Digital Vulnerability and Manipulation in the Emerging Digital Framework

Sax, M.; Helberger, N.

Publication date
2024
Document Version
Final published version
Published in
Digital Fairness for Consumers

II. Digital Vulnerability and Manipulation in the Emerging Digital Framework

Marijn Sax & Natali Helberger[1]

1 Institute for Information Law, University of Amsterdam. Both authors contributed equally.

Introduction and recap: Digital vulnerability

The use of the term '*digital* vulnerability', as opposed to just 'vulnerability', highlights how our technological circumstances require us to adopt a more dynamic approach to vulnerability. It no longer suffices – if it ever did at all – to think in terms of stable, permanent characteristics or circumstances that render a person vulnerable. A *digital* vulnerability approach is based on the insight that not only fixed characteristics of a person can render her vulnerable, but that in the continuous interplay with one's (digital) environment one can – sometimes only momentarily, or only in specific contexts – move in and out of states of vulnerability. As a result, the classic distinction between the 'normal, non-vulnerable' versus 'the vulnerable' consumer is collapsed. Every consumer is potentially vulnerable, depending on the (digital) circumstances and environments she finds herself in. Vulnerability becomes the rule, rather than the exception.

With special attention for the ways in which vulnerability plays out in digital choice environments also increasingly comes attention for manipulative influences exert in and through digital choice environments. Manipulative influences are exerted precisely by the targeting and exploitation of known or presumed vulnerabilities in order to (try to) make manipulation targets serve the ends of the manipulator.[2] It therefore comes as no surprise that in the EU's recent digital technology legislative agenda manipulation and vulnerability are often mentioned and addressed in close connection to one another.

The thematisation of (digital) vulnerability and manipulation in close connection to one another is promising, but also comes with challenges. The main challenge is that precisely because vulnerability and manipulation are so closely related, it is especially important to both conceptually and definitionally highlight not only the similarities but also the differences. Both vulnerability and manipulation are complicated concepts in their own right. Their interrelation is even more complex. The aim of this piece is to provide a first exploration of how (digital) vulnerability and manipulation, as separate concerns but also increasingly as interrelated concerns, play a role in the emerging digital framework of the Digital Services Act (DSA), Digital Markets Act (DMA), draft AI Act (AIA), draft Political Advertising Regulation (PAR).

Elsewhere, we have defined digital vulnerability as "a universal state of defencelessness and susceptibility to (the exploitation of) power imbalances that are the result of increasing automation of commerce, datafied consumer-seller relations and the very architecture of digital marketplaces."[3] And we argued that digital vulnerability is related to the power or ability of commercial actors to affect the decisions, desires, and behaviour of the consumer in ways that the consumer, all things considered, does not condone, but are also not in a position to prevent. That this is so is the result of what Kaptein *et al.*[4] have referred to as an "adaptive persuasive system". In more concrete terms this means that to be able to evaluate commercial practices in terms of their fairness, it is not enough to evaluate the message; the systemic set-up and the way technology shapes the relationship between consumer and advertiser should also figure

2 Sax, M. (2021). *Between Empowerment and Manipulation: The Ethics and Regulation of For-Profit Health Apps*. Kluwer.

3 Helberger, N., Lynskey, O., Micklitz, H.-W., Rott, P., Sax, M., Strycharz, J. (2021). *EU Consumer Protection 2.0: Structural Asymmetries in Digital Consumer Markets*, report for BEUC - The European Consumer Organisation; Helberger, N., Sax, M., Strycharz, J., & Micklitz, H.-W. (2022). Choice Architectures in the Digital Economy: Towards a New Understanding of Digital Vulnerability. *Journal of Consumer Policy, 45*(2), 175–200.

4 Kaptein, M., Markopoulos, P., De Ruyter, B, & Aarts, E. (2015). Personalizing Persuasive Technologies: Explicit and Implicit Personalization Using Persuasion Profiles. *International Journal of Human-Computer Studies, 77*, 38–51.

prominently in such an analysis. We have therefore argued that addressing vulnerability and tackling systemic structures of commercial manipulative exploitation of those vulnerabilities is not just a question of consumer empowerment but of changing markets and addressing the digital asymmetries that enable those practices in the first place.

Definition of (digital) vulnerability across the emerging digital regulatory framework

Three elements, in particular, characterise the concept of digital vulnerability: its relational nature, its architectural nature, and the erosion of privacy. In the digital realm, consumer vulnerability can be the result of asymmetrical and potentially continuous power relationships where the productive force of those relationships can produce vulnerabilities. Vulnerability is thus inherently **relational**. As such, vulnerabilities can be not only the result of individual characteristics or the social or economic position of the consumer but also the result of the properties of a digital platform, app store or another form of digital choice architecture. Digital environments can be data-driven, dynamically adjustable and designed to infer or even create vulnerabilities. The **architectural** make-up of those digital environments – i.e., their entire technology stack – can thus be geared towards the production and exploitation of vulnerabilities. Finally, there is the element of extraction and use of exploitative data practices to segment, classify, profile and target individuals. Part of the design of digital choice architecture is finding ways to collect (more and more) data about consumers, data that can be used to target and personalise services. The interaction of consumers with these services generates new data that will flow into the system, help to adjust and optimise it, make it more responsive to the explicit or inferred signals from consumers, and ultimately the business goals that inform the overall design of the choice architecture. A structural disregard for consumer **privacy** is thus, again, an essential productive force for digital vulnerability. After having signalled the importance of understanding digital vulnerability and addressing digital vulnerability and manipulation (as an element and condition of digital asymmetry), the following section examines a) the extent to which the emerging digital framework accommodates and responds to such a more comprehensive understanding, b) gaps and inconsistencies, as well as c) possible need for improvement.

Broadening a traditional concept

The emerging digital framework (DSA, DMA, AIA, PAR) is divided. On the one hand, the traditional concept of vulnerability in the sense of a pre-defined group of consumers as an exception to the rule (average consumer) is still dominant. Particularly in the AI Act but also the DSA, there are numerous references to the elderly, minors and disabled as traditionally recognised groups of vulnerable consumers.[5] Having said so, several developments are noteworthy. First of all, the proposals for the AI Act from the Council, the European Commission and the European Parliament add new categories of vulnerable users, including migrants,[6] persons living in poverty, ethnic or religious minorities,[7] and people applying for or receiving public assistance, services or benefits.[8] Like the traditional concept of vulnerability in consumer law that singles

5 See, e.g., Art. 52a(3b) EP version: Information shall be accessible to vulnerable persons, including persons with disabilities or children.
6 Recital 16 of the AI Act in the draft version from 21/01/2024.
7 Ibid.
8 Recital 37 AIA in the draft version from 21/01/2024.

out members of particular groups in society, also the AI Act singles out particular groups of users and designates them as potentially vulnerable. Unlike the vulnerable consumer concept, however, vulnerability under the AI Act and the DSA can be found in commercial relationships as well as in consumers' relationships with public institutions (for example, as receivers of public benefits). The AI Act explicitly acknowledges that vulnerability can be the result of a dependency situation. Another and related direction in which the concept is broadened is that vulnerability implies that consumers are not only susceptible to the infringement of their rights as consumers (information, fair prices, choice, being free from harm), but also to the infringement of fundamental rights and their legitimate interests as citizens. In this context, one fundamental right in particular stands out, which is the right to non-discrimination. Both the DSA and the AI Act conceptualise consumer vulnerability repeatedly as susceptibility to undue discrimination or biases as the result of the use of digital technology.[9] And whereas consumer vulnerability has been typically referred to in the context of harm for individual consumers, in the emerging digital framework, vulnerability and the exploitation of vulnerabilities can also extend to harm to society. For example, in the draft PAR (EC version), the Commission first explains how digital technology can be used to segment individuals and exploit their characteristics or vulnerabilities to explain then that this can have detrimental effects on individual citizens' fundamental rights and freedoms (such as the right to data protection, to make political decisions and exercise voting rights), but that this can also negatively impact the overall democratic process "as it enables a fragmentation of the public debate about important societal issues, predatory voter analysis, selective outreach and, ultimately, the manipulation of the electorate", next to increasing the risk for disinformation and foreign electoral interference.[10] Recital 69 DSA reads: "In certain cases, manipulative techniques can negatively impact entire groups and amplify societal harms, for example, by contributing to disinformation campaigns or by discriminating against certain groups. Online platforms are particularly sensitive environments for such practices and they present a higher societal risk." Put differently, a broader understanding of vulnerability emerges from the digital regulatory framework. Vulnerability is used to refer to the situation of users as consumers but also as citizens. Their freedoms and fundamental rights are at stake, as are the interests of society as a whole, when exploiting vulnerabilities results in collateral harm to societal values such as democracy or an inclusive society. This broader understanding of vulnerability reflects the reality of the digital environment and of platforms in particular, where it becomes increasingly difficult to draw a clear distinction between consumers and citizens. Platforms in particular serve as both economic marketplaces and privately controlled forums of public debate and engagement. Neither do their algorithms and ad auction systems distinguish between the citizen and the consumer.

Emerging new understandings of vulnerability

Next to a more traditional conception of vulnerability in the DSA the AI Act and the PAR (the DMA does not refer to vulnerability), a new approach to user vulnerability can be observed, too. This is most apparent in the proposals for an AI Act and here in Recital 16 and the corresponding Articles 5(1) a) and b). According to Recital 16 (EP version), "AI-enabled manipulative techniques can be used to persuade persons to engage in unwanted behaviours, or to deceive them by nudging them into decisions in a way that subverts and impairs their autonomy, decision-making

9 E.g., recital 44 AIA, recital 69 DSA, Recital 47 PAR (EP version).
10 Recital 47 (EP version) PAR.

and free choices".[11] The draft law furthermore, acknowledges that such exploitation can have a temporal component by referring to harms that may be accumulated over time, thereby pointing towards a relational understanding of vulnerability in the sense of our definition of digital vulnerability. The proposed act also addresses the use of newer AI to *make* users vulnerable, in the form of using "machine-brain interfaces or virtual reality as they allow for a higher degree of control of what stimuli are presented to persons, insofar as they may be materially distorting their behaviour in a significantly harmful manner". Though the provision has a distinctive 'cyberpunk' feel to it, it does clearly acknowledge that vulnerability is not necessarily inherent to the consumer but can be optimised for. Finally, unlike in consumer law, the proposed AI Act (in the EP version) intends to protect users from economic and all kinds of harm (whereas the Council and EC version focus on physical or psychological harm).

The emerging digital framework also acknowledges that vulnerability can be the result of the design and deployment of AI systems or platforms. Even though an earlier proposal of the European Parliament to include an obligation for national supervisory authorites to investigate the design goals has not made it into the later version of the text of Article 65 AIA: "Where there is sufficient reason to consider that an AI system exploits the vulnerabilities of vulnerable groups or violates their rights intentionally or unintentionally, the national supervisory authority shall have the duty to investigate the design goals, data inputs, model selection, implementation and outcomes of the AI system." According to Recital 69 of the DSA, "Online platforms are particularly sensitive environments for such practices [targeting techniques optimised to match users interests and appeal to their vulnerabilities" and they present a higher societal risk."

Data, or the extraction and use of data, is also explicitly considered as a potential source of vulnerability. According to Recital 47 PAR (EC Version): "On the basis of the processing of personal data, in particular data considered sensitive under Regulation (EU) 2016/679 of the European Parliament and of the Council and Regulation (EU) 2018/1725 of the European Parliament and of the Council, different groups of voters or individuals can be segmented and their characteristics or vulnerabilities exploited for instance by disseminating the advertisements at specific moments and in specific places designed to take advantage of the instances where they would be sensitive to a certain kind of information/message." These examples signal that a process of rethinking of vulnerability has begun in the sense of a more relational, architectural and data-reliant conceptualisation of vulnerability.

Finally, and unlike in consumer law, where the concept of vulnerability is, in the first place, a benchmark or vantage point from which to assess a particular technology, the emerging digital framework has begun to attach legal consequences to the exploitation or causation of digital vulnerabilities (or the potential thereof). For example, the potential to cause or exploit vulnerabilities can be part of the assessment of whether an AI system is high risk or not,[12] digital vulnerability can trigger the need to undertake mitigation measures and improve systems design,[13] and can be central to the ban of particular uses of digital technology.[14]

---

11 Recital 16 (EP version) AIA.
12 Article 7f AIA.
13 Art. 29a AIA (EP version), Articles 34 and 35 DSA.
14 Art. 5 (1) AIA, Article 12 (1) PAR.

Definition of manipulation across the emerging digital regulatory framework

In recent years, the concept of manipulation has found its way into the European legislative agenda for the regulation of the digital economy. The increasing interest in manipulation as a regulatory concern is closely tied to the mission of protecting vulnerable consumers, since manipulation is typically predicated on the exploitation of vulnerabilities. So, in a digital landscape where many of consumers' interactions with commercial parties take place within digital choice environments – which are especially well suited to track, analyse, and influence behaviour – the risk of manipulation and concerns over digital vulnerability are two sides of the same coin. Legislative initiatives to address manipulation are thus also initiatives that have a direct impact on the legislative approach to vulnerability.

As the recent surge in philosophical literature on (digital) manipulation clear shows, manipulation is difficult to define.[15] These definitional challenges do, of course, carry over to the legal context. So much so that even though several recent legislative initiatives – e.g., the DSA and the AI Act – contain explicit manipulation clauses, none of these legal instruments contain a legal definition of manipulation. It thus remains unclear how manipulation should be interpreted as a *legal* concept in the EU's legislative agenda on the digital economy.

The lack of a proper definition in the DSA

Article 25(1) of the DSA contains a straightforward manipulation ban:

> "Providers of online platforms shall not design, organise or operate their online interfaces in a way that deceives or manipulates the recipients of their service, or in a way that otherwise materially distorts or impairs the ability of the recipients of their service to make free and informed decisions".

Even though manipulation is explicitly mentioned, it remains unclear what it means in the context of Article 25(1). The structure of this article is best understood by starting at the end of the article. That which is ultimately safeguarded is people's ability "to make free and informed decisions". To that end, the article mentions two specific forms of influence – manipulation and deception – that can "distort or impair" free and informed decisions, while also acknowledging that there can be other "ways" in which people's free and informed decisions can be undermined. Lastly, there is the open-ended, very inclusive formulation of "shall not design, organise or operate their online interface in a way that deceives or manipulates". Clearly, the DSA aims to address the digital choice environments that betray a manipulative potential *in their entirety*.

With no definition of manipulation being mentioned elsewhere in the DSA, this specific manipulation clause does little to explicate what manipulation means in this context. It is clear that manipulation is understood as a form of influence that can impair free decision-making, but

15 Susser, D., Roessler, B., & Nissenbaum, H. (2019). Online Manipulation: Hidden Influences in a Digital World. *Georgetown Law Technology Review, 4*(1), 1–45; Sax, M. (2021). *Between Empowerment and Manipulation: The Ethics and Regulation of For-Profit Health Apps*. Kluwer; Jongepier, F., & Klenk, M. (Eds). (2022). *The Philosophy of Online Manipulation*. Routledge.

that can be said of many different forms of influence – coercion and blackmail also impair free decision-making but are clearly not cases of manipulation (or deception). So, the current framing of manipulation is too generic to be helpful. One could turn to Recital 67 which deals specifically with dark patterns. Dark patterns are not mentioned in any of the articles in the DSA, so the recital on dark patterns is the most plausible source of guidance for understanding manipulation in the DSA.

This recital, however, does little to explain how – conceptually speaking – influences such as deception, nudging, nagging, and manipulation are be understood and, importantly, differentiated from each other. The recital mentions these different forms of influence, seemingly as examples of dark patterns. It remains unclear, however, whether the concept of dark patterns is treated as just an umbrella term for several types of influences (deception, manipulation, nudging, nagging) that can somehow distort decision-making. Because no (approximations of) definitions of these forms of influences are provided, we only know that in the context of Recital 67 – and the DSA more generally? – these forms of influence are somewhat similar to each other because they are all collected under the umbrella of 'dark patterns'. What also doesn't help the reader is the fact that the DSA only contains vague gestures to what makes these different forms of influence undesirable. In Recital 67, deception, manipulation, nudging, and nagging are all described as forms of influence that impair/distort/unreasonably bias the decision-making of the consumer. If anything, this makes it even more unclear how one should differentiate between these forms of influence because they all seem to share the same wrong-making feature.

One is also kept wondering why Recital 67 is explicitly framed in terms of dark patterns, with deception, manipulation, nudging, and nagging seemingly being specific instances of dark patters, but why Article 25(1) is not framed in terms of dark patterns and *only* mentions manipulation and deception. Should deception and manipulation in Article 25(1) be read as incomplete short hands for dark patterns? If so, why aren't nudging and nagging included? If, however, manipulation and deception should not be read through the lens of the dark patterns recital, then where is one supposed to gather the interpretational resources to understand what is meant by these terms in Article 25(1)?

In sum, even though manipulation is explicitly mentioned in the DSA, it remains unclear how this challenging concept should be understood. And the DSA offers little 'interpretative materials' to work with in this regard.

AI Act: strong and limited harm focus

Manipulation is also mentioned explicitly in the AI Act. Article 5(1)(a) forbids

> "The placing on the market [or] putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques with the objective to or to the effect of materially distorting a person's or a group of persons' behaviour by appreciably impairing the person's ability to make an informed decision, thereby causing the person to take a decision that that person would not have otherwise taken in a manner that causes or is likely to cause that person, another person or a group of persons significant harm."

The similarities with the DSA are clear. Again, manipulation is mentioned as form of influence that can distort someone's decision-making capabilities. In the AI Act, there is a harm requirement added to the article. This is especially interesting because, again, like the DSA, the AI Act does not define manipulation anywhere. So, without a definition of manipulation and, as a result, without guidance on what sets manipulation apart from deception and other forms of influence, it is also challenging to formulate what type of harm manipulation is or can result in.

The connection to (digital) vulnerability is made especially clear in Article 5(1)(b). This article repeats the language of Article 5(1)(a), but instead of mentioning "purposefully manipulative and deceptive techniques, the Article 5(1)(b) mentions

> "An AI system that exploits any of the vulnerabilities of a person or a specific group of persons, [...]".

From a digital vulnerability perspective, the phrasing of this article seems promising since it explicitly moves beyond the 'non-vulnerable average consumer *versus* the vulnerable consumer' framing. Explicit attention is paid to the ways in which not only (semi-)permanent characteristics, but also particular (temporary) situations can render people (temporarily) vulnerable. What is curious, however, is the decision to draft two separate articles, one on manipulative AI systems (5(1)(a)) and one on AI systems exploiting vulnerabilities (5(1)(b)). This raises the question how the relationship between manipulation and vulnerability is understood by the legislator. Most philosophical conceptualizations of manipulation emphasize how manipulation is predicated precisely on the deliberate exploitation of vulnerabilities in order to make targets serve the ends of the manipulator. Seen from this perspective, it would stand to reason to see the threat of manipulative AI systems and the threat of AI systems exploiting vulnerabilities as one and the same threat – whenever an AI system is manipulative, it will necessarily also (seek to) exploit vulnerabilities. The fact that the exploitation of vulnerability is explicitly mentioned in a separate article, raises the question how manipulation is understood by the legislator if manipulation is also seen as distinct from the exploitation of vulnerabilities?

With the harm requirement present in the AI Act, a further question raised is whether we should be (mainly) concerned with *manipulation itself* as a harm, or with harms that can be the result of manipulation. Because we lack an underlying theory of manipulation, it also remains unclear how the relationship between manipulation as an undesirable form of influence and harm should be understood.

Critical commentary

A common thread in both the DSA's and the AI Act's incorporation of manipulation and vulnerability language is the inability to provide conceptual clarity on not only the meaning of these concepts, but also their interrelation. The dominant philosophy seems to be that in order to regulate phenomenon X, one should explicitly mention that X is in fact forbidden. If that approach is adopted without doing the necessary underlying conceptual work, this seemingly straightforward approach is bound to fail. At a minimum, there should be some clarity on what it is about the phenomenon of manipulation that makes it manipulation. Because if it is not clear how manipulation differs from, e.g., persuasion, deception, or dark patterns, it also does not help to write the term manipulation into law as the term will not help us distinguish between phenomena.

The urge to mention manipulation explicitly in new legislation is understandable given the popularity of the term in popular critical discourse. However, precisely because of the conceptual and definitional unclarities around the term, it could be wiser to opt for an approach where manipulation is addressed indirectly. One can, for instance, address some of the possible (necessary) preconditions for manipulation without explicitly conceptualizing manipulation. Another option is to rely on existing legal concepts that can be used to 'capture' manipulation concerns.

The Unfair Commercial Practices Directive (UCPD) is an informative example. This Directive precedes the recent turn to manipulation as a dominant concern in the digital economy, so the directive doesn't thematise or even mention manipulation. Still, the Directive contains a lot of interpretational resources to address manipulation in the consumer-vendor relationship.[16] Put briefly, the core aim of the UCPD is to "keep and maintain the consumer's autonomy".[17] If there is one value that is clearly threatened by manipulation, it is personal autonomy. So, manipulation worries are very relevant in the UCPD framework. If we look at the specific articles of the UCPD, especially Articles 8 and 9 dealing with aggressive practices contain a lot of material to understand and capture manipulative commercial influences. For example, the concept of 'undue influence'[18] plays a key role, as does the circumstance of vendors using an asymmetrical power relation to apply to undue influence to exploit vulnerabilities or circumstances of consumers. The attentive reader will have already noticed that without ever mentioning the concept of manipulation, the UCPD approach to aggressive commercial practices already captures most of the elements of a manipulation relationship.

All of this is not to say that the DSA and the AI Act should have opted for the UCPD approach. The UCPD is merely meant to show that one does not have to explicitly mention manipulation to capture manipulation worries.

Challenges and Potential Shortcomings of the Current Approach

The digital vulnerability framework highlights the relational and architectural nature of vulnerabilities in the digital economy. This perspective adds a certain dynamism and fluidity to (the approach to) vulnerabilities. Consumers can move in and out of states of vulnerability, and different digital environments can either trigger or exploit vulnerabilities differently. This dynamism is mirrored in our thinking about manipulation. The design of manipulative influences through digital choice environments is greatly helped by the agile nature of those same choice environments, which allows for the constant explorative search and iterative testing of the most efficacious manipulation techniques.

The architectural nature of digital vulnerabilities – and their exploitation for manipulative influences – also implies the importance of critically analysing the organizations that enable digital

16 For an elaborate analysis along these lines, see Sax, M. (2021) *Between Empowerment and Manipulation*. Kluwer.
17 Micklitz, H.-W. (2006). The General Clause on Unfair Practices. In. G. Howells, H.-W. Micklitz & T. Wilhelmsson (Eds.), *European Fair Trading Law: The Unfair Commercial Practices Directive* (pp. 83–122), p. 104.
18 Defined in Art 2(j) UCPD.

vulnerability exploitation *in their entirety*. Put differently: throughout the *entire stack* of service providers that deal in digital vulnerabilities and manipulation points of intervention can and should be found. With "stack" we refer to the fact that a digital choice environment is the result of combining (or stacking) a number of inter-related technical or organisational processes at different levels within and outside a company or organisation, involving a multitude of actors that decide about select parts of the service architecture. The user will often only see the public-facing user interface, but the fact that the user is presented with a particular service or user interface is the result of diverse design decisions at the operational, development, business or infrastructure level, or the result of decisions of external actors, such as standardisation bodies. For example, optimization logics permeate the entire stack. Such logics are decided on by management, but can inform every part of the organization down the line. It informs the KPIs that structure business decisions. It informs which user data are to be collected, how models are trained by those (and other) data to be rendered productive towards the optimization logics. It informs how the user interface – with all its features – are not only designed but constantly redesigned. This also implies that the optimization logics inform which (often iterative) software design philosophy is embraced. The optimization logics which inform all of the above will even determine which people end up being hired to work on, again, all of the above. Put simply: every layer of the stack that makes up a digital service should be a potential target of regulatory intervention if one takes digital vulnerability and manipulation seriously.

The inherent fluidity of digital vulnerability, as well as the need for a full stack approach to the manipulative exploitation of digital vulnerabilities, does raise the question how these concerns can be operationalised in static legal provisions. To be fair, the realization that addressing digital vulnerability and manipulation requires a wide, stack-like approach seems – albeit partly and indirectly – to be present in Article 25 of the DSA. The Article mentions that "Providers of online platforms shall not design, organise or operate their online interfaces in a way that deceives or manipulates". The "design, organise or operate" can indeed be read as an attempt to capture not just specific, isolated surface-level implementations, but also address the underlying organization. This broad reading is, however, directly limited again by only focusing on "online interface" which is just one part of the stack, and not typically the part in which decisive service design choices are made.

The phrasing of Article 25 DSA thus exemplifies the operationalisation challenge at hand. A vague and indirect gesture towards a stack-like approach is made, but the Commission ultimately fails to actually spell this approach out. In a more general sense, it would be beneficial if the legislator managed to tie piecemeal legislation that addresses different layers of the stack together in a wider, more coherent narrative. For example, data protection law addressing data practices and unfair commercial practices law addressing undue influences exerted on consumers both contain the legal resources to be part of a larger stack-like approach to manipulative digital consumer environments. Another example is the prohibition in Article 5 (2)(b) DMA to "combine personal data from the relevant core platform service with personal data from any further core platform services or from any other services provided by the gatekeeper". Limiting the ability to combine data from different services also limits the possibility of making inferences which again contributes to an asymmetrical relationship in which a company has much more knowledge about the consumer than the consumer about the company (and what it knows about him). Yet another example is Article 6 (3) DMA according to which designated gatekeepers must enable consumers to easily change default settings in a virtual assistant that directs or steers end users to produce and services of that gatekeeper. This is an example of a provision that reaches further down the stack to the operational and development

level. In other words, sprinkled across the emerging regulatory framework are different behavioural, structural and design requirements that address different aspects or layers of a digital choice environment and, doing so, tackle some of the underlying digital asymmetries that enable manipulation and exploitation of digital vulnerability in the first place. But when only dealt with separately, dealing with separate sub issues in isolation and without a more coherent approach, it will remain difficult to get into focus how the stack *as a whole* enables the production and exploitation of digital vulnerability and manipulative designs, or how the law can play a role in remedying the underlying structural asymmetries.

The stack approach is also useful in thinking about monitoring compliance and enforcement of the provisions of regulations like the DSA. For example, in order to monitor compliance with Article 25, competent authorities would need to look further than the design of the user interface, and also require information about the underlying algorithmic models, business decisions regarding optimisation decisions, user testing, input data, etc. In this context it is striking that Articles 67–69 of the DSA do already give the European Commission far-reaching means to collect the relevant information regarding the general organisation of a service, the algorithmic level but also business and data handling practices of Very Large Online Platforms and Very Large Online Search Engines, but that the Commission's investigative powers are not mirrored in the entitlements for national competent authorities and Digital Services Coordinates under the DSA.[19]

The Looming Privatization of Consumer Protection

The emerging regulatory framework must provide regulators and various societal actors with the means to address situations of digital asymmetry and more generally, scrutinize private control over the digital infrastructure of our algorithmic society. At the same time, it also lays the foundations for a growing privatisation of digital consumer protection, putting private companies increasingly in a position to (try to) *make or break* consumer protection. Private ordering through contracts but also technology design is not a new phenomenon in consumer law and protection.[20] In parts, regulations such as the DSA and the DMA can be seen as attempts to subject private ordering to new levels of regulatory scrutiny.[21] At the same time, the emerging digital regulatory framework, and here in particular the DSA and the AI Act, also embrace and institutionalise the conditions for private companies to define and operationalise consumer protection in digital choice environments. Consider the following examples of how they lay the foundations for new levels of privatisation of consumer protection:

---

19 Compare Art. 49 – 51 DSA.

20 Bakos, Y., Marotta-Wurgler, F., & Trossen, D.R. (2014). Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts. *The Journal of Legal Studies, 43*(1), 1–35; Belli, L., & Venturini, J. (2016). Private Ordering and the Rise of Terms as Service as Cyber-Regulation. *Internet Policy Review, 5*(4); Reidenberg, J. (1997). Lex Informatica: The Formulation of Information Policy Rules through Technology. *Texas Law Review, 76*(3), 553–594.

21 For example, as part of the systemic risk monitoring and mitigation obligations under Articles 34 and 35 specifically include an obligation to monitor terms and conditions upon their potential to create systemic risks, and where necessary, adjust those. Another example is Article 14(3) of the DSA, ordering providers to take into regard the rights and legitimate interests of users when enforcing their terms and conditions. Article 14 (3) DSA is, at the same time, also an example of how the DSA continues to leave acknowledge and even legitimate the use of contracts for private ordering by not submitting the terms and conditions themselves to regulatory scrutiny, see Quintais, J.P., Appelman, N., & Ó Fathaigh, R. (2023). Using Terms and Conditions to Apply Fundamental Rights to Content Moderation. *German Law Review, 24*, 881–911.

A first example are the systemic risk provisions in the DSA, and more generally the risked-based approach in the AI Act. Under Articles 34 and 35 in the DSA, it is an obligation of Very Large Online Platforms and Very Large Online Search Engines to 'diligently identify, analyse and assess' any systemic risks that are the result of the design or functioning of the digital choice environments they operate, including risks to a high level of consumer protection, and decide on the necessary mitigation measures. The DSA leaves it in the first place large discretion of VLOPS and VLOS to decide a) what risks to look into, b) interpret what a 'high level of consumer protection' entails, c) what metrics to use in testing their systems and d) what effective mitigation measures are.[22] Similarly, under the proposed AI Act it is the responsibility of the developers of high-risk AI systems to conduct a risk assessment. It is here that the problem of un(der) defined concepts such as vulnerability or manipulation become obvious: in the absence of a clear definition of what manipulation entails how will VLOPS or VLOS be able (or even attempt to) identify how their algorithmic systems engage in unethical or unlawful manipulation? Powerful commercial players in the digital economy will have to start engaging with increasingly important yet un(der)defined concepts such as (digital) vulnerability and manipulation.

The second example is the reliance on due diligence obligations and code of conducts (as the results of self- or co-regulation). One instrument in the DSA to "contribute to the proper application" of the regulation is voluntary codes of conduct, for example, in the area of systemic risks.[23] The Commission and the Board "shall aim to ensure that the codes of conduct clearly set out their specific objectives, contain key performance indicators to measure the achievement of those objects and take due account of the needs and interests of all interested parties, and in particular citizens, and in case of failure the Commission and board "may invite signatories to the codes of conduct to take the necessary action". In light of the interests and fundamental rights at stake, the phrasing of this paragraph signals polite resignation and reliance on the goodwill and expertise of technology providers. Even more striking is the approach under the AI Act where codes of conduct are the primary means of governance of all non-high risk AI systems (including most applications in the consumer sector).[24]

The third example is the prominent role of standardisation and adherence to (technical) standards as a form of demonstrating compliance with both, the DSA and the AI Act.[25] For example, according to the AI Act, high-risk AI systems that are in conformity with harmonised standards shall be presumed to be in conformity with the requirements of the AI Act.[26] The result is that (technical) standardisation bodies will play a critical role in interpreting and operationalising the regulatory framework, including the rules meant to protect vulnerable consumers against misleading or deceptive practices. The lack of necessary expertise in fundamental and

22 Leerssen, P.J. (2023). *Seeing What Others Are Seeing: Studies in the Regulation of Transparency for Social Media Recommender Systems*, PhD dissertation, University of Amsterdam; Leerssen, P.J. (2023). Counting the Days: What to Expect from Risk Assessments and Audits under the DSA – and When?. *DSA Observatory blog*, https://dsa-observatory.eu/2023/01/30/counting-the-days-what-to-expect-from-risk-assessments-and-audits-under-the-dsa-and-when/; Mantelero, A. (2022). Fundamental Rights Impact Assessments in the DSA: Human Rights and the Risk-Based Approach of the New EU Regulations on the Digital Society. *Verfassungsblog*, https://verfassungsblog.de/dsa-impact-assessment/.

23 Article 45 DSA.

24 Article 69 draft AI Act.

25 See for a comprehensive analyse the following excellent report: Micklitz, H.-W. (2023). *The Role of Standards in Future EU Digital Policy Legislation. A Consumer Perspective*. Report for BEUC (The European Consumer Organisation). https://www.beuc.eu/sites/default/files/publications/BEUC-X-2023-096_The_Role_of_Standards_in_Future_EU_Digital_Policy_Legislation.pdf

26 Article 40 draft AI Act (EC, EP and Council version).

human rights, and adequate representation of the interests of consumers has been flagged by academics and civil society as a serious concern.[27] Again, the failure to clearly define concepts such as vulnerability or manipulation can be instrumental in the failure of the regulatory framework to realise the values it seeks to protect.

The new playing field that emerges is one where we can expect the big commercial players that are the subject of regulation are also the ones that will proactively try to shape the interpretation and implementation of the un(der)defined concepts. The eventual definition and legal operationalisation of, e.g., 'manipulation' that will be settled on matters to the opportunities as well as constraints for online service providers. So absent guidance on how manipulation is supposed to be understood legally, one should expect the addressees of the new legislative agenda to volunteer interpretations that are especially business-friendly.

The privatization of consumer protection raises new challenges for consumer law and policy:

- new roles: with their rich expertise of consumer law and consumer concerns, consumer organisations will have an important role in issuing guidance for the concretisation of abstract terms such as vulnerability or manipulation by private organisations but it also can be necessary to consider new roles, for example representing the interests of consumers in standardisation efforts or the auditing of mandatory risk assessments from the perspective of consumer protection.
- new powers: One question around the privatisation of consumer protection is how far the authority and intervention rights of consumer organisations go, and if they are sufficient to monitor compliance of privatised acts of consumer protection. The question of powers and the reach of existing tools of consumer enforcement is particularly pertinent in situations in which no concrete consumer harm is materialised (yet) but the way private companies interpret and operationalise consumer protection does not take into account sufficiently the interests of consumers.
- new forms of cooperation: with the privatisation of consumer protection, new forms of alignment and cooperation between consumer authorities and private companies will emerge and be necessary. This creates new opportunities, for example for knowledge exchange and learning, but also new challenges for example of how to protect consumer authorities' independence.

The Consumer-Citizen: the Crumbling Distinction between the Consumer and the Citizen

The boundary between the consumer and the citizen is becoming increasingly porous. Today's digital marketplace is also the marketplace of ideas and podium of public discourse. The example of political microtargeting is useful to illustrate to what extent public and private functions, and consumers and citizens conflate. Commercial targeting practices are increasingly also

27 Micklitz, H.-W. (2023). *The Role of Standards in Future EU Digital Policy Legislation. A Consumer Perspective*. Report for BEUC (The European Consumer Organisation). https://www.beuc.eu/sites/default/files/publications/BEUC-X-2023-096_The_Role_of_Standards_in_Future_EU_Digital_Policy_Legislation.pdf; Ebers, M. (2022). Standardizing AI – The Case of the European Commission's Proposal for an Artificial Intelligence Act. In DiMatteo, L.A, Poncibò, C., & Cannarsa, M. (Eds.) (2022), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and* Ethics, (pp. 321–344) Cambridge University Press; Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach. *Computer Law Review International, 22*(4), 97–112.

used by political campaigns for political (micro)targeting,[28] relying on the same platforms (like Google and Facebook), even the same advertising auctioning system and the same data.[29] Political campaigns increasingly rely on the tools developed for commercial targeting practices and the same commercial parties (and here in particular the Google and Facebook duopoly) to spread their messages. The consequence is that political advertising is turning, at least from the perspective of platforms, into 'just another form of advertising', and it is becoming difficult to distinguish the citizen from the consumer. The blurring boundaries between the protection of consumers and citizens is also apparent in the way the AI Act is expanding the concept of vulnerability (see above) but also provisions like Article 25 DSA. According to Art. 25 (2) DSA, "[t]he prohibition in paragraph 1 shall not apply to practices covered by Directive 2005/29/EC or Regulation (EU) 2016/679". Effectively that leaves non-commercial forms of targeting as the main area of application of Art. 25 DSA. The crumbling distinction between citizen and consumer makes questions about the role and mission of consumer protection authorities and consumer law more pressing. Is it still realistic to distinguish between commercial and non-commercial communication? How far does the mandate of consumer protection authorities go? Do they increasingly also have a societal role, to not only consider the interests of consumers but also more collective and societal interests and fundamental rights? And what forms of cooperation are necessary between the different regulators?

Conclusion

The EU's new legislative agenda for the digital society is certainly ambitious. The package of the DSA, DMA, AI Act, and PAR is clearly aimed at addressing a wide range of issues and, moreover, tries to do so in a more structural manner. It is clear that the EU sees the exploitation of vulnerabilities and manipulative digital choice environments as serious systemic risks that warrant a systemic response. The laudable ambitions do, however, also result in a somewhat fragmented approach. In terms of vulnerability, the various legislative initiatives move between, on the one hand, a continuation of the 'average consumer versus the vulnerable consumer' approach, and, on the other hand, the first contours of a wider approach to *digital* vulnerability with more emphasize on the relational, architectural, and privacy-related nature of vulnerability. Manipulation, in turn, is making an appearance in the DSA, the AI Act, and the PAR. The appearance of manipulation of consumer and citizen behaviour as an explicit concern in legislation is unfortunately not yet accompanied by conceptual and/or definitional clarity. Nowhere is manipulation defined, and in the recitals and articles where it appears its relation to other problematic forms of influences (e.g., deception, dark patterns, nudging) remains unclear. The unclarities concerning the underlying legal theories and conceptualizations of (digital) vulnerability and manipulation are somewhat worrisome seen from the perspective of *private ordering*. The combination of 1) undertheorised and underdefined key concepts in combination with 2) an increased reliance on due diligence obligations and codes of conduct to give shape to these key concepts, potentially places quite some power in the hands of private companies to volunteer and propagate interpretations of (digital) vulnerability and manipulation that are above all commerce-friendly. These worries on private ordering in combination with vague concepts are, in turn, further exacerbated by the fact that the distinction

28 Dobber, T., Trilling, D., Helberger, N., & De Vreese, C.H. (2017). Two Crates of Beer and 40 Pizzas: The Adoption of Innovative Political Behavioural Targeting Techniques. *Internet Policy Review, 6*(4), 1–25.

29 Helberger, N., Dobber, T., & De Vreese, C.H. (2021). Towards Unfair Political Practices Law: Learning Lessons from the Regulation of Unfair Commercial Practices for Online Political Advertising. *JIPITEC, 12*, 273–296.

between the consumer under consumer law and the citizen as part of a democratic society is increasingly broken down. Take, for instance, the PAR. Political advertising on highly commercial platforms is both an issue of targeting consumers in commercial environments *and* an issue of targeting citizens involved in democratic processes. As digital choice environments increasingly break down contexts once conceived of as distinctive spheres, so do the recent legislative initiatives increasingly address people as citizen-consumers that play different roles in different context, all at the same time. The resulting conceptual messiness is understandable, but also leaves much room for improvement.