European Review of Digital Administration & Law - Erdal
2026, Volume 7, Issue 1, pp. 107-118
ISSN 2724-5969 – ISBN 979-12-218-2846-7 – DOI 10.53136/9791221828467 11

*The Rule of Law in Heavily Polluted Digital Information Ecosystems: Lessons from Industrial Pollution Analysis**

Juan Herrero
(Professor of Social Psychology, University of Oviedo)

Iván Solano
(PhD Candidate, University of Oviedo)

Alejandra Boto
(Professor of Administrative Law, University of Oviedo)

ABSTRACT The paper addresses the legal implications of the extensive implementation of the EEPR cycle (Engagement-Extraction-Processing-Revenue) by the technological industry and the constant dumping of contaminated information, capable of harming individuals and society. The data industry is contextualised within the framework of the polluting industry, thereby facilitating the use of the scientific and legal scholarship developed over the past decades.

KEYWORDS: Digital Ecosystems - EEPR Cycle - Digital Regulation - Human Digital Agency - Information pollution

TABLE OF CONTENTS: 1. Introduction. – 2. Target Scope and Regulation Challenges. The Battle Related to the Scientific Evidence. – 3. Regulation Approach of the EU and Digital Industry Responses. – 3.1. The Dissemination of Illegal and Harmful Content. – 3.2. The Commodification of Human Behavior and the Associated Threats to Users' Privacy and Agency. – 4. Self-Regulation and Regulation: An Ideological Battle with Effects on Legal Regulatory Activities. – 5. The Fight for Privacy. – 6. Conclusion.

1. *Introduction*

In the last decades, large technology corporations have grown to reach economic dimensions comparable to those of many industrialized economies. The market capitalization of the major technology companies (Meta, Alphabet, Apple, Microsoft, Amazon) in 2023 accounted nearly 55% of the GDP of the entire European Union. These are striking figures for a business where data plays a pivotal role (primarily for Alphabet and Meta, but also for others).

In 2022, Zuboff identified the EEPR (Engagement-Extraction-Processing-Revenue) cycle that typically underpins the business model of these companies.[1] In brief, the data business starts with a phase of engagement. In the realm of the internet and its applications and services, individuals must download apps and purchase wearables or home devices connected to the internet. To ensure a constant data extraction flow, individuals must actively use these services, devices, and applications. Additionally, individuals must consent to the use of their data for commercial purposes, explicitly allowing for data extraction. Third, the data flow requires processing and reprocessing to create tradable products and services. Fourth, revenue is generated by marketing the processed data.

This paper argues that industry's use of the EEPR cycle has led to the continuous discharge of pollutants into information ecosystems, which has resulted in significant harm due to negative externalities for individuals and society. These include the corruption of information ecosystems (epistemic chaos) and the lack of agency experienced by individuals. Public administrations cannot ignore these dynamics, yet designing a regulatory response is far from easy.

* Article submitted to double-blind peer review.
Alejandra Boto integrates in the research project PID2024-155142NB-I00; Juan Herrero is the principal investigator of research project PID2024-161573NB-I00 (both funded by Spanish State Research Agency, AEI).
[1] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization*, in *Organization Theory*, vol. 3, no 3, 2022.

2. *Target Scope and Regulation Challenges. The Battle Related to the Scientific Evidence*

The data industry, with its extractive and processing practices, generates a considerable amount of informational waste that is continuously discarded into information ecosystems in which it operates. Before legally regulating this activity, it is essential to conduct a comprehensive analysis based on scientific evidence of the harm caused by this industry to individuals and societies. This is the first obstacle encountered by legislators for several reasons. Firstly, the data industry is a recent phenomenon whose medium- and long-term effects can only be anticipated but not yet accurately measured. Secondly, scientific evidence is cumulative and subject to constant scientific review and debate, therefore policymakers must select conclusive evidence, which is difficult to achieve. Thirdly, the data industry attempts to influence regulatory activity in several ways. One method is by challenging the empirical evidence available and attempting to focus the scientific debate on the harms it generates. To inform their decision-making in this contentious environment, where considerable economic interests are at stake, policymakers require rigorous, replicable, generalizable, and unbiased empirical evidence. Fourthly, the efficacy of ongoing legislative initiatives must be subjected to rigorous evaluation to ascertain whether they are indeed achieving the intended objectives. In this regard, scientific evidence plays a pivotal role. Finally, data industry legislation currently seeks to achieve a delicate balance between protecting individuals and society on one hand, and fostering the innovation and economic growth that the industry generates, on the other. Excessively strict regulation can limit both economic growth and competitiveness vis-à-vis other economic regions where legislation is less stringent. However, excessively lax legislation, able to attract investments and economic activities from the data industry, could irreparably damage individual and social systems , as it will be argued in this paper.

The discussion of these aspects primarily focus on the European Union, which is currently leading the legal debate and is a pioneer among the most advanced economies in regulating the data industry. This is exemplified by the European Union's General Data Protection Regulation (GDPR), which has been considered a gold standard for data protection regulation.[2] This initiative has been reinforced with the publication in 2022 of the Digital Services Act (DSA) and the Digital Markets Act (DMA), along with other associated initiatives. This set of initiatives is informed by the European Declaration on Digital Rights and Principles. This declaration focuses mainly on six areas: 1) Putting people and their rights at the center of the digital transformation; 2) Supporting solidarity and inclusion; 3) Ensuring freedom of choice online; 4) Fostering participation in the digital public space; 5) Increasing safety, security, and empowerment of individuals; 6) Promoting the sustainability of the digital future.

These proposed areas of action appear reasonable when viewed in the context of the significant implementation of the EEPR cycle by the industry. The unregulated data industry could pose a threat to all six principles.

The European Union's legislative developments represent a significant shift in the landscape of digital governance.[3] In 2022, the European Parliament's passage of the Digital Services Act and the Digital Markets Act marked a pivotal moment in the fight against the data industry. These acts have ushered in a new era of democratic governance over the tech giants and their ecosystems. The Digital Services Act's is guided by the principle: 'What is illegal offline should be illegal online', which represents a significant step forward in addressing the challenges posed by the proliferation of illegal content, online disinformation, and other societal risks. It sets clear obligations for digital service providers, such as social media and marketplaces, to act in this regard. Furthermore, the document identifies large platforms (those with over 45 million monthly users in the EU) as a primary target for regulatory action. More stringent legislation will be applied to these and other similar platforms, which will include the prevention of systemic risks such as the dissemination of illegal or harmful content,

[2] J. Andrew and M. Baker, *The general data protection regulation in the age of surveillance capitalism*, in *Journal of Business Ethics*, vol. 168, no. 3, 2021, 565-578.

[3] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of kno.wledge in our information civilization*, cit.

adverse effects on fundamental rights, electoral processes, and gender-based violence or mental health. These platforms will also be subjected to independent audits and will be required to provide users with the option of not receiving recommendations based on profiling. Finally, they will be obliged to facilitate access to their data and algorithms to authorities and vetted researchers.

This regulatory strategy can be described as aggressive, driven by a deep-seated animosity towards large technology corporations. This viewpoint was expressed by Christen Schaldemose, a spokesperson for the Digital Services Act (DSA) committee.[4]

The primary tenets of the DSA may be summarized as follows: 1) New measures to counter illegal content online and obligations for platforms to react quickly, while respecting fundamental rights, including the freedom of expression and data protection; 2) Strengthened traceability and checks on traders in online marketplaces to ensure products and services are safe; including efforts to perform random checks on whether illegal content resurfaces; 3) Increased transparency and accountability of platforms, for example by providing clear information on content moderation or the use of algorithms for recommending content (so-called recommender systems); users will be able to challenge content moderation decisions; 4) Bans on misleading practices and certain types of targeted advertising, such as those targeting children and ads based on sensitive data. The so-called 'dark patterns' and misleading practices aimed at manipulating users' choices will also be prohibited.

These are undoubtedly measures that go to the heart of the EEPR cycle and that seek, in our view, to substantially limit the spillage of pollutants into information ecosystems. Countermeasures are in place to ensure that, when such spillages occur, large platforms will 'clean up' the information ecosystem. The primary concerns addressed by this Act are the dissemination of illegal or harmful content and the commodification of human behavior. While the dumping of pollutants into information ecosystems has been regarded as a source of epistemic chaos and systemic threats to modern democracies,[5] the commodification of human behavior has been linked to a loss of user agency, autonomy and a vulnerability of fundamental rights. Below, potential advantages and disadvantages of these legal initiatives will be examined critically, with particular attention to the strategies employed by corporations to avoid liability and continue their business operations.

Upon learning of the empirical evidence provided by scientists regarding the devastating effects of industrial pollution on people and land, lawmakers set out to regulate the industry by law. However, the industry became aware that certain legal measures regulating its activities could potentially threathen its survival by limiting its actions and reducing their profits. As a result, they shifted their focus to the source of these legal measures: the scientific evidence. During that time, an industry of merchants of doubt emerged, who were paid or sponsored by various industries.[6] Their goal was to systematically cast doubt on scientific findings that threathened industries' interests. Companies in the oil, tobacco, pesticide, paper, mining, and other industries began to counter scientific conclusions that could harm their interests with their scientific reports. These skeptics cast doubt on the scientific evidence, preventing lawmakers from justifying legal restrictions on their activities. This ongoing conflict between industry and regulators suggests that any attempt to regulate the data industry based on scientific evidence may face challenges similar to those faced by past regulatory initiatives. We propose a unified framework to leverage the accumulated expertise on the relationship between polluting industries and regulators and apply it to the new relationship between the data industry and necessary adaptations of

[4] Who explained that 'For too long tech giants have benefited from an absence of rules. The digital world has developed into a Wild West, with the biggest and strongest setting the rules. But there is a new sheriff in town - the DSA. No.w rules and rights will be strengthened. We are opening up the black box of algorithms so that we can have a proper look at the moneymaking machines behind these social platforms' (as quoted in the press release of the European Parliament, 5 July 2022, available at www.europarl.europa.eu/news/en/press-room/20220701IPR34364/digital-services-landmark-rules-adopted-for-a-safer-open-online-environment).

[5] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of kno.wledge in our information civilization,* cit.

[6] N. Oreskes and E.M. Conway, *Defeating the merchants of doubt*, in *Nature*, vol. 465, no. 7299, 2010, 686-687.

the regulatory framework.

Scholars have already described how the industry responds to evidence of the potential harm caused by its operations to the environment and its inhabitants and noted that this response can vary depending on the level of social concern at a given time.[7] Three phases or steps can be distinguished: 1) the industry's lack of engagement with externalities; 2) acknowledgment of the problem and symbolic compensation; and 3) engagement with critics and participation in legislation to shape regulatory policy. Corporations that harm and pollute territories and people exercise control over regulatory threats through a variety of mechanisms. These mechanisms include denying the problem, hesitantly acknowledging the problem, and proposing symbolic changes. When everything else fails and the regulatory threat appears existential, the industry takes the lead in acknowledging the urgency and contributing to regulatory changes. However, these changes do not prohibit its operations and allow it to survive and thrive.

The initial stage primarily aims to create skepticism and generate uncertainty in various forums by disputing empirical evidence. The digital industry relies on scientists whose work fosters uncertainty and is often used to argue that the harmful effects of their activities remain a matter of scientific debate. Legal initiatives to restrict or prohibit their activities can only be based on opinion, not scientific fact, as long as the scientific debate remains inconclusive.

In the second phase, as the industry becomes unable to counter the overwhelming scientific evidence against its polluting activities, some of these scientific conclusions are acknowledged superficially, and symbolic solutions are proposed to silence critical voices. During this phase, several actions can be taken to promote sustainability, such as introducing new, environmentally friendly products and services, committing to reorienting the company's activities, announcing expert programs and working groups to address the issue, and investing in research to resolve outstanding problems. The main objective of this phase is to save time.

Finally, the third phase is a critical moment that can threaten the survival of the industry. The harmful effects of their activities are now so obvious that they can no longer be denied. In addition, public opinion is overwhelmingly against them, and legislators are aggressively threatening ever more stringent regulations. At this point, the strategies employed in phases 1 and 2 are no longer useful. The industry needs to reinvent itself. The industry collaborates, negotiates, and pressures governments to reach legislative agreements that benefit them. They both ally with activists and take the lead in transforming their operations aimed at achieving a sustainable industry. When regulations are more hostile to their goals, they relocate their operations to take advantage of other geographic areas with more permissive legislation. The stages of response are processual and maintain a dynamic relationship. The digital industry may cycle through these phases multiple times.

3. *Regulation Approach of the EU and Digital Industry Responses*

The EU legislation attempts to cover the entire EEPR cycle. The objectives of this legislation are manifold. They include preventing the industry from employing engagement techniques that restricts individuals' ability to opt-out, guaranteeing fundamental rights such as privacy and dignity, ensuring that users regain agency and can act in a free and informed manner, regulating the flow of data and its processing and reprocessing, or setting limits on the use of artificial intelligence in vulnerable areas that may affect the security of citizens. These legislative initiatives aim to restrict the industry's activities and may affect its profitability.

The digital industry is developing various strategies to limit, mitigate, or prevent the legal environment from threatening its survival. The following sections illustrate how, in the EU legal context, the data industry seeks to protect its activities by addressing empirical evidence, acknowledging minor issues, proposing symbolic solutions, and taking the lead as a socially responsible and sustainable industry that ensures freedom in democratic societies. We will explicitly focus on two dimensions frequently emphasized by recent developments in the EU legislation: 1) the dissemination of illegal and harmful content, and 2) the commodification of human

[7] P. Benson and S. Kirsch, *Capitalism and the Politics of Resignation*, in *Current anthropology*, vol. 51, no. 4, 2010, 459-486.

behavior and the associated threats to user privacy and agency.

3.1. *The Dissemination of Illegal and Harmful Content*

In its current state of development, data transmission prioritizes what is known as blindness by design: where accuracy, volume, and speed of data transmission are prioritized,[8] and the semantics of that data (its content) are relegated to a secondary place, if not omitted altogether. Major platforms are constantly fine-tuning their data transmission systems with the primary goal of ensuring that the information sent by the sender is the same as the information received by the receiver. This 'exact' information, which navigates the immense communication flows, often originates from and is intended for humans and contains not only raw information but also 'meaning'. The sheer volume of communication users encounter daily would be unmanageable. Therefore, information must be filtered at the point of origin. These tasks are performed by the platforms' recommender systems, which, through increasingly complex algorithms, seek to encourage user engagement by presenting them with information that is of interest to them and that maximizes their activity on the platform. The success of these algorithms is measured by pragmatic criteria: the dissemination of information that increases engagement (number of likes, retweets, follows, etc.) is prioritized.

While EU laws seek to limit the transmission of certain content that is illegal or harmful to the user or the social system in general, the current systems cannot detect this type of information because content detection is not in their DNA. Data transmission systems are not designed a priori to detect-filter-delete unwanted information because data semantics is not a property of their open systems. As Facebook's engineers have recently recognized, 'Data Infra[structure] will need to have some semantic awareness of its data assets. (…) Building this will take multiple years and will require data infrastructure (DI) investment'.[9] This can be interpreted as a euphemism for the assertion that their business is data extraction, transmission, and commercialization, regardless of content. When Meta concedes that their open systems lack data semantics, they implicitly assume that the meaning of data is irrelevant, allowing any data to be exploited for profit. Following the recently enacted EU legislation, these systems are now required to be self-limiting, in that they must be designed in such a way that they are not completely open to all information, but rather, they must be able to sift out any information that is illegal or harmful. This stance represents a significant challenge to the industry as it would necessitate a fundamental shift in approach.

To identify something as illegal or harmful, its meaning must be assessed. Although many large platforms include automated detection systems that screen for prohibited content, these automated systems are decidedly blunt and imprecise and thus operate as tools for human moderators rather than stand-alone (or even primary) moderation devices.[10] These human moderators flag information as harmful, which can then be removed from the stream to mitigate its impact. This can take some time, and the malicious information may have spread to a significant number of users. The damage is already done because treating all posts equally, regardless of content, will inevitably amplify disinformation, sensationalism, hate, and other social harms. Content moderation thus works post-catastrophe.[11]

Considering the corporate strategy proposed, the industry has recently come to recognize a potential problem in the EEPR cycle, where illegal or harmful information finds ways to spread and go viral.[12] Facebook scientists identified the company's algorithms as a potential source of amplification of polarizing content, which could serve to

[8] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization,* cit.

[9] Facebook Ad and Business Product Team, *Facebook Data Lineage Internal Document: ABP Privacy Infra, Long Range Investments [A/C Priv]*, 2021. Available at www.documentcloud.org/documents/21716382-facebook-data-lineage-internal-document. In particular, see pages 7.8.

[10] J. Davis, *Theorizing curation*, in D. Rohlinger and S. Sobieraj (eds.), *The Oxford handbook of digital media sociology*, New York, NY, Oxford University Press, 2022.

[11] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization,* cit.

[12] P. Benson and S. Kirsch, *Capitalism and the Politics of Resignation,* cit.

perpetuate the EEPR cycle.[13] Despite this acknowledgment, the company disavows publicly the prioritization of such information by the EEPR cycle and continues to challenge scientific evidence to maintain an open scientific debate. In 2023, twenty-nine renowned scientists published an article in *Science* where they empirically demonstrated that current algorithms used by Facebook and Instagram do not foster affective polarization.[14] However, they found that alternative algorithms that might be used to prioritize the presentation of relevant information to users (specifically, a Chronological Feed in which the most recent content defined by the publication date and time of original posts appeared at the top of those feeds) might increase issue polarization as compared to the company's algorithm. This finding, published in one of the most important scientific outlets in the world, questions EU directives regarding recommending systems. 'Where several options are available (…), recommender systems that determine the relative order of information presented to recipients of the service, providers of online platforms shall also make available a functionality that allows the recipient of the service to select and to modify at any time their preferred option. That functionality shall be directly and easily accessible from the specific section of the online platform's online interface where the information is being prioritized' (Article 27(3) of the DSA). Several authors were or had been employed by Meta at the time of the research and others received investment funds from the company or were sponsored by the company with payments for conferences, lodging, travel, etc. This does not invalidate the conclusions of the article, which must be evaluated according to scientific criteria. However, it illustrates the company's interest in permanently leaving the scientific debate on its commercial practices. The message to EU legislators is clear: there is no conclusive empirical evidence to support EU legislative initiatives. Furthermore, the EU legislative proposal allowing users to choose the recommender system on the platform might lead to more polarization and information pollution in the ecosystem. This is what can be identified as phase 1:[15] contesting empirical evidence. In addition to this response strategy characteristic of phase 1, the industry is continuing to address the issue of illegal and harmful content that its operating practices neither prioritize nor generate. They are taking the lead in research on the dissemination of illegal and harmful content while proposing some symbolic solutions.[16] Human content moderation is a symbolic solution. Internal Facebook documents indicate that engineers had proposed effective methods to reduce polarization and the dissemination of illegal or hurtful content on the platform. These included modifying affordances, such as the reshare button.[17] However, the company rejected these proposals due to their negative impact on the EEPR cycle, particularly in the engagement phase.[18] This example illustrates how the dissemination of corrupted information can lead to economic gain, perpetuating the EEPR cycle. It is often more lucrative to invest a modest sum in the cleanup of spills than to attempt to prevent them entirely, as a portion of the profit is derived from the spill itself. The data industry processes data, regardless of contamination (e.g., illegal, illicit, dangerous, etc.).

The EU legislation aims to hold digital platforms responsible of pollutant spills into the information ecosystem, requiring the clean-up of the ecosystem whenever it happens. The legislative approach has been contested by the industry, which has argued that, given the limitations of current technology, this is not possible. This is part of the so-called phase 2 (at some point, the

[13] J. Horwitz and D. Seetharaman, *Facebook executives shut down efforts to make the site less divisive*, in *The Wall Street Journal*, 26 May 2020, available at: www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499.

[14] A.M. Guess, N. Malhotra, J. Pan, P. Barberá, H. Allcott, T. Brown, A. Crespo-Tenorio, D. Dimmery, D. Freelon, M. Gentzkow, S. González-Bailón, E. Kennedy, Y.M. Kim, D. Lazer, D. Moehler, B. Nyhan, C. Velasco Rivera, D.R. Thomas, E. Thorson, R. Tromble, A. Wilkins, M. Wojcieszak, B. Xiong, C.K. de Jonge, A. Franco, W. Mason, N.J. Stroud and J.A. Tucker, *How do social media feed algorithms affect attitudes and behavior in an election campaign?*, in *Science*, vol. 381, no. 6656, 2023, 398-404.

[15] Following P. Benson and S. Kirsch, *Capitalism and the Politics of Resignation,* cit.

[16] P. Benson and S. Kirsch's phase 2.

[17] K. Hagey and J. Horwitz, *Facebook tried to make its platform a healthier place. It got angrier instead*, in *The Wall Street Journal*, 15 September 2021. Available at www.wsj.com/tech/facebook-algorithm-change-zuckerberg-11631654215.

[18] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization,* cit.

problem is acknowledged) and phase 3 (which would have a severely detrimental impact on the industry). Meta engineers compared data flows to a bottle of ink poured into a lake, suggesting that it is not possible to 'put that ink back in the bottle' or to 'organize it again, such that it only flows to the allowed places in the lake'.[19] Once the industry has polluted the information ecosystem, as they have acknowledged, it is not possible to reverse the situation. The informational ecosystem will almost certainly evolve to become more polluted.

Digital platforms have also argued that information contamination has been a feature of human communication systems for as long as humans have existed. As such, social platforms, which are designed to emulate these communication systems, will inevitably contain some degree of contaminated information. This is simply how humans behave. Once more, this can be regarded as a corporate strategy to obfuscate the facts.[20] In face-to-face social environments, information with hurtful or illegal meanings is also transmitted through communication flows. Still, in many cases, the agency of the individual can have an important inhibiting effect on its subsequent transmission. Forethought, self-regulation, and self-reflection constitute the core components of human agency.[21] This agency is significantly reduced in digital information flows governed by algorithms designed to enhance engagement.[22] Scholars of digital communication have focused on the ability of users to decide when to engage or disengage from digital communication environments,[23] but this rational decision is strongly influenced by corporate goals to increase engagement to ensure the collection of personal data.[24] This corporate strategy leads to the resignation and loss of human agency in users' interactions with the digital industry. Unless users are empowered to regain agency in their interactions with the digital industry, the impact of harmful information on individuals and society at large will continue and even increase over time.

3.2. *The Commodification of Human Behavior and the Associated Threats to Users' Privacy and Agency*

The initial phase of the EEPR cycle, engagement, is of particular importance to the interests of the data industry. Greater engagement has been shown to increase data extraction, which in turn allows for the inclusion of more data points in processing algorithms. These algorithms generate semantics for the data and build tradable products. This favors the marketing of data-based products and services (profiling of users, for example) and increases revenue. This phenomenon can be illustrated by the aggregation of data from 'likes' on various themes and issues, the subsequent inference of psychological traits (data semantics), and the sale of this profiling to political actors, such as in the Cambridge Analytica scandal.[25]

Through the process of data extraction, the commodification of human behavior is enabled. This transformation results in the creation of a commodity with an associated economic value. Although the data industry initially extracted personal data with minimal users' consent, these practices have been gradually prohibited by different legal systems. In these legal systems, and the EU legal system is no exception to this rule, the practice of obtaining informed consent from users before the extraction and commercialization of their data has become widespread. The conditions under which data is extracted and commercialized are a matter of concern for the EU (Article 7 of the GDPR; Article 14 of the DSA; etc.). The GDPR places a significant emphasis on the protection of citizens in their interactions with the digital

[19] L. Franceschi-Bicchierai, *Facebook Doesn't Know What It Does With Your Data, Or Where It Goes: Leaked Document*, in www.vice.com/en/article/facebook-doesnt-know-what-it-does-with-your-data-or-where-it-goes.
[20] P. Benson and S. Kirsch's phase 1.
[21] A. Bandura, *Toward a Psychology of Human Agency: Pathways and Reflections*, in *Perspectives on Psychological Science,* vol. 13, no. 2, 2018, 130-136.
[22] G. Ceylan, I.A. Anderson and W. Wood, *Sharing of misinformation is habitual, not just lazy or biased*, in *Proceedings of the National Academy of Sciences*, vol. 120, no. 4, 2023, e2216614120.
[23] M.M. Vanden Abeele, *Digital wellbeing as a dynamic construct*, in *Communication Theory*, vol. 31, no. 4, 2021 932-955.
[24] S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization,* cit.
[25] J. Hinds, E.J. Williams, and A.N. Joinson, *"It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal*, in *International Journal of Human-Computer Studies*, vol. 143, 2020, 102498.

industry. The EU claims that it has established clear legislation regarding the consent to the processing of users' data. The user's consent must be specific for any specific matter, intelligible, accessible, and written in plain and clear language. Moreover, the consent must be freely given and the user must be well informed; however, companies offering products and services may not require consent for the processing of personal data that are not necessary for the performance of the contract. The objective of this initiative is to safeguard the user's online experience.

When users provide access to their data, they usually get free access to the apps and services, and as soon as they deny their data extraction, processing, and commercialization, the enjoyment of the service or app ends (such as in the pay-or-consent approach[26]). The industry trades services and products for privacy. For the industry, whether they want to or not, users must allow access to the data they generate to feed the EEPR cycle. The EU specifically targets limiting the industry tactics regularly used to bend the users' will. The DSA is seeking to prohibit the use of so-called 'dark patterns' and deceptive practices that are designed to influence users' choices.[27] The EU defines dark patterns as 'online interfaces of online platforms that materially distort or impair, either on purpose or in effect, the ability of recipients of the service to make autonomous and informed choices or decisions' (DSA, recital 67).[28] Insofar as they may impair users' agency, dark patterns are prohibited. The issue of clearly defining these choice architectures and dark patterns and listing them precisely to indicate which practices are specifically prohibited is raised. The EU directives do indeed provide a potential list of these practices.[29] The prohibition of these practices by law is intended to achieve two objectives. The first is to protect users from engaging in unwanted behaviors or making decisions that negatively affect them. An example of this would be making it extremely difficult to opt out of a service. The second objective is to protect users from losing agency and the ability to make autonomous and informed choices or decisions. This objective is more general and seeks to restore users' agency in their relations with the digital industry. The list of deceptive design patterns, best practices, and use cases is not exhaustive, and a specific technique might not be present. In this case, social media providers remain responsible and accountable for ensuring the GDPR compliance of their platforms (DSA, recital 67). A more detailed examination of this topic will be presented in the context of self-regulatory and regulatory legal initiatives.

4. *Self-Regulation and Regulation: An Ideological Battle with Effects on Legal Regulatory Activities*

The spectrum of regulatory and self-regulatory approaches[30] encompasses all potential avenues for orienting the industry's activities toward sustainable objectives. At the most radical end of the argument in favor of self-regulation, the existence of polluted information ecosystems is recognized, but the industry's role is usually contested. While information ecosystems that are contaminated may pose a risk to certain individuals and perhaps some vulnerable groups of society, one way to limit these risks is by fostering the resilience of users, empowering their autonomy, and ensuring their agency. This is

[26] S. Nam and Y. Kwon, *Is pay-or-consent for privacy justifiable? Evidence from different users' privacy attitudes toward behavioral data collection in South Korea*, in *Telecommunications Policy*, vol. 48, no. 7, 2024, 102794.

[27] J. Ponce Solé, *Law, Digital Nudging and Manipulation: Dark Patterns, Artificial Intelligence and the Right to Good Administration*, in *European Review of Digital Administration & Law*, vol. 3, no. 1, 2022, 31-44.

[28] European Economic and Social Committee, *Digital Services Act and Digital Markets Act – Stepping stones to a level playing field in Europe*, https://data.europa.eu/doi/10.2864/28842.

[29] European Data Protection Board, *Guidelines 03/2022 on Deceptive design patterns in social media platform interfaces: how to recognise and avoid them*, www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-032022-deceptive-design-patterns-social-media_en.

[30] Please refer to N. Richards and W. Hartzog, *The Pathologies of Digital Consent*, in *Washington University Law Review*, vol. 96, no. 6, 2019, 1461-1504; L. Floridi, *The end of an era: from self-regulation to hard law for the digital industry*, in *Philosophy & Technology*, vol. 34, 2021, 619-622; M. Fasoli, *The Overuse of Digital Technologies: Human Weaknesses, Design Strategies and Ethical Concerns*, in *Philosophy & Technology*, vol. 34, no. 4, 2021, 1409-1427; R. Medzini, *Enhanced self-regulation: The case of Facebook's content governance*, in *New Media & Society*, vol. 24, no. 10, 2022, 2227-2251 and U. Smartt, *Upholding Digital Rights and Media Plurality: Does Self-regulation by Social Media Platforms Contravene Freedom of Expression?,* in K. Fowler-Watt and J. McDougall (eds.), *The Palgrave Handbook of Media Misinformation*, Cham, Springer International Publishing, 2022, 31-44.

the position of the major platforms that dominate the data industry. The industry argues that the extensive implementation of information and communication technologies is a significant driving force for the improvement of living conditions in highly digitized societies, a powerful driver of economic growth, a positive vector for the transformation of societies, and the vehicle for a promising future. Defenders of self-regulation reluctantly acknowledge that during its successful development, the digital industry may be accompanied by some dysfunctions, in any case, unwanted and alien to the industry's interests and commitments. The industry stresses that its fundamental motivation is the improvement of human living conditions and that regulatory agencies should let them pursue that goal without too much interference. Their primary interest is in people and societies, and they are particularly vigilant of potential undesirable dysfunctions that may arise. They use all available resources to limit the impact of such dysfunctions.

From a regulatory standpoint, the glass can be seen as half-empty. While acknowledging the social and economic progress that these technologies promote, there is a tendency to distrust the industry's commitments and focus primarily on the damage its activity causes, including threats to individuals and societies. The increasing contamination of information ecosystems and the systematic erosion of users' autonomy and agency in their interactions with the industry represent a matter of concern that must be addressed. Allowing the industry to grow without external control has caused a multitude of negative externalities, justifying the implementation of strict regulation. In the context of the EU, legislators have declared a point of no return.

The overwhelming empirical evidence on the negative effects of industry operations on individuals' psychological well-being, the dissemination of disinformation, the polarization of political attitudes, or other political interferences has led the EU, for instance, to pursue stricter regulations and to leave to self-regulation only those aspects in which regulation is not entirely possible. This is the case for those dark patterns that have not yet been identified by the regulator and that are not included in the list. In such instances, it is left to industry self-regulation through the implementation of best practices. Given the rapid pace of technological advancement in the industry, it would be prudent for legislators to consider implementing periods during which practices not yet prohibited are subject to self-regulation.

Self-regulation represents the preferred path for the digital industry. This approach rests on an industry-imposed vision that there are no harmful products but that individuals use them inappropriately. According to this perspective, the most effective approach to curtailing the dissemination of polluted information may be the implementation of legislative measures designed to enhance digital media literacy.[31] Furthermore, the development of a media-critical citizenry has the potential to encourage data industry actors to prioritize the exchange of high-quality and non-polluted information.

This strategy bears a resemblance to the above identified as second phase in the corporate response to liability in the industry's operations. This point has been as well summarized as the 'McDonald's model': faced with imminent regulation, fast-food companies embarked on a mass marketing campaign to argue that their products were safe if they were consumed as part of a healthy lifestyle. Since consumption (or engagement) was strongly rooted in a personal decision, it meant that providers were not liable for any detrimental health effects caused by the food they sold (or the digital products and services that they provide). This has the effect of blaming users for problems in their emotional well-being while absolving the providers of these apps of any responsibility.[32] The amount of research aimed at identifying the individual mechanisms that explain behavioral addiction to mobile devices and their communication applications is vast.[33] Aspects such as personality, self-control, or the psychosocial needs of users have proven to be interesting markers of addiction. This line of empirical evidence could serve to

[31] F. Nuñez, *Disinformation legislation and freedom of expression*, in *University of Carolina Irvine Law Review*, vol. 10, 2019, 783.

[32] M.J. Dennis, *Towards a theory of digital well-being: Reimagining online life after lockdown*, in *Science and Engineering Ethics*, vol. 27, no. 3, 2021, 1-19.

[33] For reviews and meta-analyses, see Busch and McCarthy, 2021; Casale *et al.*, 2022; Elhai, *et al.,* 2017; Horwood and Anglim, 2021; Sullivan and George, 2023.

justify self-regulation.

The primary legislative strategy of the EU is based on regulation, although it also incorporates self-regulation. Contemporary technologies underpinned by supportive data protection laws, such as the GDPR, seek to increase the agency of individual users to manage their data.[34] The EU establishes a general regulatory framework, which becomes increasingly strict as it incorporates additional laws. This framework aims to guarantee that the user can give informed consent. In this context, the final decision is left to the users' self-regulation. The key element in this process is free and voluntary consent. If users are properly informed and freely decide to consent, the process will be considered lawful. Their rights will be protected. This opens a promising avenue for further inquiry, whereby an indirect strengthening of industry regulation can occur because of increased user resilience.

The notion of a 'marketplace of ideas' is a compelling example of self-regulation as an optimal strategy to counteract the dissemination of polluted information within the informational ecosystem. In this theoretical framework, polluted information competes with factual information until the latter prevails. The viability of this process depends on the availability of reliable information, on users' ability and willingness to distinguish it from polluted information, and on the assumption that information pollution occurs inadvertently.[35] The probability of this situation occurring is low, as the three preconditions are not likely to apply to contemporary digital ecosystems. Firstly, users are not exposed to factual information that contradicts their views. Secondly, under the condition of low digital agency, distinguishing between polluted and non-polluted information becomes more difficult. Thirdly, polluted information increases revenue as compared to non-polluted information, so it is not an accidental spill. Because the marketplace of ideas- i.e., complete self-regulation- has been ineffective in dealing with the spread of polluted information, at some point, regulatory interventions may be necessary. These interventions, however, should be kept to a minimum from the self-regulatory perspective, which is predicated on the notion that, under optimal conditions, users will engage in judicious decision-making processes, thereby leading, for instance, to the complete self-regulation of detrimental or unlawful content. Critics of this perspective might counterargue that in highly polluted information ecosystems, where epistemic chaos prevails, engagement techniques employed by the industry would likely diminish users' agency before these informed decisions transpire.

[34] See L. Floridi, *The end of an era: from self-regulation to hard law for the digital industry,* cit.

[35] F. Nuñez, *Disinformation legislation and freedom of expression*, cit.

5. *The Fight for Privacy*

As the EEPR cycle is contingent upon the user's consent, the primary and most contentious issue is privacy. Regulators are deeply concerned about potential privacy vulnerabilities that might result in an infringement of fundamental rights. Although not publicly disclosed, the industry sees privacy as an obstacle to its operations. From the EEPR cycle point of view, this is completely understandable: the user's privacy donation heavily conditions the available data to be extracted. In the absence of effective measures to ensure the acceptance of user privacy donations, the viability of industry operations may be significantly compromised.

The GDPR defines personal data as any information relating to an identified or identifiable natural person (see Article 4). The problem with this definition is that there is a significant difference between what the law considers to be data –or specifically personal data – which can identify the user, and what the industry can do with any data that can identify the user, even without the user's conscious intervention. Most data generated during an online interaction can be used to identify the user. When EU law seeks to assure the user that the industry will not traffic in his or her personal information unless he or she clearly expresses informed consent to do so, it threatens the industry with limiting access to some of the data most relevant to its business. The industry's response includes Benson and Hirsch's stages 1 and 2: denying that personal information is of special interest (stage 1) and creating new products that are more privacy-friendly (stage 2).

Over time, the sophistication of predictive models and the increase in computing power have allowed metadata to become an increasingly reliable source of information.

Today, the industry can obtain accurate personal data without the voluntary disclosure of the user: the correct triangulation of different metadata sources makes it possible to de-anonymize the massive, aggregated data stored on servers. Although users are constantly told that it is not possible to identify individuals from large, anonymized datasets,[36] three spatial-temporal points are enough to allow a re-identification in 90% of the cases.[37] According to a study conducted in 2013, Facebook 'likes' predict gender, political orientation, ethnicity, and sexual orientation with an 85% success rate. Behavioral data can be used to identify the core psychological traits of a user.[38] A more salient issue is the possibility of obtaining data from one user through the data of another user. This results in the user's data being transmitted through the information ecosystem without the user's knowledge or consent.

The industry has developed new products as part of Benson and Hirsch's Phase 2 characteristic responses that are more related to user privacy. Although the industry continues to deny that its products invade users' privacy, it is sensitive to the potential threats of privacy invasion and is leading the public discourse on the need to protect this asset for users. In 2019, for instance, Meta's CEO, Zuckerberg, unveiled the company's new strategic direction at his annual developer conference: 'The future is private,' he proclaimed.[39] This has been accompanied by the development of products that better protect privacy, such as the secure browsing implemented in virtually all major browsers. The tobacco industry, which was singled out by health authorities and public opinion for deliberately adding addictive substances (engagement) such as nicotine to its products, decided – without publicly admitting that these were addictive substances and that they were added to its products for that reason – to invest in products that were in principle less addictive and less harmful – for example, light cigarettes. But these products are just as harmful, as empirical evidence has shown. Something similar is happening with secure browsing, which uses this euphemism to increase engagement and ensure data mining in areas that users consider more private.

In EU legislation, the ownership of personal data belongs to the subjects, and it can only be traded with their consent. We use the term 'trade' to clarify that the industry's preferred term, 'privacy policies', primarily refers to the 'trading' of personal data. This is the first instance of an industry employing a 'dark pattern' that misleads users into believing that the clauses are about safeguarding their privacy rights when, in fact, they are about the trading of their data. Instead of asking, 'How would you prefer to safeguard your privacy?' they should ask, 'How would you prefer that we do business with your data?' This dark pattern is not on the list provided under the DSA regulation. This is probably one of the best wins for the industry. The Privacy Policy section should be renamed to align with the new data protection regulations, for instance, Your data trading rights.

We have drafted a section on Your personal data trading rights that complies with GDPR legislation: 'the request for consent shall be presented in (…) in an intelligible and easily accessible form, using clear and plain language' (GDPR, Article 7(2)). It runs as follows:

'We buy and sell personal data. To be profitable to us, you must agree to let us use your personal information. We sell your data in commercial transactions. Our policy is to pay for your data through products and services designed for your entertainment, social connectivity, information search, etc. These products and services are designed to maximize your time using them. This way, we can learn more about you and improve our predictive models about your tastes, interests, political ideology, sexual identity, close and intimate confidants, and other aspects that you may consider intimate but that we can do business with. Our company may use your data, but it may also be sold to third parties whose policies sometimes do not exactly match ours. Your data can be used for other purposes. It is sold and resold through continuous processing with different purposes to different companies that also do business

[36] P. Benson and S. Kirsch's phase 1.
[37] Y.A. de Montjoye, L. Radaelli, V.K. Singh and A.S. Pentland, *Unique in the shopping mall: On the reidentifiability of credit card metadata*, in *Science*, vol. 347, no. 6221, 2015, 536-539.
[38] M. Kosinski, D. Stillwell and T. Graepel, *Private traits and attributes are predictable from digital records of human behavior*, in *Proceedings of the national academy of sciences*, vol. 110, no. 15, 2015, 5802-5805.
[39] In S. Zuboff, *Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization,* cit.

with your data. Please indicate below if you accept our terms and conditions. Otherwise, you will not be able to access our service.'

The discrepancy between the actions of the industry, which are legally protected, and the actions that should be taken, from a perspective that is more favorable to the interests of users, invites a critical analysis of the current methods used for data mining.

6. *Conclusion*

We have critically examined the consequences of some industry practices for the user, with a particular focus on the negative externalities that result. These externalities can be divided into two categories: increased contamination of the information ecosystem and erosion of users' agency and autonomy. The paper has shown how these two negative externalities are a direct consequence of the current implementation of the EEPR cycle within the industry. On the one hand, the dissemination of contaminated information, such as disinformation, generates significantly more engagement than truthful information. This phenomenon results in the dumping of polluted information in information ecosystems, which generates greater profit for the industry than truthful information. Conversely, the loss of user agency and autonomy results in increased engagement, which in turn boosts extraction and the entire EEPR cycle. The depletion of individual and social resources carried out through this EEPR cycle and its capacity to contaminate information poses a significant risk to social systems, potentially leading to a state of epistemic chaos. In such a scenario, the quality of information is severely compromised, with truthful and corrupted information reaching an equivalent epistemic status. This, in turn, has the potential to undermine individual and social functioning.