ARTICLES

FUTURE-PROOF REGULATION AND ENFORCEMENT FOR THE DIGITALISED AGE

Edited by Gavin Robinson, Sybe de Vries *and* Bram Duivenvoorde

DIGITAL DETECTIVES: A RESEARCH AGENDA FOR CONSUMER FORENSICS

CATALINA GOANTA*

TABLE OF CONTENTS: I. Introduction. – II. Consumer law enforcement through computational investigations. – II.1. The CPC Regulation. – II.2. DSA. – III. Computational Measurements of Influencer Activity: a case study for digital enforcement. – IV. Consumer forensics: a new field for digital detectives. – IV.1. Consolidating data collection and analysis methods for consumer law. – IV.2. Classifying new forms of consumer harms. – IV.3. Developing new approaches to consumer protection. – IV.4. Exploring the legal and ethical implications of consumer forensics. – V. Conclusion.

ABSTRACT: Effective enforcement on digital markets is one of the most essential considerations for contemporary consumer law and policy. On digital markets, traders engage in very sophisticated commercial practices that are opaque to the average consumer. Online information is necessary for the monitoring of how companies engage in legal compliance, and where public authorities should intervene. This puts a lot of pressure on public administration to develop investigation and enforcement approaches that match the different consumer harms and needs arising out of digital markets. As it is virtually impossible to police technology practices without understanding the technologies and business models behind them, public authorities need to arm themselves with the means necessary to detect digital violations. This *Article* focuses on the digital enforcement of consumer protection law in the European Union and proposes a new field of research focused on the investigation and enforcement of consumer violations on digital markets in the form of "consumer forensics". In the author's opinion, consumer forensics is the answer to the question of how consumer enforcement regulation can become future-proof. As digitalization is rapidly affecting the way in which consumers are protected on the Internet, both the substantive and procedural dimensions of regulatory effectiveness will be impacted by evidence gathering to understand and further prove the existence of new online harms. To show the potential of this topic, the *Article* will offer some in-depth insights from a very specific topic of administrative scrutiny, namely measuring influencer marketing activities that are relevant for consumer protection.

KEYWORDS: consumer protection – consumer forensics – influencer marketing – digital enforcement – European private law – computer science.

* Associate professor, Utrecht University, e.c.goanta@uu.nl.

www.europeanpapers.eu
ISSN 2499-8249
doi: 10.15166/2499-8249/680

I. Introduction

Effective enforcement on digital markets is one of the most essential considerations for contemporary consumer law and policy. In the offline world, authorities do on-site inspections, track shipments and test substances to make sure that consumer products are safe and practices fair.[1] Yet the current state of digital markets challenges existing paradigms of law enforcement. On digital markets, traders engage in very sophisticated commercial practices that are opaque to the average consumer. For instance, consumer interfaces are now masterfully designed in ways that are said to nudge users to engage in transactions;[2] websites comparing product prices give consumers more comparative information about existing offers but also raise product prices altogether;[3] and reviews can give more insights into the qualities of products or services, but they can also be gamed by malicious actors.[4] These are only a handful of examples of how in addition to their role of facilitating transactions in goods and services, consumer digital markets are increasingly designed to convey information, in ways and to ends that we do not yet fully grasp, and which may not always favour consumers.

The online information given to consumers is necessary for the monitoring of how companies engage in legal compliance, and where public authorities should intervene. This puts a lot of pressure on public administration to develop investigation and enforcement approaches that match the different consumer harms and needs arising out of digital markets. On 17 February 2023, the United States Federal Trade Commission (FTC) launched a new Office of Technology, aiming to "support FTC investigations into business practices and the technologies underlying them".[5] The FTC's approach to safeguarding consumer interests on digital markets is an example of a worldwide trend, which emphasizes one very clear direction: developing technology to investigate and assess other

[1] U Wollein and others, 'Potential Metal Impurities in Active Pharmaceutical Substances and Finished Medicinal Products: A Market Surveillance Study' (2015) European Journal of Pharmaceutical Sciences 100.

[2] CM Gray and others, 'The Dark (Patterns) Side of UX Design' *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (ACM 2018) dl.acm.org; CM Gray and others, 'Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective' *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (ACM 2021) dl.acm.org; A Mathur and others, 'Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites' (2019) Proceedings of the ACM on Human-Computer Interaction 1.

[3] D Ronayne, 'Price Comparison Websites' (2021) International Economic Review 1081.

[4] EF Cardoso, RM Silva and TA Almeida, 'Towards Automatic Filtering of Fake Reviews' (2018) Neurocomputing 106; M Juuti and others, 'Stay On-Topic: Generating Context-Specific Fake Restaurant Reviews' in J Lopez, J Zhou and M Soriano (eds), *Computer Security* (Springer International Publishing 2018); J Malbon, 'Taking Fake Online Consumer Reviews Seriously' (2013) Journal of Consumer Policy 139; J M Martínez Otero, 'Fake Reviews on Online Platforms: Perspectives from the US, UK and EU Legislations' (2021) SN Social Sciences 181; R Mohawesh and others, 'Fake Reviews Detection: A Survey' (2021) IEEE Access 65771.

[5] Federal Trade Commission, 'FTC Launches New Office of Technology to Bolster Agency's Work' (16 February 2023) www.ftc.gov.

technologies. As it is virtually impossible to police technology practices without understanding the technologies and business models behind them, public authorities need to arm themselves with the means necessary to detect digital violations. In addition, the sheer scale at which harmful practices (*e.g.* dark patterns)[6] may be deployed requires an overhaul of how investigations are done.

This *Article* focuses on the digital enforcement of consumer protection law and proposes a new field of research focused on the investigation and enforcement of consumer violations on digital markets in the form of "consumer forensics". While forensic science has traditionally been affiliated with the criminal field,[7] the proposed concept of consumer forensics deals with uncovering consumer violations though various methods and procedures. In the author's opinion, consumer forensics is the answer to the question of how consumer enforcement regulation can become future-proof. As digitalization is rapidly affecting the way in which consumers are protected on the Internet, both the substantive and procedural dimensions of regulatory effectiveness will be impacted by evidence gathering to understand and further prove the existence of new online harms. To show the potential of this topic, the *Article* will offer some in-depth insights from a very specific topic of administrative scrutiny, namely measuring influencer marketing activities that are relevant for consumer protection.

The *Article* is structured as follows. Section II makes an overview of the characteristics of EU consumer law enforcement on contemporary digital markets, and briefly discusses some administrative powers enabled through EU sectoral regulation such as the Consumer Protection Cooperation Regulation (CPC Regulation)[8] and the Digital Services Act (DSA).[9] Section III discusses influencer marketing and the need to monitor influencer activities, as an example of how digital enforcement tasks can be dealt with by existing technologies. In doing so, the section also offers insights into the various categories of computational approaches which may be used to monitor the activities of social media influencers. Section IV defines and elaborates upon the concept of consumer forensics as a multidisciplinary field of research, and proposes a research agenda for the future. Section V concludes.

[6] A Mathur and others, 'Dark Patterns at Scale' cit.

[7] A Årnes (ed.), *Digital Forensics: An Academic Introduction* (John Wiley & Sons Inc 2018); John Sammons, *The Basics of Digital Forensics: The Primer for Getting Started in Digital Forensics* (Elsevier/Syngress 2012); K Nance, B Hay and M Bishop, 'Digital Forensics: Defining a Research Agenda' (2009) 42nd Hawaii International Conference on System Sciences ieeexplore.ieee.org.

[8] Commission Regulation 2017/2394 of the European Parliament and of the Council of 12 December 2017 on Cooperation Between National Authorities Responsible for the Enforcement of Consumer Protection Laws and Repealing Regulation (EC) No 2006/2004.

[9] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC.

II. Consumer law enforcement through computational investigations

Digital enforcement is a development rooted in the increased digitalization of our societies. In the past years, due to the rising interest of monitoring digital markets, Open Source Intelligence (OSINT) and academic initiatives have proposed new approaches for Internet investigations that are relevant for digital enforcement.[10] The FTC has been a trail-blazer in this respect. In the past years, the FTC has supported academic events where scientific labs working on privacy and security, network science, natural language processing, and other relevant computer science fields, present their cutting-edge computational research relevant for consumer protection. For instance, in the 2022 edition of one such events, the 6th Workshop on Technology and Consumer Protection,[11] one of the papers presented (and co-authored by staff of the FTC's Office of Technology Research and Investigation) featured the largest dataset of Yelp reviews currently available in academic research. The dataset consists of a total of two million unique reviews of more than 10,000 businesses monitored over periods from four months to eight years, resulting in 12.5 million data points.[12] This dataset was used to study an upcoming problem with fake review monitoring, namely reclassification – the dynamic algorithmic filtering and reallocation of quality review labels by platforms. The dataset is publicly available, and so is the code used in the project, including the crawler used to gather the data.[13] Another study that showed the potential of computational approaches for digital monitoring was the dark patterns study conducted at the Center for Information Technology Policy at Princeton University.[14] The study analysed 53,000 product pages from 11,000 shopping websites, and uncovered 1,818 dark pattern instances. 183 websites were identified to engage in deceptive practices.[15]

These are examples of how even with their existing limitations, computational approaches can contribute to the development of digital monitoring methodologies. Data collection, facilitated by automated crawlers that visit pre-determined Internet paths and scrape html code, photos, url links, etc. from webpages, or collect meta data, can be used to scale Internet investigations. The dark patterns study by Mathur et al. mimicked the steps a consumer would take when purchasing goods or products on the Internet. In the study, no transactions were completed, but data relating to consumer options available up to the very moment of clicking on the purchase button were registered through the crawl.[16]

[10] See for instance the evidence review undertaken by the UK Consumer and Markets Authority. CMA, 'Evidence Review of Online Choice Architecture and Consumer and Competition Harm' (5 April 2022) www.gov.uk.

[11] Workshop on Technology and Consumer Protection www.ieee-security.org.

[12] R Amos and others, 'Reviews in Motion: A Large Scale, Longitudinal Study of Review Recommendations on Yelp' (2022) 6th Workshop on Technology and Consumer Protection (ConPro).

[13] Princeton Longitudinal Reviews Dataset sites.google.com.

[14] A Mathur and others, 'Dark Patterns at Scale' cit.

[15] *Ibid.*

[16] *Ibid.*

Such research projects are not necessarily cost-prohibitive (*e.g*. they often do not entail consumer experiments, which are generally expensive), but the type of computer science expertise needed for them can sometimes be a limitation. This is why public authorities still have a long way to go to be able to conduct such wide-scale investigations, and why the academic world has been so far the cradle of development for state-of-the-art methodologies and results. Most importantly, public authorities cannot undertake investigations at scale without considering their legitimacy to collect data or to have broader, explicit powers for digital monitoring. Against this background, a lot of recently adopted regulatory instruments introduce new procedural frameworks for digital enforcement. Two such instruments are the CPC Regulation and the DSA. The first regulatory instrument included in this analysis has been a much-needed upgrade to consumer law enforcement, and the DSA, while not being a part of the consumer *acquis* as such, is one of the most modern frameworks in EU platform liability, particularly due to its enforcement mechanisms.

II.1. The CPC regulation

The CPC Regulation is a European Union regulation that aims to improve consumer protection and strengthen the cooperation between consumer protection authorities in the EU. It was first introduced in 2004[17] and updated in 2017 to take into account changes in the digital market.[18] One of the key features of the CPC Regulation is that it allows consumer protection authorities in the EU to work together to monitor and enforce consumer protection laws in a more coordinated and effective way.[19] This includes sharing information and coordinating enforcement actions across different member states, which can be particularly useful in cases where digital companies operate across multiple countries.

Digital enforcement by public authorities under the CPC Regulation typically involves the use of various tools and techniques to monitor, investigate, and take action against individuals or organizations that violate European consumer protection rules.[20] Generally, the powers granted to consumer authorities can be bundled in five main categories.[21] Public authorities may use *monitoring and surveillance techniques* to collect digital

[17] C Poncibò, 'Networks to Enforce European Law: The Case of the Consumer Protection Cooperation Network' (2012) Journal of Consumer Policy 175.

[18] V Balogh, 'Digitalization and Consumer Protection Enforcement' (2022) Institutiones Administrationis – J Adminstrative Science 85. See also DM Rao, 'International Consumer Protection Framework & Policy' (2021) International Journal of Law Management & Humanities 2439.

[19] C Goanta and G Spanakis, 'Discussing the Legitimacy of Digital Market Surveillance' (2022) Stanford Computational Antitrust 44.

[20] *Ibid*. 49; V Balogh, 'Digitalization and Consumer Protection Enforcement' cit. See also M Damjan and K Lutman, 'Administrative Enforcement of EU Consumer Law: A Disoriented Tiger in the Regulatory Jungle of E-Commerce' (2022) Journal of European Consumer and Market Law 130-138.

[21] For a comprehensive overview of the CPC Regulation, see C Goanta and G Spanakis, 'Discussing the Legitimacy of Digital Market Surveillance' cit. The proposed classification is the author's interpretation of the various types of investigation and enforcement options offered by the CPC Regulation.

data related to individuals or organizations suspected of violating digital laws or regulations. For instance, authorities may monitor online commercial communications, track digital footprints, and using data analysis tools to identify suspicious activity (*e.g.* art. 9(3)(a) CPC Regulation). Authorities may also use *blocking and filtering techniques* to restrict access to websites, content, or services that violate digital laws or regulations. This may include blocking access to websites that distribute illegal content or services that violate consumer protection laws (*e.g.* art. 9(4)(g)(i) CPC Regulation). In addition, public authorities may take *legal actions* against individuals or organizations that violate digital laws or regulations. This may include the power to bring about the cessation or the prohibition of consumer law infringements (*e.g.* art. 9(4)(f) CPC Regulation). Authorities may impose additional *administrative penalties* and fines on individuals or organizations that violate digital laws or regulations (*e.g.* art. 9(4)(h) CPC Regulation), and may work closely with the private sector, including internet service providers, social media companies, and e-commerce platforms to enforce digital regulations (*e.g.* art. 9(4)(b) CPC Regulation).

The CPC Regulation also brings new provisions to digital monitoring, compared to earlier regimes.[22] It provides consumer protection authorities with new powers to request information from online platforms and marketplaces, and to take enforcement actions against companies that violate European consumer regulation.[23]

An example of such a new provision is art. 9(d), which explicitly offers authorities the right to engage in mystery shopping. Mystery shopping is an approach traditionally used by companies to measure quality of service, staff performance, compliance with regulatory framework, and other specific topics related to effective functions related to the customers. It is a common assessment method across all commercial industries including travel, food, retail, and banks.[24] Given the proliferation of e-commerce, mystery shopping has become an increasingly important method on digital markets.[25] According to art. 9(3)(d) of the CPC Regulation, mystery shopping is defined as "the power to purchase goods or services as test purchases, where necessary, under a cover identity, in order to detect infringements covered by this Regulation and to obtain evidence, including the power to inspect, observe, study, disassemble or test goods or services". While the Regulation codifies earlier practices of mystery shopping, certain limitations remain. On the one hand, national interpretations of the CPC Regulation have led to questions of whether the powers granted to consumer authorities need to be further grounded in

[22] Regulation (EC) 2006/2004 of the European Parliament and of the Council of 27 October 2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws (the Regulation on Consumer Protection Cooperation).

[23] *Ibid*.

[24] R Chen and C Barrows, 'Developing a Mystery Shopping Measure to Operate a Sustainable Restaurant Business: The Power of Integrating with Corporate Executive Members' Feedback' (2015) Sustainability 12279.

[25] UNCTAD Working Group on Consumer Protection in E-commerce (June 2022) unctad.org.

national administrative law.[26] This is due to the applicability of administrative law across a wide spectrum of enforcement authorities, and the need to systematically clarify and harmonize sectoral practices. On the other hand, mystery shopping is resource-heavy. Resources include money for the purchases (which may vary depending on the type of good/service or sector), familiarity with investigation methods, or availability of automation in terms of processes, staff and computing power.

Mystery shopping has been traditionally undertaken manually. Scaling up these operations is possible, such as in the Commission's 2015 study on cross-border shopping and geo-blocking, where a total of 10,537 observations for 147 country pairs were analysed.[27] Large scale mystery shopping exercises have generally reflected coordinated shopping across Member States. However, mystery shopping can be further scaled using computational approaches. Large scale mystery shopping surveys have traditionally entailed the purchase of goods or services. Web crawling for the purpose of data collection does not, but it could be set up for purchases as well. An important addition is that the type of digital monitoring that can be performed through mystery shopping might also benefit from the development of data access. Web scraping entails collecting data (*e.g.* html code) from selected trader websites. However, marketplaces and other relevant market actors may also choose to standardize access to that data through making an Application Programming Interface (API) available to relevant authorities. APIs offer access to industrial level data, which can be useful to monitor activity on social media. The future of digital compliance entails that digital companies need to be asked to share data, and infrastructures and procedures for such data access need to be made available in administrative law, and further financial and knowledge resources need to be invested in digital compliance. The CPC Regulation is a good example of a broad legislative mandate that was given to consumer authorities to conduct Internet investigations. The practical implementation of this Regulation will pose considerable issues, particularly because of the absence of a systematic and cohesive approach in how to undertake Internet investigations found at the intersection of substantive legal frameworks and available technologies. A solution to these problems which can ensure the fitness of the CPC Regulation for the coming decade is proposed under section IV.

II.2. DSA

The DSA is a recent platform regulation that aims to update and modernize the existing rules governing intermediary liability for Internet companies in the EU, and impose novel

[26] C de Rond, 'De toezichthouder als *mystery shopper*' (17 August 2018) www.recht.nl.

[27] M Cardona, 'Geo-blocking in Cross-border e-Commerce in the EU Digital Single Market' (2016) Institute for Prospective Technological Studies Digital Economy Working Paper www.econstor.eu. It is noteworthy that while the DSA is not a European consumer *acquis* instrument as such, it is highly relevant for consumer protection.

transparency and due diligence obligations, as well as novel enforcement architectures.[28] The DSA introduces an updated legal framework for online platforms, which places greater responsibility and liability on these platforms for illegal content and activity on their sites.[29] This will help to incentivize platforms to take more proactive measures to monitor and remove illegal content, and will make it easier for authorities to hold them accountable for any illegal activity that occurs on their sites. The DSA requires online platforms to implement measures to monitor and report on illegal content and activity on their sites. This includes requirements for platforms to establish complaint mechanisms for users to report illegal content, and to provide regular reports to authorities on their efforts to remove such content, as well as due diligence obligations such as conducting internal audits (art. 37) and setting up internal compliance mechanisms (art. 41).

The DSA also introduces new provisions to improve cooperation and coordination among authorities in different EU member states (Chapter IV). This includes establishing a new EU-wide regulatory body to oversee the enforcement of the DSA (the European Board of Digital Services – art. 61), as well as provisions for greater information-sharing and cooperation between member state authorities (*e.g.* art. 60 on joint investigations). The DSA also establishes new national regulatory bodies in this respect (the Digital Services Coordinators), which will be responsible for overseeing the enforcement of the new rules in each Member State (art. 49). The DSA also gives authorities new oversight and enforcement powers to ensure compliance with the new rules (art. 51).

One important aspect of the DSA framework is its data access provision (art. 40), which is designed to ensure that authorities have the necessary tools to monitor and enforce compliance with the new rules, and that independent academic researchers who are vetted can be granted access to platform data for analysis purposes. Under the DSA, online platforms will be required to provide authorities with access to relevant data in order to monitor and enforce compliance with the new rules. This provision may enable a new era for the collaboration between academic researchers and public authorities in the development of public interest technology, given that national authorities will have to undertake vetting processes and appoint researchers who ought to have access to platform data. Under art. 40, a legal compliance API could facilitate Internet investigations into illegal content.[30] While the DSA does not specify a standard API for accessing data, in practice, a DSA API for academic research could be designed to provide secure and

[28] Ecommerce Europe, 'EP IMCO Adopts DSA Report: Some Good Progress, but Key Concerns Remain to Be Addressed' (14 December 2021) ecommerce-europe.eu.

[29] C Cauffman and C Goanta, 'A New Order: The Digital Services Act and Consumer Protection' (2021) European Journal of Risk Regulation 1; C Goanta, T Bertaglia and A Iamnitchi, 'The Case for a Legal Compliance API for the Enforcement of the EU's Digital Services Act on Social Media Platforms' (2022) *ACM Conference on Fairness, Accountability, and Transparency* (ACM 2022) dl.acm.org.

[30] C Goanta, T Bertaglia and A Iamnitchi, 'The Case for a Legal Compliance API for the Enforcement of the EU's Digital Services Act on Social Media Platforms' cit.

controlled access to specific data points within digital platforms.[31] A DSA API would require researchers to authenticate themselves and obtain authorization before they can access the data. This could involve providing credentials as issued by the relevant authorities vetting the researchers who are supposed to be granted data access. The usefulness of a DSA API could be rooted in the collection of data points on user behaviour data, content moderation data, or platform policies, based on specific tasks outlined in the project applications submitted for the vetting process, as well as the standardization of data formats, which would enable researchers to easily process and analyse the data. Additional considerations need to be given to limits imposed on the rate of data retrieval, to prevent overloading the platform's infrastructure. This would ensure that researchers can access the data they need without disrupting the platform's operations. Most importantly, the API would need to comply with data protection regulations applicable to academic research, and could involve implementing measures to protect the confidentiality and security of the data, such as data encryption, anonymization, and access controls.

The DSA framework on data access is designed to ensure that authorities have the necessary tools and powers to monitor and enforce compliance with the new rules. By providing greater transparency and oversight of online platform practices, it can contribute to the development of technologies to monitor consumer harms on digital markets. The standardization of data access through products such as a DSA API for legal compliance could further facilitate this goal. The DSA is a good example of how the fitness of enforcement – particularly as far as data access is concerned – is seen as a technology problem, to which technology-related solutions must be applied (*e.g.* data access).

III. Computational measurements of influencer activity: a case study for digital enforcement

Section II focused on briefly discussing recent digital enforcement trends, including regulatory practices in the European Union. This section will continue this discussion by giving a concrete example of a monitoring activity which has become increasingly popular with consumer authorities: the monitoring of social media influencers. From the perspective of the CPC, influencer activities are relevant for the enforcement of consumer law, while for the DSA, influencer marketing could be argued to be a systemic risk (art. 34(1) DSA) which may raise a lot of data access and investigation questions in the future.

The growth of digital markets has been leading to new iterations of consumer harms, and nowhere is that clearer than in the case of consumer manipulation through native

[31] See the proposal for a legal compliance API, and more basic descriptions of what an API is and what it can achieve, C Goanta, T Bertaglia and A Iamnitchi, 'The Case for a Legal Compliance API for the Enforcement of the EU's Digital Services Act on Social Media Platforms' cit.

advertising. Based on electronic word-of-mouth marketing (eWOM)[32] and parasocial relationships,[33] influencer marketing is an increasingly popular form of native advertising. It entails advertising which is embedded in non-advertising content. In the case of influencer marketing, advertising is embedded in the organic content made by influencers – also known as content creators – who make online content on a professional basis.[34] In other words, influencer marketing is a form of marketing where brands collaborate with individuals who have a significant social media following (known as influencers) to promote their products or services.

Public authorities are struggling to understand and investigate the scope of this market and its harmful implications. This is due to the fact that authorities often have to deal with a very high degree of information asymmetry, explained – amongst others – by the opacity of platform governance. For instance, even the most basic question of "how many social media influencers exist in jurisdiction x?" is very challenging to answer in practice, especially without access to structured data from social media platforms.

While influencer marketing can be an effective way for brands to reach new audiences, and for individuals to engage in new entrepreneurial activities, it can also lead to a number of harms. Four examples of harms are discussed in what follows. First, influencer marketing can sometimes be deceptive if it is not clear that a post or endorsement is sponsored.[35] This can mislead consumers into thinking that an influencer's endorsement is based on their genuine opinion, rather than being a paid advertisement.[36] Second, this type of marketing often promotes beauty standards that are unrealistic and unattainable for most people, which can contribute to body image issues and low self-esteem, especially among young people. This has led some jurisdictions such as Norway

[32] AB Rosario, K de Valck and F Sotgiu, 'Conceptualizing the Electronic Word-of-Mouth Process: What we Know and Need to Know about EWOM Creation, Exposure, and Evaluation' (2020) Journal of the Academy of Marketing Science 422; S Chu and Y Kim, 'Determinants of Consumer Engagement in Electronic Word-of-Mouth (EWOM) in Social Networking Sites' (2011) International Journal of Advertising 47; S Doh and J Hwang, 'How Consumers Evaluate EWOM (Electronic Word-of-Mouth) Messages' (2009) CyberPsychology & Behaviour 193; M Lee and S Youn, 'Electronic Word of Mouth (EWOM): How EWOM Platforms Influence Consumer Product Judgement' (2009) International Journal of Advertising 473.

[33] C Lou and H K Kim, 'Fancying the New Rich and Famous? Explicating the Roles of Influencer Content, Credibility, and Parental Mediation in Adolescents' Parasocial Relationship, Materialism, and Purchase Intentions' (2019) Frontiers in Psychology 2567; AN Tolbert and KL Drogos, 'Tweens' Wishful Identification and Parasocial Relationships With YouTubers' (2019) Frontiers in Psychology 2781.

[34] M De Veirman and others, 'Unravelling the Power of Social Media Influencers: A Qualitative Study on Teenage Influencers as Commercial Content Creators on Social Media' in C Goanta and S Ranchordás (eds), *The Regulation of Social Media Influencers* (Edward Elgar Publishing 2020) www.elgaronline.com; C Goanta and S Ranchordás, *The Regulation of Social Media Influencers* cit.

[35] M Swart and others, 'Is This An Ad?: Automatically Disclosing Online Endorsements On YouTube With AdIntuition' *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (ACM 2020) dl.acm.org.

[36] LE Bladow, 'Worth the Click: Why Greater FTC Enforcement Is Needed to Curtail Deceptive Practices in Influencer Marketing' (2017) William & Mary Law Review 1123.

to adapt consumer legislation in such a way that it now also requires influencers to disclose synthetic media (*e.g.* the use of filters) in order to alleviate mental health concerns.[37] Third, another health concern is physical. Some influencers use their platform to promote unhealthy behaviours, such as extreme dieting or excessive exercise, which can be harmful to their followers, or can be based on the promotion of harmful products which would pose issues for product liability laws. Fourth, some influencers may exploit their followers by promoting products or services that are overpriced or of poor quality, or by engaging in fraudulent practices such as fake giveaways or scams.[38]

The popularity of influencer marketing[39] has also created new challenges for public authorities in their efforts to enforce consumer protection laws and prevent misleading or deceptive advertising. One potential solution to these challenges is the computational analysis of influencer activities which can provide valuable insights into the reach and impact of influencer marketing campaigns. Such computational approaches include the use of data analysis tools and techniques to measure and analyse the performance of social media influencers. This includes data on their audience demographics, engagement rates, and the impact of their content on consumer behaviour. By using such computational approaches, public authorities can gain a better understanding of the reach and impact of influencer marketing campaigns, and can more effectively identify and target cases of illegal or deceptive advertising.

There are several ways that public authorities can use computational approaches for digital enforcement in relation to harmful influencer activities. First, they can use these tools to identify potential cases of misleading or deceptive advertising by influencers. By analysing data on engagement rates and audience demographics, authorities can identify cases where influencers are promoting products or services in a way that is likely to mislead consumers. This can help to target enforcement efforts more effectively and ensure that consumers are protected from fraudulent or deceptive advertising.

Second, computational approaches can be used to monitor the activity of social media influencers and ensure that they are complying with relevant consumer protection laws over time. By tracking the performance of influencers over time, authorities can identify trends and patterns in their behaviour and take action when necessary. This can lead to requesting information from influencers or their sponsors, conducting investigations and taking legal action when appropriate.

Finally, such approaches can be used to educate consumers about the risks and benefits of influencer marketing. By analysing data on consumer behaviour and attitudes, authorities can identify the most effective ways to communicate with consumers and

[37] BBC, 'Influencers React to Norway Photo Edit Law: "Welcome Honesty" or a "Shortcut"?' (6 July 2021) www.bbc.com.

[38] Federal Trade Commission, 'Social Media a Gold Mine for Scammers in 2021' (25 January 2022) www.ftc.gov.

[39] S Skalbania, 'Advising 101 for the Growing Field of Social Media Influencers Comments' (2022) WashLRev.

help them make more informed decisions about the products and services they purchase. This can include developing educational campaigns, creating online resources or guides, and providing consumers with tools and resources to protect themselves from fraudulent or deceptive advertising.

More specifically, the computational approaches referred to above reflect a wide array of methods grounded in computer science, which can be used for data collection and analysis for the sake of identifying consumer protection violations. Two examples of such approaches are Natural Language Processing (NLP) and network analysis. NLP entails using machine learning and computational linguistics techniques to analyse and understand human language. It can be used to monitor social media influencers by analysing the sentiment of their posts and identifying common themes or topics. NLP is also often used to understand whether influencers comply with disclosure obligations. NLP techniques can be used to search for specific keywords in social media posts that indicate that the post is sponsored or an advertisement. Words like "sponsored," "ad," or "paid" may be used as indicators of a sponsored post.[40] NLP approaches can also analyse the context of a social media post to determine whether it is an advertisement or a sponsored post. NLP can analyse the language used in a post to detect whether the influencer is promoting a product or service in exchange for compensation.[41] Named Entity Recognition (NER) is a technique used in NLP to identify and classify named entities in text, such as people, organizations, or locations. NER can be used to identify when an influencer or brand is mentioned in a social media post, indicating that it may be a sponsored post. Lastly, NLP techniques can be used to analyse the sentiment of social media posts to determine whether the influencer is endorsing a product or service in a positive or negative way. This can help determine whether sponsored posts are considered positive from the perspective of consumers.[42]

The second field of computer science that can be relevant in the context of computationally measuring influencer activities is network analysis. Network analysis involves studying the connections between entities, such as people or organizations. It can be used to monitor social media influencers by analysing, for instance, the network of followers and connections they have, identifying potential fraud or fake activity, and uncovering new opportunities for collaboration or partnership. Network analysis can be used to identify

[40] D Ershov and M Mitchell, 'The Effects of Influencer Advertising Disclosure Regulations: Evidence From Instagram' *Proceedings of the 21st ACM Conference on Economics and Computation* (ACM 2020) dl.acm.org; X Fang and T Wang, 'Using Natural Language Processing to Identify Effective Influencers' (2022) International Journal of Market Research 611.

[41] JP Santos Rodrigues, AC Munaro and E Cabrera Paraiso, 'Identifying Sponsored Content in YouTube Using Information Extraction' *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (IEEE 2021) ieeexplore.ieee.org.

[42] E Lahuerta-Otero and R Cordero-Gutiérrez, 'Looking for the Perfect Tweet: The Use of Data Mining Techniques to Find Influencers on Twitter' (2016) Computers in Human Behaviour 575; B Auxier, C Buntain and J Golbeck, 'Analysing Sentiment and Themes in Fitness Influencers' Twitter Dialogue' in NG Taylor and others (eds), *Information in Contemporary Society* (Springer International Publishing 2019) link.springer.com.

influential users in a social media network based on factors such as the number of followers, engagement rates, and the frequency and type of content posted.[43] These influencers may be more likely to engage in influencer marketing and promote products or services. Influential influencers (based on size or engagement) could be an enforcement priority. Network analysis can also be used to detect collaborations between influencers and brands based on patterns of interactions and connections in the network. For example, if multiple influencers are promoting the same product or service at the same time, it may be an indication of a coordinated campaign. This approach can shed light into how the market for influencer marketing campaigns looks like, and which stakeholders (other than influencers) are involved. A recent study used network analysis in this way to identify vaping campaigns and brands in influencer networks.[44] Lastly, network analysis can be used to identify fake followers and bots in a social media network.[45] By analysing patterns of interactions and connections in the network, it may be possible to detect accounts that are not genuine and are being used to inflate follower counts and engagement rates. Such accounts may be further investigated and monitored for engaging in deceiving practices.

In addition, other computer science approaches may be used to undertake web measurement studies. One such study dealt with the collection of 500,000 YouTube videos and 2.1 million Pinterest pins including affiliate marketing, in order to measure how many of them included consumer disclosures. The findings revealed that only around 10% of the affiliate marketing content on both platforms contained disclosures.[46]

These computer science methods and approaches can be combined and tailored to specific needs to monitor the activity of social media influencers in a more effective and efficient way.

IV. Computational measurements of influencer activity: a case study for digital enforcement

Legal enforcement has always had a procedural or administrative framework in which evidence plays an important role. For digital markets however, with investigations having

[43] SA Ríos and others, 'Semantically Enhanced Network Analysis for Influencer Identification in Online Social Networks' (2019) Neurocomputing 71; C Yi-Hsuan Chen, WK Härdle and Y Klochkov, 'SONIC: SOcial Network Analysis with Influencers and Communities' (2022) Journal of Econometrics 177.

[44] J Vassey and others, 'E-Cigarette Brands and Social Media Influencers on Instagram: A Social Network Analysis' (2022) Tobacco Control.

[45] S Cresci and others, 'Fame for Sale: Efficient Detection of Fake Twitter Followers' (2015) Decision Support Systems 56; A Mehrotra, M Sarreddy and S Singh, 'Detection of Fake Twitter Followers Using Graph Centrality Measures' *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)* (IEEE 2016) ieeexplore.ieee.org; Y Zhang, J Lu, 'Discover Millions of Fake Followers in Weibo' (2016) Social Network Analysis and Mining 16.

[46] A Mathur, A Narayanan and M Chetty, 'Endorsements on Social Media: An Empirical Study of Affiliate Marketing Disclosures on YouTube and Pinterest' (2018) Proceedings of the ACM on Human-Computer Interaction 1.

to increasingly rely on technology, there is currently a lot of unclarity in terms of the strategies, goals and practices of Internet investigations in different regulatory sectors. In 2010, Garfinkel wrote about digital forensics that "[w]ithout a clear strategy for enabling research efforts that build upon one another, forensic research will fall behind the market, tools will become increasingly obsolete, and law enforcement, military and other users of computer forensics products will be unable to rely on the results of forensic analysis".[47] The two examples given in this quote – law enforcement and the military – accurately reflect the focus on criminal investigations and national safety that digital forensics has had in its early days. In the same paper, Garfinkel also outlined a short history of digital forensics, which was in 2010 around forty years old.[48] Other regulatory sectors – such as consumer protection – have much to learn from digital forensics, given the vast array of research and best practices that has been amassed in this field previously.

As explored earlier in this *Article*, there is no Internet policing without digital investigations and enforcement, and without policing, consumers are literally left to their own devices and their own literacy to escape the novel harms of the digital economy. Against this background, it is important to acknowledge the need to support a cohesive research agenda around the development of forensic science for the purpose of monitoring and identifying consumer protection violations. A lot of the technology that is needed in the exercise of investigation and enforcement powers by consumer or DSA authorities simply does not yet exist. For instance, as appealing and inspiring as the dark patterns study has been for regulators and administrative agencies around the world, making detection tools based on the proposed methodology remains a highly complex and constantly evolving task. This *Article* puts forth that consumer forensics should be recognized as a new, growing field of multidisciplinary research that aims to explore the application of digital forensic techniques and tools to the study of consumer harms and consumer protection. The goal of consumer forensics is to develop new insights and computational approaches that can help regulators, policymakers, and industry stakeholders better understand and prevent consumer harms on digital markets. The consolidation of such a field can contribute to ensuring the long-term fitness of the legal enforcement frameworks such as those addressed in section II.

The remainder of this section is dedicated to structuring a proposed research agenda for consumer forensics, focusing on four key points: the consolidation of data collection and analysis methods for consumer law; the classification of new forms of consumer harms for digital markets; the development of new approaches to consumer protection; and the exploration of the legal and ethical implications of consumer forensics.

[47] SL Garfinkel, 'Digital Forensics Research: The next 10 Years' (2010) Digital Investigation S64.
[48] *Ibid.*

IV.1. Consolidating data collection and analysis methods for consumer law

One important area of research in consumer forensics should be the development of new methods for analysing consumer data. This could involve the use of advanced statistical methods, machine learning, and other data science techniques to uncover patterns and trends in consumer behaviour. It could also involve the development of new data sources, such as consumer-generated content, that can provide insights into consumer preferences and behaviour.

While many of the computer science methods referred to in section III are not as such new, their application to consumer protection is. Fields such as digital forensics can provide insights into how data collection and analysis has been done so far in the context of other regulatory sectors, such as criminal law. It is essential that sectors which are new to digital forensics liaise with the relevant regulatory frameworks, authorities and practices pre-dating their own interest in and need for computational approaches for digital investigations and enforcement. This task can bring together various stakeholders such as public authorities, academia or civil society.

IV.2. Classifying new forms of consumer harms

Another important area of research in consumer forensics would be the investigation of new forms of consumer harms. This is probably one of the most difficult aspects of enforcing consumer protection on digital markets. Apart from regular harms which are known to offline economies as well (*e.g.* non-conformity in goods), the digital implications of practices that ought to be considered unfair, manipulative, aggressive, generally remain difficult to grasp (*e.g.* which dark patterns are harmful to consumers?).

Even the wide-spread example of dark patterns, which have inspired regulators around the world to take measures against them, are difficult to pinpoint in terms of their impact.[49] While some dark patterns such as obstructions (*e.g.* blocking consumers from specific actions) are harmful because they have been used to make it difficult for consumers to legally get out of subscriptions, not the same can be said for all other categories. For instance, dark patterns such as confirm-shaming, which are supposed to allegedly shame consumers into choosing a transaction, may be somewhat blown out of context. The claim that a button with such a question would amount to a consumer harm, because it would influence the consumer to take a decision they otherwise would not, is not a claim that has sufficient scientific grounding.[50] At the same time, the scientific modelling of consumer behaviour is becoming increasingly difficult to undertake, due to the many different variables that influence it.

[49] See for instance J Luguri and L Jacob Strahilevitz, 'Shining a Light on Dark Patterns' (2021) Journal of Legal Analysis 43.

[50] *Ibid*.

Multidisciplinary methodologies bringing together behavioural sciences, computer science and socio-legal inquiries into consumer decision-making should be used to identify criteria for the classification of digital harms and their impact, as well as criteria for the identification of harm seriousness.

IV.3. Developing new approaches to consumer protection

Consumer forensics could also explore new approaches to consumer protection.[51] What exactly would count as a new approach? Currently, positive consumer protection law has led to the creation of standards such as the "average consumer" which projects a certain level of knowledge and diligence onto consumer behaviour.[52] Similarly, rules on mandated disclosures are considered commonplace in European consumer law, but the reality of consumers not reading terms of service really challenges transparency paradigms which are based on the provision of information.[53] As a field of study, consumer forensics can help by iterating evidence-based policy-making until new approaches can be developed. What are the harms identified on digital markets, and what characterizes them? Such data-driven information can help clarify business practices and consumer expectations in ways which can lead to new paradigms in consumer regulation.

IV.4. Exploring the legal and ethical implications of consumer forensics

Finally, research in consumer forensics would need to consider the legal and ethical implications of using forensic techniques in the study of consumer behaviour. Some of these implications include algorithmic transparency, consumer privacy and proportionality. Data collection and analysis always include some form of automation. In administrative procedural frameworks, automation has led to some concerning outcomes, such as algorithmic bias.[54] For consumer forensics, since the focus of investigations is on markets and not individuals, this risk is somewhat reduced.[55] However, it is not inexistent. The case study discussed in section III dealt with influencer marketing. Influencers currently are commodified identities: on the one hand they may be commercial actors, but on the other hand, they retain an individual identity. Any automated investigations or enforcement that lead to measures against such stakeholders can learn from the earlier frameworks on algorithmic accountability developed within public administration. The

[51] See for instance Omri Ben-Shahar and Ariel Porat, *Personalized Law* (Oxford University Press 2021).

[52] R Incardona, and C Poncibò, 'The Average Consumer, the Unfair Commercial Practices Directive, and the Cognitive Revolution' (2007) Journal of Consumer Policy 21.

[53] O Ben-Shahar and C Schneider, *More Than You Wanted to Know: The Failure of Mandated Disclosure* (Princeton University Press 2014).

[54] S Ranchordás and L Scarcella, 'Automated Government for Vulnerable Citizens: Intermediating Rights' (2018) William & Mary Bill of Rights Journal 30(2).

[55] C Rosca and others, 'Digital Monitoring of Unlawful Dark Patterns: What Role for Public Interest Technology?' (CHI Position Paper 2021) drive.google.com.

influencer case study also reveals why individual privacy is equally an important consideration,[56] in which case consumer authorities should take appropriate measures to protect personal data. Similarly, proportionality is an important aspect of consumer forensics due to the fact that the methods must align to the objectives pursued.[57]

V. Conclusion

This *Article* took digital enforcement as a starting point and discussed how Internet investigations are increasingly reliant on technology, not only in fields such as criminal law, but increasingly in other fields as well, such as consumer protection. Cohesive public policies relating to the implementation of digital enforcement frameworks will determine whether regulatory reforms such as the CPC Regulation and the DSA are future-proof. The *Article* combined insights from computer science literature with some brief discussions of European reforms in procedural law, to showcase the overlapping interests between academic fields busy with the study of consumer protection violations through computational methods, and public (consumer) authorities with growing procedural powers for carrying out digital investigations and enforcement. The *Article* further exemplified this overlap by taking the measurement of influencer marketing activities as a case study for the discussion of how computer science approaches can be honed to monitor consumer protection violations. The goal of this case study was to also to visualize in a more concrete way what computational approaches can achieve. Lastly, the *Article* proposed a new field of study in the form of consumer forensics and outlined a research agenda focusing on four key points: the consolidation of data collection and analysis methods for consumer law; the classification of new forms of consumer harms for digital markets; the development of new approaches to consumer protection; and the exploration of the legal and ethical implications of consumer forensics. Overall, the goal of consumer forensics would be to develop new insights and tools that can help prevent consumer harm and promote more transparent, accountable, and consumer-friendly markets. By bringing together researchers from a range of disciplines, including computer science, economics, law, and psychology, consumer forensics has the potential to be a powerful new field for promoting consumer protection and enhancing our understanding of consumer behaviour.

[56] CW Savage, 'Managing the Ambient Trust Commons: The Economics of Online Consumer Information Privacy' (2019) Stanford Technology Law Review 95; BA Martin, 'The Unregulated Underground Market for Your Data: Providing Adequate Protections for Consumer Privacy in the Modern Era Notes' (2019) Iowa Law Review 865.

[57] M Oswald and others, 'Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and "Experimental" Proportionality' (2018) Information & Communications Technology Law 223; M Schuilenburg and R Peeters (eds), *The Algorithmic Society: Technology, Power, and Knowledge* (Routledge/Taylor & Francis Group 2021).