Administration**

Juli Ponce
(Full Professor of Administrative Law at the University of Barcelona)

ABSTRACT The use of behavioural insights in the digital area has grown considerably in importance during the COVID-19 pandemic and can be an element of promoting the right to good administration in the public sector. On the other hand, digital nudging has a dark side, both in private and public sector. The possible use by public and private sectors of so-called dark patterns, concerning which the European Parliament has recently proposed to include a ban in the future Digital Services Act, and what is known as hypernudging raises legal questions regarding a possible manipulation that violates freedom of thought, as indicated by the Committee of Ministers of the Council of Europe in a statement of 2019. This article deals with the definition of those concepts and their possible legal regulation, by means of considering some international examples.

1. *Introduction*

The Covid-19 pandemic has led to a dramatic rise in Internet use. Recent research shows, in fact, that in the past year there has been an increase of up to 30% in digital consumption.[1]

In this digital environment, choice architectures are constantly made, either actively or passively. For those who are not familiar with the term, choice architecture is a concept reflecting the awareness that the choice between different options is affected by the way in which such options have been proposed.[2]

Humans face choices every day, but the result of every decision is influenced not only by rational deliberations regarding the available options. The design of the choice environment in which the information is presented can exert a subconscious influence on the outcome. In other words, the decision often depends on how the choice is presented; hence, decision architecture alters people's behaviour in predictable ways. The simplest changes in the choice environment —in which options are presented— can influence people's decision and "nudge" them to behave in certain ways. In fact, there is no neutral way to present options. For example, it has been proven that the mere act of changing the default options for organ donation —from opt-in to opt-out— has almost doubled the percentage of people who consent to donate organs.

There is always a design of the context of decision, which is created and modelled —consciously or unconsciously— by an architect of the choice: a context in which consumers and users of public services choose between specific options and come to decisions (buying, getting vaccinated, etc.) and the same happens in the digital ambit.

Accordingly, it is inferable that there can be —and there actually are— private and public activities aimed at encouraging or discouraging consumers and users' behaviour both outside and —most importantly— within the digital world. These architectures of choice can be transparent or not and have purposes that may turn out to be acceptable and even positive (e.g. encouraging consumption without scams, respecting the will of the consumers, customizing public services to provide better public management, etc.) or ethically and legally unacceptable[3] (e.g. increasing the sale of products or services to consumers, guiding or hindering the use of public services,[4] obtaining personal

* Article submitted to double-blind peer review.
This article is one of the results of the Spanish National Project PID2020-115774RB-I00, Citizen-Centric Services, Biases and Artificial Intelligence: Towards a Consolidation of Digital Rights in Public Administrations, funded by mcin/aei/10.13039/501100011033.

[1] See figures provided by WARC: *Data Global Ad Trends: The State of the Industry 2020/21*, in www.warc.com

[2] R. H. Thaler, C.R. Sunstein and J.P. Balz, *Choice architecture,* in E. Shafir (ed.), *The behavioral foundations of public policy*, Princeton, N.J., Princeton University Press, 2013, 428.

[3] M. Lavi, Evil nudges, in Vanderbilt Journal of Entertainment & Technology Law, vol. 21, issue 1, 2018, 1.

[4] R.H. Thaler, *Nudge, not sludge*, in *Science*, vol. 361, issue 6401, 2018, 431.

data without a clear and explicit consent —thus manipulating people—, etc.).

In other words, in those and other cases, we are dealing with digital nudges. A nudge is, according to Thaler and Sunstein's well-known definition, any cheap and easy-to-avoid aspect of the architecture of decisions in the digital environment that modifies people's behaviour in a predictable way, without prohibiting any option and without changing economic incentives.[5] Therefore, digital nudging, for the purpose of this reflection, is the use of user interface design elements to guide people's behaviour in digital choice environments. In turn, digital choice environments are user interfaces —such as web forms and ERP (Enterprise Resource Planning) screens— that require people to make judgments or decisions.[6]

Given the current spectacular growth in the use of digital media, the architecture of digital choice is gaining importance and the same happens in the case of the digital incentives or nudges that persuade consumers and users of public services.

The present article is a brief analysis on how these incentives are developing in the private and public sectors. It will address the possible use, by governments, digital platforms and companies, of behavioural insights achieved in recent decades. In the case of public sector, those digital incentives can be a way of promoting the right of good administration (art. 41 of the European Charter of Fundamental Rights and equivalent national regulations[7]), but their digital application can be also used to take advantage of people's cognitive biases with the consequent risk of inacceptable manipulation —specifically identified by the Committee of Ministers of the Council of Europe in the *Declaration on the manipulative capabilities of algorithmic processes* of 13th February 2019.[8]

It is important to highlight that a recent publication promoted by the Council of Europe warned: "Special attention should also be paid to the potential use of AI in human-machine interaction to implement nudging strategies. Here, due to the complexity and obscurity of the technical solutions adopted, AI can increase the passive role of citizens and negatively affect the democratic decision-making process. Otherwise, an active approach based on conscious and active participation in community goals should be preferred and better managed by AI participation tools. Where adopted, nudging strategies should still follow an evidence-based approach".[9]

2. *The dark side*

Going to our first ambit of analysis, applications and websites in the private sector —without specific regard to AI —, it is necessary to reflect, first of all, on the expression that designates the manipulative use of digital nudges to the detriment of consumers and users: dark patterns.

The article will subsequently endeavour to answer some essential questions regarding dark patterns:

- What are dark patterns?
- Which are the most common dark patterns?
- What cognitive biases do dark patterns exploit?
- What can be done against them?
- Should public intervention be necessary against dark patterns? Which type of intervention?
- Is there any other open issue regarding dark patterns?

2.1. *What are dark patterns?*

According to the insights of various specialists,[10] dark patterns are designs of User Interface (UI) and User Experiences (UX) that try to exploit people's vulnerabilities through manipulation and scamming with the intention of pushing them towards a certain outcome.

This definition highlights the breadth of the concept of dark patterns as well as the vast number of purposes they can serve (e.g., obtaining more personal data, money, influencing a vote or, in general, any decision).

[5] R.H. Thaler and C.R. Sunstein, *Nudge. The Final Edition*, London, Penguin Books, 2021.
[6] M. Weinmann, C. Schneider and J. vom Brocke, *Digital Nudging*, in *Business & Information Systems Engineering*, vol. 58, issue 6, 2016, 433.
[7] J. Ponce, The Right to Good Administration and the role of Administrative Law in promoting good government, in A. Cerrillo and J. Ponce (eds.), Preventing Corruption and Promoting good Government and Public Integrity, Bruxelles, Bruylant, 2017, 25.
[8] See Declaration by The Committee of Ministers on the manipulative capabilities of algorithmic processes, in www.coe.int.
[9] Council of Europe, Towards regulation of AI systems, 2020.
[10] See *Dark Patterns,* in www.darkpatterns.org.

Similarly, the *California Privacy Rights Act* (CPRA) of 2020, which, as will be explained, was recently amended to ban dark patterns, reports: "Dark pattern" means a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision making, or choice, as further defined by regulation".[11]

2.2. *Which are the most common dark patterns?*

Dark patterns have been detected, studied and labelled with names that are undoubtedly original. Some concrete examples of these obscure designs, extracted from various sources will be presented hereafter to better understand the phenomenon.[12]

Confirmshaming. "Confirmshaming" is a dark pattern in which the user must choose between activating specific options/signing up for some service or not. In case of dissent, the consumer is made to feel bad, guilty or ashamed.[13]

Disguised Ads. This is a dark pattern in which ads appear "disguised", confused in the midst of normal content, video players or navigation elements, in order to mislead the user into clicking on them without noticing it.[14]

Forced Continuity. "Forced continuity" occurs when money is charged without warning at the end of a free trial of a service or in the case of subscriptions that are automatically renewed without asking for explicit consent.[15]

Friend Spam. The platform asks for permissions to access email, phone and/or social networks' contacts for a specific action —for example finding friends— but such permissions are used to send spam to the user's contacts.[16]

A few years ago, LinkedIn, which regularly resorted to this design was given a 13-million-dollar fine, as it was considered a clearly abusive practice.[17]

Misdirection. As suggested by the name, "misdirection" is a dark pattern consisting of a distraction of users aimed at making them follow a path that leads to a pre-set outcome and not to the one they really wanted to achieve.[18]

Price comparison prevention. This dark pattern hinders the comparison between one item and another in order to prevent users from making informed decisions.[19]

Privacy Zuckering. The name of this design combines —for obvious and well-known reasons— the surname of Mark Zuckerberg, Facebook's CEO, with the informal term "sucker". In fact, it takes place when users are tricked into sharing more private information than they really want. This is because the small print hidden in the terms and conditions that users accept in order to access online services gives permission to sell their personal data to other companies.[20]

Roach motel. Behind this name hides a very common practice that consists of facilitating the entry or subscription to a service and then making cancellation extremely difficult.[21]

Bait and switch. A dark pattern arising in those cases in which the user wants to realize an operation, but performs a completely different one, which is the one that interests the "misleading" website.[22]

Sneak into basket. This is an online sales systems' practice in which some extra items are included in the shopping basket to make people inadvertently buy them. Extra items are usually added via a checkbox or a radio button that is hardly visible during one or more steps of the purchase. It has been a very common practice on the websites of low-cost airlines.[23]

Hidden Costs. This dark pattern is very similar to the previous one, as it consists of the sudden inclusion of some extra costs, such

[11] See California Privacy Rights Act (CPRA) of 2020.
[12] C. Álvarez, Dark Patterns: the dark side of the UX, in www.wildwildweb.es; Dark Data — Zines, in www.parsons.edu.; Dark patterns - Types of dark pattern, in www.darkpatterns.org.
[13] See examples at *confirmshaming*, in www.tumblr.com.
[14] See example at *confirmshaming*, in www.tumblr.com.
[15] See examples at *Forced continuity - a type of dark pattern*, in www.darkpatterns.org.
[16] See examples at Forced continuity - a type of dark pattern.
[17] See After Lawsuit Settlement, LinkedIn's Dishonest Design Is Now A $13 Mil, in www.fastcompany.com.
[18] An example of this hard-to-define design can be found at *Misdirection*, in www.darkpatterns.org.
[19] See examples at Price comparison prevention - a type of dark pattern, in www.darkpatterns.org.
[20] See examples at Price comparison prevention - a type of dark pattern.
[21] See examples at *Roach motel - a type of dark pattern*, in www.darkpatterns.org.
[22] See examples at *Bait and switch - a type of dark pattern*, in www.darkpatterns.org.
[23] See examples at Bait and switch - a type of dark pattern.

as delivery costs or taxes. The main difference is that this one appears at the end of the sale process.[24]

2.3. *What cognitive biases do dark patterns exploit to manipulate consumers?*

The aforementioned examples of dark patterns seek to use consumer biases in a dishonest way to induce people to make mistakes/operations —or prevent them from doing specific actions— by manipulating them.[25]

Before going into the question in greater depth, it could be useful to say some words about cognitive biases.

In recent decades, thanks to the well-known work of the Israeli psychologists Amos Tversky —died in 1996— and the 2002 Nobel laureate Daniel Kahneman, psychology has contributed most to make it widely accepted that:[26]

- The absolute rationality of the person, of the *homo economicus* does not exist. First of all, because rationality is limited (as highlighted by Herbert Simon a long time ago) and, secondly, because it is a concept that does not take into account perfectly rational behaviors such as reciprocity and altruism (which give rise to a model of *homo reciprocans* that makes decisions based on social norms, in which reciprocity, altruism and trust matter).
- Rationality is interfered with by *heuristics* and *cognitive biases*. The works of the authors cited above point out that cognitive schemes and heuristics are rules that simplify the selection and processing of information. These are like intuitive *shortcuts*, which function as adaptive mechanisms against the limits of our cognitive resources (so a red octagon generally means "stop", while an outstretched hand expresses "greeting") and, in situations of risk and uncertainty, lead to certain assessment and prediction biases. Heuristics can provide fast and efficient shortcuts in information processing, but sometimes they also lead to systematic and predictable errors. Thus, heuristics produce errors, and biases are errors that occur systematically. Nevertheless, not all errors are biases, even though all biases are errors.

Due to these biases, *it is not unusual for our brains to mislead and* turn us into individuals who make mistakes and bad decisions, even when we have complete information. Although it may come as a surprise, since the deviations of people's rationality have already been studied well, the scientific advances of the last decades show us that people are not perfect decision-makers who maximize their interest in an absolutely rational way. Kahneman explained very educationally that two systems of decision-making coexist inside us: one is automatic and fast, the so-called system 1; while the other, system 2, is an effort linked to previous deliberation. System 1 is activated unconsciously and works well on various occasions, but on many others it leads to cognitive errors caused by those heuristics and biases that are used by our mind to make quick decisions without excessive energy consumption.

The main premise of the theory of cognitive psychology, therefore, is understanding that the human brain is a limited processor of information unable to successfully process all incoming stimuli.

Dark patterns are thus made to exploit cognitive biases; to affect humans and consumers' decision-making and emotions to their detriment and to the advantage of private companies that act as architects of people's choice.

"Confirmshaming", for example, exploits framing bias and sense of guilt, while "misdirection" takes advantage of lack of attention, the anchoring effect and scarcity bias, and so on... Some recent studies have already linked dark patterns with the cognitive biases that manipulate people to the advantage of companies.[27]

An overview of these connections can be found in the following table:

[24] See examples at *Hidden costs*, in www.darkpatterns.org.

[25] Those designs proved to be so significant that a member of the editorial board of the *New York Times* dedicated an article to them, available at: *Opinion | The Internet's 'Dark Patterns' Need to Be Regulated - The New York Times*, in www.nytimes.com.

[26] D. Kahneman, *Thinking, fast and slow*, Penguin Books, London, 2011.

[27] See A. Mathur, G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty and A. Narayanan, *Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites,* in CSCW, Article 81, 2019.

Legend: ●= Always, ◐= Sometimes, ○= Never

Category	Type	Description	# Instances	# Websites	Asymmetric?	Covert?	Deceptive?	Hides Info?	Restrictive?	Cognitive Biases
Sneaking	Sneak into Basket	Adding additional products to users' shopping carts without their consent	7	7	○	○	◐	●	○	Default Effect
	Hidden Costs	Revealing previously undisclosed charges to users right before they make a purchase	5	5	○	○	◐	●	○	Sunk Cost Fallacy
	Hidden Subscription	Charging users a recurring fee under the pretense of a one-time fee or a free trial	14	13	○	○	◐	●	○	None
Urgency	Countdown Timer	Indicating to users that a deal or discount will expire using a counting-down timer	393	361	○	◐	◐	○	○	Scarcity Bias
	Limited-time Message	Indicating to users that a deal or sale will expire will expire soon without specifying a deadline	88	84	○	◐	○	●	○	Scarcity Bias
Misdirection	Confirmshaming	Using language and emotion (shame) to steer users away from making a certain choice	169	164	●	○	○	○	○	Framing Effect
	Visual Interference	Using style and visual presentation to steer users to or away from certain choices	25	24	◐	●	◐	○	○	Anchoring & Framing Effect
	Trick Questions	Using confusing language to steer users into making certain choices	9	9	●	●	○	○	○	Default & Framing Effect
	Pressured Selling	Pre-selecting more expensive variations of a product, or pressuring the user to accept the more expensive variations of a product and related products	67	62	◐	◐	○	○	○	Anchoring & Default Effect, Scarcity Bias
Social Proof	Activity Message	Informing the user about the activity on the website (e.g., purchases, views, visits)	313	264	○	◐	◐	○	○	Bandwagon Effect
	Testimonials	Testimonials on a product page whose origin is unclear	12	12	○	○	◐	○	○	Bandwagon Effect
Scarcity	Low-stock Message	Indicating to users that limited quantities of a product are available, increasing its desirability	632	581	○	◐	◐	◐	○	Scarcity Bias
	High-demand Message	Indicating to users that a product is in high-demand and likely to sell out soon, increasing its desirability	47	43	○	◐	○	○	○	Scarcity Bias
Obstruction	Hard to Cancel	Making it easy for the user to sign up for a service but hard to cancel it	31	31	○	○	○	◐	●	None
Forced Action	Forced Enrollment	Coercing users to create accounts or share their information to complete their tasks	6	6	●	○	○	○	●	None

From: A. Mathur, G. Acar, M.J. Friedman, E. Lucherini, J. Mayer, M. Chetty and A. Narayanan, Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites, in CSCW, article 81, 2019.

2.4. *What can be done against dark patterns?*

Logically, the first step is to become aware of their existence, which, as mentioned previously, is often hard to detect. Drawing on the abovementioned research of Nobel Prize winner Daniel Kahneman, we, as people and consumers, must direct our personal effort towards enhancing system (2) of thinking, thereby avoiding system (1) characterized by quick and intuitive decision-making. This implies a de-biasing effort.

Undoubtedly, the personal effort required is going to be titanic, because we, as consumers, will face an army of designers equipped with knowledge of our biases and manipulability. This is clearly confirmed by some publications that list hundreds of intelligent design strategies based on people's way of thinking.[28]

[28] S. M. Weinschenk, *100 Things Every Designer Needs*

It is therefore necessary to reflect on whether people can be left alone against such evil architects of their choices; whether or not they can be involved in a David-Goliath struggle in which they represent the former, the one that rarely wins.

2.5. *What kind of public scrutiny can be exercised over dark patterns? What is the existing regulation of dark patterns in the European Union and the United States?*

In view of the above and given that dark pattern development actually involves market faults and corporate abusive practices, public interventions against them seem necessary.

Obviously, there can be different types of intervention, ranging from sermons to carrots and sticks.[29] Thus, public interventions can involve consumer information campaigns and promotion of companies' self-regulation (something that the EU has been trying with relatively limited success —in view of how widespread these practices are— since 2018 with the Code of Disinformation Practices[30]), as well as the legal regulation of digital decision architecture (by prohibiting and by establishing specific requirements for consumer protection, including the definition, where appropriate, of infringements and penalties —traditional "command and control" activity).

Both the European Union and the United States have regulations on dark patterns, but their approach is different.

In the US, some laws that specifically define and prohibit dark patterns are coming into force, as in the case of the State of California, as we have seen before.

Conversely, in the European Union there is neither a definition of the phenomenon nor an ad-hoc regulation yet. Several voices seem to agree that this would not be necessary, since existing European norms on data protection and consumer protection already regulate it in general terms.[31]

2.6. *What issues regarding dark patterns remain open?*

Shedding more light on dark patterns is necessary, especially in the case of the European Union, but it also essential to reflect on some questions that remain still unanswered, such as:

Do dark patterns subsist because of the absence of specific regulation or due to the lack of effective enforcement of existing general regulation in areas such as data protection or consumer protection?

Is self-regulation —like the EU Code of Disinformation Practices— a truly effective instrument in this area? Which model can be more effective, the American model of prohibition and explicit regulation, or the European one? Are there enough mechanisms in place in the EU to develop effective public policies against dark patterns? If not, what should be done in the future?

Finally, although the analysis of artificial intelligence has been excluded from this first reflection, it is impossible not to wonder what the use of machine-learning algorithms in combination with nudges will bring in the near future. "Darker" patterns? Does the recent proposal for EU Regulation in the field of Artificial Intelligence address this question? Should it? Or is it a question for the future Digital Services Act?

In that regard, European Parliament included amendments in the original text of this bill at the beginning of 2022 banning dark patterns.[32]

to Know about People, Indianapolis, IN, New Riders Publishing, 2011.

[29] R. Rist, Carrots, Sticks and Sermons. Policy Instruments and Their Evaluation, London, Routledge, 2003.

[30] See the text of the Code at Code of Practice on Disinformation | Shaping Europe's digital future, in www.archive-it.org.

[31] S. Rieger and C. Sinders, *Dark Patterns: Regulating Digital Design*, Stiftung Neue Verantwortung, 2020; S. Berbece, *Let There Be Light! Dark Patterns Under the Lens of the EU Legal Framework*, KU Leuven Student Master´s work, 2019, available at SSRN: https://ssrn.com/abstract=3472316.

[32] Specifically, the amendments are the following:
Amendment 105 introduces a proposal for a new regulation of article 2 – paragraph 1 – point q a: "(qa) 'dark pattern' means a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision-making or choice".
Amendment 158 introduces a proposal for a new regulation of article 12 – paragraph 2 c: "2c. Providers of intermediary services shall refrain from any dark patterns or other techniques to encourage the acceptance of terms and conditions, including giving consent to sharing personal and non-personal data".
Short justification included in the *Opinion of The Committee on Civil Liberties, Justice and Home Affairs* (28 July 2021): "Behavioural and personalised targeting for non-commercial and political advertising should be phased out to protect users and ensure the existence of traditional media, and be replaced by contextual advertising. The same should apply to targeting people based on sensitive data, or to targeting minors. Behavioural and personalised targeting for commercial advertising should only be possible where users have freely opted in, without exposure to 'dark' patterns or the risk of be-

After considering the problem of dark patterns and the possible legal solutions, the second part of this reflection will focus on the role of Artificial Intelligence and public intervention specifically.

3. *Decision architecture and digital nudges. Are they against humanity?*

The previous section introduced the idea that digital decision architectures and the use of digital nudges can represent serious risks of manipulation, as was declared to be the case by the Council of Europe in 2019, as we have seen. However, the focus of that reflection was on their use by the private sector, on the so-called dark patterns, thus no specific reference was made to artificial intelligence.

In this part, the analysis will be complemented by adding the public sector and AI to the discussion on possibilities and risks related to digital nudging.

3.1. *The great digital manipulation of our cognitive biases in the attention economy*

As mentioned in the previous section, the use of nudges, whether digital or not, does not necessarily imply manipulation or opacity. Digital nudges can be perfectly ethical and lawful, transparent, and encourage the consumption of "desirable" products or the pursuit of the general interest. Actually, in the case of governments and public administrations, digital nudges if transparent and design respecting the law can be an element of promoting the right to good administration, by serving people with a citizen-centric approach.[33]

Unfortunately, in the case of private sector, according to some authoritative voices, like Williams[34] currently digital nudges imply a

ing excluded from services, and without being fatigued by consent banners if they have already made a clear choice in their browser/device settings".

Amendment 11, proposal for a new regulation of Recital 15 b: "(15b) Targeting individuals based on personal data, including behavioural data, should not be permitted for non-commercial and political purposes. Misleading or obscure advertising for non-commercial and political purposes is a special class of online threat because it influences the core mechanisms that enable the functioning of our democratic society. Targeting minors on the basis of their personal data or targeting individuals on the basis of special categories of data which allow for targeting vulnerable groups should not be permitted. Targeting recipients for commercial purposes should require the recipients' consent. To ensure that recipients have a real choice, refusing consent should be no more complicated than giving consent, "dark patterns" should not be used to undermine the recipient's choice and refusing consent should not result in access to the functionalities of the platform being disabled. In order to avoid fatiguing recipients who refuse to consent, terminal equipment settings that signal an objection to processing of personal data should be respected. Displaying contextual advertisements does not require processing personal data and is thus less intrusive".

According to amendment 40, introducing a proposal for a new Recital 39: "(39a) Recipients of a service should be able to make a free, autonomous and informed decisions or choices when using a service and providers of intermediary services shall not use any means, including via its interface, to distort or impair that decision-making. In particular, recipients of the service should be empowered to make such decision sinter alia regarding the acceptance of and changes to terms and conditions, advertising practices, privacy and other settings, recommender systems when interacting with intermediary services. However, certain practices typically exploit cognitive biases and prompt recipients of the service to purchase goods and services that they do not want or to reveal personal information they would prefer not to disclose. Therefore, providers of intermediary services should be prohibited from deceiving or nudging recipients of the service and from distorting or impairing the autonomy, decision-making, or choice of the recipients of the service via the structure, design or functionalities of an online interface or a part thereof ('dark patterns'). This should include, but should not be limited to, exploitative design choices to direct the recipient to actions that benefit the provider of intermediary services, but which may not be in the recipients' interests, presenting choices in a non-neutral manner, such as giving more visual prominence to a consent option, repetitively requesting or urging the recipient to make a decision such as making the procedure of cancelling a service significantly more cumbersome than signing up to it. However, rules preventing dark patterns should not be understood as preventing providers to interact directly with users and to offer new or additional services to them. In particular it should be possible to approach a user again in a reasonable time, even if the user had denied consent for specific data processing purposes, in accordance with Regulation (EU) 2016/679. The Commission should be empowered to adopt a delegated act to define practices that could be considered as dark patterns".

The Explanatory statement underlines that: "In addition, the Rapporteur believes that the algorithms used in recommender system should be designed in a way that prevents dark patterns and rabbit holes from happening. Moreover, the Rapporteur suggest a "must-carry" obligation to ensure that information of public interest is high-ranked in the platform's algorithms".

[33] J. Ponce (ed.), *Nudges, Good Governance and Good Administration. Behavioral Insights, Nudging and Public and Private Sectors*, Athens, European Public Law Organization (EPLO), 2022 (upcoming book, soon to be published).

[34] In this section, I will consider the remarkable reflections offered by James Williams' book, *Stand out of our Light. Freedom and Resistance in the Attention Economy*, Cambridge, Cambridge University Press, 2018. This interesting publication of the co-founder of "Time Well Spent" —a movement that led to the creation of the Center for Humane Technology— won the Nine Dots Award. James Williams worked as a strategist at Google

large-scale manipulation project that has been undergoing development for a long time and which has barely been recognized until now. This project recalls religious/mythical/totalitarian systems, is in the hands of very few people in the world and is aimed at the consumer: an objective pursued by spending huge amounts of money on advertising (in 2017 advertising expenditure was 223 billion and it is growing by 10% annually). Consumers, as said, are the target of cognitive biases' manipulations and this has also been reported by Williams, who explicitly cites Kahneman and Tzaversky in support.

This large-scale system of manipulation operates in the attention economy, an environment in which digital products and services compete relentlessly to capture and exploit consumers' attention.[35] Obviously, the same risk in the case of the public sector can be also identified for other purposes.[36]

3.2. *Digital manipulation against the right to freedom of thought, the dignity of the person, free development of the personality and the Social and Democratic Rule of Law*

It is important to underline that attention is linked to freedom and human will, and the system of large-scale manipulation by digital design that has been described harms both.

First of all, this is because that there cannot be freedom of thought without freedom of attention. In the classic "On Liberty", John Stuart Mill's book published in 1859, already explained a very similar concept about freedom of thought (p. 15 ff.):

"It comprises, first, the inward domain of consciousness; demanding liberty of conscience in the most comprehensive sense; liberty of thought and feeling; absolute freedom of opinion and sentiment on all subjects, practical or speculative, scientific, moral, or theological. The liberty of expressing and publishing opinions may seem to fall under a different principle, since it belongs to that part of the conduct of an individual which concerns other people; but, being almost of as much importance as the liberty of thought itself, and resting in great part on the same reasons, is practically inseparable from it. Secondly, the principle requires liberty of tastes and pursuits; of framing the plan of our life to suit our own character; of doing as we like, subject to such consequences as may follow: without impediment from our fellow creatures, so long as what we do does not harm them, even though they should think our conduct foolish, perverse, or wrong.

[…]

Not that it is solely, or chiefly, to form great thinkers, that freedom of thinking is required. On the contrary, it is as much and even more indispensable to enable average human beings to attain the mental stature which they are capable of".

When digital interactions manipulate the freedom of attention, they also affect the freedom of thought: a right to freedom protected by the Universal Declaration of Human Rights and the European Convention of Human Rights (art. 9).

Taking the Spanish case as an example, although the Spanish Constitution does not explicitly mention this right, it can be assumed that its art. 20, which safeguards the freedom of expression, also protects the freedom of thought. In addition, as recalled by the Spanish Constitutional Court in sentence number 76/2019, the ideological freedom guaranteed by art. 16.1 of the Spanish Constitution has two dimensions: one is internal and involves the right to adopt a certain intellectual position before life and other life-related issues, and to represent or judge reality according to personal convictions. The other, the external one, is the dimension of *agere licere*: the right to act according to one's own ideas without incurring any penalty or demerit and without suffering compulsion or interference on the part of public authorities.

Therefore, it can be said that a —public or private— digital design that takes advantage of biases to manipulate and capture people's attention can undermine their constitutional

for 10 years and, as a result of this experience, he decided to leave the company to pursue a PhD at Oxford and conduct research on the philosophy and ethics of technology.

[35] "…when we use the term "attention" in day-to-day parlance, we typically mean what cognitive scientists call the "spotlight" of attention, or the direction of our moment-to-moment awareness within the immediate task domain", J. Williams, *Stand out of our Light*, 44-45.

[36] *The Guardian*, article published in 8 September 2021, underlines in relation to UK government that some studies show a growing government use of sensitive data to nudge behaviour. See *TechScape: Should government use the web to nudge our behaviour? | Technology | The Guardian,* in www.theguardian.com.

right to freedom of thought, which is intimately linked to the value of dignity and to the freedom of development of one's personality (following with the Spanish example, art. 10.1 of the Spanish Constitution).

The second reason is that this manipulation of attention implies hindering human will. There can be no human will without attention, because will, which is the faculty of deciding and ordering one's own conduct, can only exist if there is attention and absence of manipulation. From a legal standpoint, digital manipulation can thus be considered a threat to the freedom of individuals to establish rules of conduct for themselves and others within the limits of the law, hence, an impediment to the autonomy of the will safeguarded by many legal systems (using again the example of Spain, by art. 1255 of the Civil Code).

Thirdly, it should be noted that if freedom of thought and individual will be played upon through the digital manipulation of millions of people, then the general will is also affected, thus damaging democracy and the rule of law. Accordingly, in the decision cited above — and in many other similar cases— the Spanish Constitutional Court has stressed that, without freedom of thought, neither would there be a place for the fundamental principles of a legal system based on democratic values and the rule of law.

3.3. *An anti-Enlightenment project: modes of digital manipulation*

The attention economy and digital manipulation harm both people and social, legal and political systems. Their negative impact is therefore not trivial. The stakes are high.

In his abovementioned book, Williams[37] proposes a useful threefold distinction about this harmful effect to understand it better: the impact on "the doing" (that he calls spotlight), the impact on "the being" (on the values that guide us, which he calls starlight) and the impact on "the knowing" (which he calls daylight):

Digital manipulation and distractions of attention regarding "the doing". This is the typical loss of concentration due to digital designs aimed at distracting the individual (with the awareness that, after each distraction, attention is generally recovered in approximately 23 minutes). The author points out that the impact on "the doing" is not only individual: it can also have social significance, as it may affect political life. In fact, digital designs can distract from the relevant information that allows one to be politically informed. Among the examples provided by Williams there are practices implemented by China and the former US President Donald Trump. This type of manipulation, however, is not the only form in the digital environment and, although serious, it is not the most corrosive for democratic coexistence.

Digital manipulation and distractions of attention regarding "the being". This second type of digital distraction is aimed at making people lose their values through the promotion of pettiness (that is, the assignment of intrinsic value to goals and objectives with no intrinsic value, that are often marked by a poverty of spirit and short-sightedness, and which reveal lack of prudence), narcissism and social fragmentation, with the consequent erosion of values such as social cohesion.

Digital manipulation and distractions of rationality regarding "the knowing". According to the author this is the "epistemic distraction", which affects reflection, memory, prediction, calm, logic and goal-setting. The digital environment does this through fake news, impairment of intelligence and emotional capacities, by generating stress and other pathologies, by affecting reflection through notifications and applications, by promoting continuous moral indignation and by leading to dehumanization and populism. Accordingly, several studies have led Nicholas Carr to declare to the BBC that we are becoming less intelligent, more closed-minded and intellectually limited by technology.[38] A technique at the service of this impairment of the "daylight" is precisely the use of dark patterns already addressed in the previous section.

It should also be recalled that in a well-known article written in 1784, Kant pointed out that enlightenment is characterized by the decision and the courage to use one's own understanding without the guidance of others; the famous *sapere aude*, which can bring people out of a self-guilty dependence caused by laziness and cowardice.[39] Centuries later,

[37] J. Williams, *Stand out of our Light*, 2018.

[38] See Nicholas Carr: "Nos estamos volviendo menos inteligentes, más cerrados de mente e intelectualmente limitados por la tecnología", in *BBC News Mundo*.

[39] *Beantwortung der Frage: Was ist Aufklärung?*, often

these concepts have become topical again because of digital manipulation: a market-driven project for bringing us back to dependence by means of distractions that sap our attention, and hidden nudges that push us where we don't want to go.

Therefore, digital manipulation can be seen as an "anti-Enlightenment" project.

3.4. *Compulsion Incentivizing Technologies. The impact of Artificial Intelligence (AI)*

In other words, and in relation to what has been said in the previous section, digital manipulation acts by taking advantage of biases, by playing upon them, by exploiting and enhancing system 1 of thinking and by deactivating system 2.

This scenario is likely to get worse in the future, for at least two reasons. The first is the potential increase in available leisure and the consequent rise in consumption of technologies that incentivize compulsion.

Secondly, because of the impact of AI. In this regard, Yeung has introduced the concept of "hypernudge": nudging empowered by Big Data and algorithms that has the ability to move from the one-size-fits-all design to "tailored" —precision— nudges, which target specific individuals according to their specific characteristics through machine learning.[40]

Yeung warned that Big Data-driven nudging is agile, discreet and very powerful. It provides data holders with the ability to generate a highly personalized choice architecture by guiding people's decisions, no matter whether they are consumers or users of a public service. In fact, the author conceives hypernudges as instruments of control based on design and, to give a straightforward example about them, she pointed to the order of the results pages provided by search engines —e.g. Google, Bing etc. These instruments do not force us to look only at the first websites of the list —which happen to be also the most favourable for search engine marketing—, nor to forgo the other hundreds of thousands of websites, but that is exactly what we do, and the search engine knows that, because of our cognitive and temporal limitations.

The hypernudge is based on the highlighting of algorithmically determined correlations between elements of data that human cognition cannot observe, not even with the help of standard computing technology. This confers an undisputed prominence to the highlighted data patterns, as they allow the dynamic configuration of the informational choice of the user and her/his decisions to be swayed by taking advantage of priming: the psychological effect whereby the exposure to one stimulus —e.g. images, sounds, words etc.— influences the response to a subsequent stimulus, hence also future behaviours and actions.

Big Data-driven nudging can be very useful in medicine[41] and public services management,[42] but also in fields like tax compliance and tax administration, as proved by the example of the Strategic Plan 2020-2023 designed by the Spanish Tax Agency.[43]

Nevertheless, it should not be overlooked that Big Data-driven nudging can also put people's rights at risk.

This is particularly evident in terms of personal data protection, as remarked by the District Court of The Hague at the beginning of 2020 with an express mention of art. 8 of the European Convention on Human Rights (ECHR).[44] In the case in question, SyRI, the Dutch algorithmic System for Risk Indication, and the public authorities that implemented and managed it, stood —and were— accused of having collected, for several years, a disproportionate amount of taxpayers' personal data —on income, pensions, insurance, type of house, taxes, fines, integration, education, debts and

referred to simply as "What Is Enlightenment?", is a 1784 essay, published in December 1784 in the *Berlinische Monatsschrift* (*Berlin Monthly*).

[40] K. Yeung 'Hypernudge': Big Data as a mode of regulation by design, Information, in Communication & Society, vol. 20, issue 1, 2017, 118.

[41] D. Misawa, J. Fukuyoshi and S. Sengoku, Cancer Prevention Using Machine Learning, Nudge Theory and Social Impact Bond, in International Journal of Environmental Research and Public Health, vol. 17, No. 3, 2020, 790.

[42] J. Ponce, El derecho a una buena administración y la personalización de los servicios públicos. Sesgos, "nudging" e inteligencia artificial in B. Puentes Cociña and A. Quintiá Pastrana (eds.), El derecho ante la transformación digital: oportunidades, riesgos y garantías, Barcellona, Atelier, 2019, 51.

[43] See the *Strategic Plan 2020-2023* of the Spanish Tax Agency at: *adenda_plan_objetivos.pdf*, in www.agenciatributaria.es.

[44] On 5 February 2020, the District Court of The Hague (*Rechtbank Den Haag*) held that the System Risk Indication (SyRI) algorithm system, a legal instrument that the Dutch government uses to detect fraud in areas such as benefits, allowances, and taxes, violates article 8 of the European Convention on Human Rights (ECHR) (right to respect for private and family life).

unemployment benefits— to calculate who was more likely to defraud the welfare system

Personal data, however, are not the only things at stake. Yeung highlighted that manipulation and deception are another two critical issues and that users' acceptance of information and requests for consent for the use of digital environments are not apt to solve them. This ties in with the potential violation of the right to freedom of thought and of the democratic principles already mentioned.

It is therefore clear that adequate mechanisms to prevent these serious digital risks must urgently be designed.

3.5. *What should —and should not— be done?*

In the light of the situation described above and of the likelihood of dangerous future developments, it is useful to consider the *Onlife Manifesto* funded by the European Commission,[45] which emphasizes that:

"In the digital economy, attention is approached as a commodity to be exchanged on the market place, or to be channelled in work processes. But this instrumental approach to attention neglects the social and political dimensions of it, i.e., the fact that the ability and the right to focus our own attention is a critical and necessary condition for autonomy, responsibility, reflexivity, plurality, engaged presence, and a sense of meaning. To the same extent that organs should not be exchanged on the market place, our attentional capabilities deserve protective treatment. Respect for attention should be linked to fundamental rights such as privacy and bodily integrity, as attentional capability is an inherent element of the relational self for the role it plays in the development of language, empathy, and collaboration. We believe that, in addition to offering informed choices, the default settings and other designed aspects of our technologies should respect and protect attentional capabilities".

As already pointed out, defending attention from manipulation and deception means defending freedom of thought and human will, both at the individual and at the collective level. This is a political task that requires a prior reform of the current totalitarian system of information technologies, because digital design is the politics behind the politics.

Before proposing concrete measures, there are some actions and attitudes that should be avoided in order to face the described dangers. "Doing nothing" is the first inadvisable posture, because the existing evidence suggests the need to take an active and precautionary approach (based on the precautionary principle) towards technologies, especially in the social sphere. Neither can the problem be solved by advising users to disconnect or adapt to the current situation. Moreover, we consider it unwise and inconvenient to rely only on technological companies' self-regulation and ethics, as these can be just a facade and an attempt to push aside the law.[46]

Then what should be done? The solution lies in the introduction of incentives for technology design that benefit consumers and users and contribute to making technologies more human.

The main interventions that could help move the attention economy in the right direction are: (1) rethinking the nature and purpose of advertising, (2) conceptual and linguistic reengineering, (3) changing the upstream determinants of design, and (4) advancing mechanisms for accountability, transparency and measurement.

Advertising: in this field, Williams, in his book, suggests make ad blocking software mandatory and activated by default, with users being able to unblock it if they wish. Thus, although he does not use the term "nudge", this is exactly what he means by proposing a default option or, more specifically, an opt-in: not receiving advertising unless I choose to receive it. This is an important proposal that leads us to reflect as well on the legal battle that has been going on for years against Internet ad-blocking applications.[47]

Language: Williams identifies various terms related to the language of digital persuasion, which he groups into triads, from lesser to greater impact on people's attention and will: invite-tempt-seduce / suggest-persuade-demand / direct-guide-drive. It is necessary to make progress in specifying what

[45] See L. Floridi, The Onlife Manifesto. Being Human in a Hyperconnected Era, Berlin, Springer, 2015,13, § 4.6, open access at: The Onlife Manifesto | SpringerLink.

[46] See the interview with Prof. Karen Young at *AI and the law*, in www.birmingham.ac.uk.

[47] A.M. Russell, *The Legal Fate of Internet Ad-Blocking*, in Boston University Journal of Science & Technology Law, 24, 2018, 299.

these contexts of persuasion are, because, as Wittgenstein said, "the limits of my language mean the limits of my world (*Tractatus Logico-Philosophicus*, proposition 5.6)." In order to achieve a more human design of technologies it is therefore overwhelmingly important to name, classify and reflect on digital nudges.[48]

Modification of technological design: policymakers should have a fundamental role to play in responding to the crisis of the digital attention economy. We can take inspiration from pre-digital regulations on junk mail or telemarketing calls, which, after all, tried to avoid unwanted intrusions into private life. Transparency about the objectives of digital design is paramount.[49]

Examples of how policymakers and judges can protect citizens from manipulation and digital deception are already observable. Some of them can be found in the decision of 1st October 2019 from the Court of Justice of the European Union (CJEU), Grand Chamber, in which it was established that the existence of genuine consent implies avoiding ticked boxes by default, in accordance with EU Law (Privacy Directive 2002/58/EC, art. 6 GDPR and Directive 95/46/EC).[50]

Another interesting line of intervention could be creating digital media platforms that could play a similar role to the one that public broadcasting has played in television and radio. In accordance with this approach, which regards the provision of digital public services to counteract the aforementioned manipulations and increase the "lights" (that nonetheless should be controlled for their potential to create similar risks), it is worth drawing attention to the Italian experience ITsART.[51] This is a new platform promoted by the Italian Ministry of Culture and *Cassa Depositi e Prestiti* (Italy's deposits and loans fund) for world-wide distribution of artistic and cultural content in digital form. The business partner of the project is CHILI Spa, a company selected for its industrial and technological know-how. ITsART is managed through a company with 51% public shareholding; a public-private partnership in which CHILI Spa only owns 49% of shares.

The latter proposal challenges the widespread idea that the state should always withdraw from the provision of public services and become a mere guarantor or regulator, given its shortcomings and inadequacies vis-à-vis the private sector. This is neither true nor necessary in all cases, unless it is advocated with a specific ideological goal, as the Nobel Prize winner Herbert Simon pointed out some time ago.[52] Avoiding such an ideological bias is crucial, as well as analysing on a case-by-case basis if the intervention of Administrations is to become necessary and apt to serve the general interest, both in the digital world and outside it. The idea of a formal democracy as a guarantor of formal rights and freedoms must give way to a material democracy that enables everyone to enjoy such rights and freedoms on an egalitarian basis; something that would be impossible without reinforcing the principle of equality. Freedom without equality is an empty concept. Hence the need for governments to direct economic life and to strive for the achievement of the maximum general welfare.

Accountability and measurement: although blaming designers for lowering our "lights" is unwise —as it is the result of a systemic functioning that incentivizes manipulation—, the introduction of a professional oath for digital designers, similar to the Hippocratic Oath, may be a good option, according to Williams. However, he also admits that its implementation would not be free of complications, especially due to the plurality of professions involved in digital design, including people without specific training, and

[48] J. Williams, *Stand out of our Light*, 114: "Clarifying the language of persuasion will have the added benefit of ensuring that we don't implicitly anchor the design ethics of attention and persuasion in questions of addiction", which is a core problem, but also "a convenient distraction from deeper questions about a design's fundamental purpose".

[49] In his book, William goes so far as to propose the introduction of a fee for exceeding certain levels of "attention offsets". This idea, which implies punishing companies for provoking intentional harm, was not further detailed and developed by the author; nevertheless, it clearly reflects the important role that the law should play.

[50] It is precisely due to these regulations that, in 2014, Spain added art. 60 *bis* to the Royal Legislative Decree 1/2007, of 16th November, which approved the revised text of the General Law for the Defence of Consumers and Users and other complementary laws. In particular, art. 60 *bis* establishes that consumers and users are entitled to the reimbursement of additional payments charged by the trader without their express consent through default options.

[51] See www.ITsART.tv.

[52] As the Nobel Prize winner Herbert Simon pointed out some time ago, H.A. Simon, *Why Public Administration?,* in *Journal of Public Administration Research and Theory*, vol. 8, issue 1, January 1998, 1.

the lack of professional associations.[53]

3.6. *Inadequacies of ethics and self-regulation (even regulated): lobbies and regulation*

It is now clear that the combination of digital design, *nudges* (including dark patterns), exploitation of cognitive biases, Big Data and AI can create an explosive cocktail for citizens' freedom and free will and for the functioning of social and democratic states governed by the rule of law. Nevertheless, we should not "throw the baby out with the bathwater" and deny or waste the potential of all these techniques and technologies to serve the general interest. It seems clear that threats to democracy and people's rights described above cannot be tackled merely through private companies' self-regulation and enthusiastic calls for ethics; just as serious illnesses cannot be cured with love and prayer alone.

The role of law and, within it, of "positive nudges" in defence of citizens is an issue more pressing than ever. In Europe we have already had bitter experiences with self-regulation — including regulated self-regulation—, for example, in the banking sector. This was made clear by the Great Recession and the European Commission recognized it, by pointing out that financial actors have wrongly determined their actions and business policies with dire consequences.[54]

Hopefully, the same mistakes will not be repeated in the digital sphere, and the frustrations generated by the ineffectiveness of self-regulation will be learnt from, as in the case of the EU Disinformation Code mentioned in the previous section. This example of self-regulation has been assessed by the European Commission, which considers that the assessment "has revealed significant shortcomings. These include inconsistent and incomplete application of the Code across platforms and Member States, limitations intrinsic to the self-regulatory nature of the Code, as well as gaps in the coverage of the Code's commitments. The assessment also highlighted the lack of an appropriate monitoring mechanism (…), lack of commitments on access to platforms' data for research on disinformation and limited participation from stakeholders, in particular from the advertising sector". Therefore, the European Commission concludes that it is necessary "to transform the Code into a stronger instrument for addressing disinformation and creating a safer and more transparent online environment".

In the same line, the OECD has pointed out that "Industry self-regulation can be an advantageous complement to government policies, but it also poses a number of challenges" and that "the use of ISR to help address consumer issues needs to be considered systematically when policy makers and enforcement authorities are developing options for taking action. As discussed in the Consumer Policy Toolkit, ISR could be part of a multi-faceted response to a problem, supporting other measures that governments might take. With respect to the development, monitoring and evaluation of such mechanisms, it appears that stakeholder involvement has been limited, and that it may be beneficial to explore whether there are ways that involvement could be strengthened, in ways that would benefit all stakeholders".[55]

Although powerful market forces opposed to regulation in the general interest will probably continue to act as lobbies against a stronger regulation when and if necessary, it is a matter of being timely and avoiding large-scale opaque and negative psychological mutation of consumers, citizens and democratic political systems.

The future European regulation on AI offers an ideal opportunity to discuss these issues, and hence to go beyond the necessary but insufficient concern for personal data protection.[56] Title II of the proposal of regulation made public in April 2021 sets out a list of prohibited AIs. This draft regulation follows a risk-based approach by differentiating between AI uses that constitute (i) unacceptable risk, (ii) high risk, and (iii) low or minimal risk. The list of prohibited practices included in Title II comprises all

[53] To overcome them, the author suggests how to elaborate the oath and what content it should have, making a concrete proposal: J. Williams, *Stand out of our Light*, 120.

[54] European Commission, Green Paper - Corporate governance in financial institutions and remuneration policies {COM(2010) 286 final} {SEC(2010) 669}.

[55] OECD, Industry Self-Regulation: Role and Use in Supporting Consumer Interests, 2015, available at: Industry Self-Regulation: Role and Use in Supporting Consumer Interests, in www.oecd.org.

[56] See the regulation at Proposal for a Regulation laying down harmonised rules on artificial intelligence | Shaping Europe's digital future, in www.europa.eu.

those AI systems whose use is considered unacceptable because they contravene EU values; among them, "the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm" (art. 5.1.a)

It should be also noted that the proposal points out that other manipulative or exploitative practices facilitated by AI systems could be covered by data protection, consumer protection and digital services legislation ensuring that individuals are properly informed and can freely choose not to be subjected to profiling or other practices that may affect their behaviour.

The reference to physical or psychological harm, however, is not particularly appropriate, given the significance of digital designs in relation to potentially manipulative AI. This should be replaced by a simple mention of the possibility of causing or inducing error or deception, thereby affecting the autonomy of the will. In this regard, it was mentioned in the previous section about the recent amendment of the California Privacy Rights Act (CPRA) of 2020 to ban dark patterns.

Accordingly, it would also be worth reformulating the European draft regulation by including a ban on any AI system that deploys subliminal techniques beyond a person's consciousness in order to distort her/his behaviour to subvert or impair her/his autonomy, decision making or choice. This is the line of the amendments to the draft of the Digital Services Act introduced by the European Parliament in January 2022, banning dark patterns, as we have seen above.

4. *Conclusions*

The use of behavioural insights in the digital domain has become extremely significant during the COVID-19 pandemic. Although digital nudging can be useful for making effective the right to good administration, it can create unacceptable manipulations. In this area, the possible use by the public and private sector of the so-called dark patterns, concerning which the European Parliament has recently proposed to include a ban in the future Digital Services Act, and what is known as hypernudging raises legal doubts regarding a possible violation of freedom of thought, as indicated by the Committee of Ministers of the Council of Europe in a recent statement of 2019. The future Digital Services and Artificial Intelligence regulations could and should introduce provisions avoiding the worst effects of digital manipulation.

The door is open to use the best of artificial intelligence and to avoid the worst, through reasonable EU and national regulations avoiding that we, the citizens, become digital *zombies* in the hands of governments and corporations.