What Evidence is Needed to Prove the Existence of Dark Patterns?

JOHANNA GUNAWAN, Northeastern University, USA
COLIN M. GRAY, Purdue University, USA
CRISTIANA SANTOS, Utrecht University, The Netherlands
NATALIIA BIELOVA, Inria Centre at Université Côte d'Azur, France

1 INTRODUCTION AND BACKGROUND

"Dark patterns" are manipulative, deceptive design practices deployed in online services aimed, either by purpose on in effect, at influencing the decisions of users about their purchases, use of time, and disclosure of personal data. Effective enforcement of dark patterns requires a sound methodology for the identification and characterization of dark patterns, including potential harms. To measure dark patterns, empirical research has evaluated both users' reactions to deceptive designs [8, 9, 11, 13, 15–17] and their presence across digital services (spanning modalities including mobile apps [10, 14], websites [14, 18], voice assistants [21], e-commerce sites [18], social networks [19, 22], games [7], and privacy control mechanisms like consent banners [12, 20, 23].

However, few enforcement cases refer directly to dark patterns practices, and the only rare emerging examples are US-based (e.g., Epic Games [4], Credit Karma [6], Vonage [5]). Also, it is hard to trace when users studies and other empirical research are actually used by regulators in their regulatory actions and court proceedings: a study by Nouwens et al. [20] was used by an EU regulator (CNIL) to support a case against Facebook [3] and an industry study was used to sustain sanctions against Google [2] to demonstrate that newly-introduced requirements in the CNIL's recommendation on cookies and other trackers [1] had an impact on the consent acceptance rate when users were faced with such deceptive consent banners. Moreover, while findings from research studies are likely to have substantial probative value to regulators and policymakers, these studies are scattered across various disciplines, such as computer science, social science, design, and law, and are often not known to regulators at the moment of decision making, making research-to-policy translation harder.

In this position paper, we consider the current status of dark patterns research and enforcement in an effort to tighten the feedback loop between technical implementations of dark patterns within consumer technologies, scholarly findings of harms or negative outcomes, and subsequent policy responses. First, we explore open questions in the dark patterns regulatory space to understand what gaps remain in uniting legally relevant evidentiary standards and existing empirical work. Then, we present our proposed methodology for ongoing and future work at the intersection of technical research and policy. Finally, we discuss the intended outcomes of our work and their relevance to this workshop.

2 OPEN QUESTIONS IN EMPIRICAL METHODS AND EVIDENTIARY REQUIREMENTS

In spite of the growing body of both definitional and empirical work, *the methodological questions behind dark patterns enforcement cases are not yet resolved or fully articulated.* Within this framing, we question what sources of evidence and related methodologies are sufficient to show the presence of dark patterns practices that are now being declared illegal

Authors' addresses: Johanna Gunawan, gunawan.jo@northeastern.edu, Northeastern University, USA; Colin M. Gray, gray42@purdue.edu, Purdue University, 401 N Grant Street, West Lafayette, Indiana, USA, 47907; Cristiana Santos, c.teixeirasantos@uu.nl, Utrecht University, The Netherlands; Nataliia Bielova, nataliia.bielova@inria.fr, Inria Centre at Université Côte d'Azur, France.

under EU and US law, and are there certain methods more well suited to detect specific kinds of dark patterns? Are there different standards of evidence when comparing data protection and consumer deception claims or harms related to dark patterns? To what extent does the finding of the presence of dark patterns a matter of expertise and experience? And how can we build synergies between regulation and technology design practices in ways that consider societal impact and technological realities?

3 FUTURE WORK: SIMULTANEOUS DESIGN AND POLICY ANALYSIS

In this paper, we will work *to close this gap* by informing regulators and legal professionals of existing and ongoing research across disciplines (computer science, social science, design and law) about factual and reliable evidence of dark patterns to rely upon with certainty, therefore proactively contributing to a technology-and-policy-design integration.

To achieve this goal, we will pursue a bi-directional approach, analyzing two bodies of knowledge. First, we collect EU and US case law, identifying cases that have relied upon the presence of dark patterns as part of legal enforcement. This dataset includes recent regulatory enforcement decisions by the European Data Protection Board, Data Protection Authorities, FTC, and other poignant cases that draw on similar design issues regardless of a specific "dark patterns" reference.

Second, we evaluate existing methodologies used to detect, identify, or characterize dark patterns, constructing genres of potential evidence related to the presence or harms of dark patterns based on prior empirical work. We base this set of genres on a systematic review of dark patterns literature (a dataset which presently includes over 80 such sources published from 2013 to 2022). This range of methods includes expert inspection, small- and large-scale user studies, automated detection, and user/log data as potential areas for further investigation relating to the presence or harms of dark patterns. From these two sources, we then discuss regulators' evidentiary requirements and thresholds of proof that are actually used in case law, and compare these to the methods used in empirical dark patterns research.

4 INTENDED OUTCOMES: TOWARDS IMPROVED DARK PATTERNS AUDITING AND COLLABORATIVE EFFORTS

Following our analyses, we will propose which methods are most appropriate for regulators to assess dark patterns, and articulate the kinds of evidence that need to result from these methods to provide a reasonable burden of proof. In this section, we outline which type(s) or level(s) of dark patterns are most likely (and best suited) to be identified and/or characterize dark patterns by method type. Our bi-directional analysis correlates methodology to legal evidence, building towards a unified understanding of how regulators and researchers may audit dark patterns collaboratively and design effective policies against their deployment. Our work will present opportunities for future synergies across multiple stakeholder groups and identify existing best practices for dark patterns auditing, which may include strategies for planning and packaging dark patterns HCI research for greater policy influence.

REFERENCES

[1] 2020. Délibération n. 2020-092 du 17 septembre 2020 portant adoption d'une recommandation proposant des modalités pratiques de mise en conformité en cas de recours aux « cookies et autres traceurs. https://www.cnil.fr/sites/default/files/atoms/files/recommandation-cookies-et-autres-traceurs.pdf

[2] 2021. Deliberation of the restricted committee No. SAN-2021-023 of 31 December 2021 concerning GOOGLE LLC and GOOGLE IRELAND LIMITED. https://www.cnil.fr/sites/default/files/atoms/files/deliberation_of_the_restricted_committee_no._san-2021-023_of_31_december_2021_concerning_google_llc_and_google_ireland_limited.pdf.

[3] 2021. Deliberation of the restricted committee No. SAN-2021-024 of 31 December 2021 concerning FACEBOOK IRELAND LIMITED. https://www.cnil.fr/sites/default/files/atoms/files/deliberation_of_the_restricted_committee_no._san-2021-024_of_31_december_2021_concerning_facebook_ireland_limited.pdf.

[4] 2022. Fortnite video game maker Epic Games to pay more than half a billion dollars over FTC allegations of privacy violations and unwanted charges. https://www.ftc.gov/news-events/news/press-releases/2022/12/fortnite-video-game-maker-epic-games-pay-more-half-billion-dollars-over-ftc-allegations. https://www.ftc.gov/news-events/news/press-releases/2022/12/fortnite-video-game-maker-epic-games-pay-more-half-billion-dollars-over-ftc-allegations Accessed: 2023-1-25.

[5] 2022. FTC action against Vonage results in $100 million to customers trapped by illegal dark patterns and junk fees when trying to cancel service. https://www.ftc.gov/news-events/news/press-releases/2022/11/ftc-action-against-vonage-results-100-million-customers-trapped-illegal-dark-patterns-junk-fees-when-trying-cancel-service. https://www.ftc.gov/news-events/news/press-releases/2022/11/ftc-action-against-vonage-results-100-million-customers-trapped-illegal-dark-patterns-junk-fees-when-trying-cancel-service Accessed: 2022-11-4.

[6] 2023. FTC finalizes order requiring credit karma to pay $3 million and halt deceptive 'pre-approved' claims. https://www.ftc.gov/news-events/news/press-releases/2023/01/ftc-finalizes-order-requiring-credit-karma-pay-3-million-halt-deceptive-pre-approved-claims. https://www.ftc.gov/news-events/news/press-releases/2023/01/ftc-finalizes-order-requiring-credit-karma-pay-3-million-halt-deceptive-pre-approved-claims Accessed: 2023-1-24.

[7] Jacob Aagaard, Miria Emma Clausen Knudsen, Per Bækgaard, and Kevin Doherty. 2022. A Game of Dark Patterns: Designing Healthy, Highly-Engaging Mobile Games. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3491101.3519837

[8] Carlos Bermejo Fernandez, Dimitris Chatzopoulos, Dimitrios Papadopoulos, and Pan Hui. 2021. This Website Uses Nudging: MTurk Workers' Behaviour on Cookie Consent Notices. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 346 (oct 2021), 22 pages. https://doi.org/10.1145/3476087

[9] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. 2021. "I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous!" - Dark Patterns from the End-User Perspective. In *Designing Interactive Systems Conference 2021* (Virtual Event, USA) *(DIS '21, Vol. 1)*. Association for Computing Machinery, New York, NY, USA, 763–776. https://doi.org/10.1145/3461778.3462086

[10] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376600

[11] Colin M Gray, Jingle Chen, Shruthi Sai Chivukula, and Liyang Qu. 2021. End User Accounts of Dark Patterns as Felt Manipulation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), Article 372. https://doi.org/10.1145/3479516

[12] Colin M Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. 2021. Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI'21)*. ACM Press. https://doi.org/10.1145/3411764.3445779

[13] Paul Graßl, Hanna Schraffenberger, Frederik Zuiderveen Borgesius, and Moniek Buijzen. 2021. Dark and Bright Patterns in Cookie Consent Requests. *Journal of Digital Social Research* 3, 1 (Feb. 2021), 1–38. https://doi.org/10.33621/jdsr.v3i1.54 Number: 1.

[14] Johanna Gunawan, David Choffnes, Woodrow Hartzog, and Christo Wilson. 2021. A Comparative Study of Dark Patterns Across Mobile and Web Modalities. In *Conference Companion Publication of the 2021 Computer Supported Cooperative Work and Social Computing* (Virtual Event, Canada) *(CSCW '21)*. Association for Computing Machinery.

[15] Hana Habib, Megan Li, Ellie Young, and Lorrie Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22, Article 621)*. Association for Computing Machinery, New York, NY, USA, 1–27. https://doi.org/10.1145/3491102.3501985

[16] Jamie Luguri and Lior Jacob Strahilevitz. 2021. Shining a Light on Dark Patterns. *Journal of Legal Analysis* 13, 1 (March 2021), 43–109. https://doi.org/10.1093/jla/laaa006

[17] Dominique Machuletz and Rainer Böhme. 2020. Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR. *Proceedings on Privacy Enhancing Technologies* 2020, 2 (2020), 481–498. https://doi.org/10.2478/popets-2020-0037

[18] Arunesh Mathur, Gunes Acar, Michael J Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), Article No. 81. https://doi.org/10.1145/3359183

[19] Thomas Mildner and Gian-Luca Savino. 2021. Ethical User Interfaces: Exploring the Effects of Dark Patterns on Facebook. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21, Article 464)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3411763.3451659

[20] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376321

[21] Kentrell Owens, Johanna Gunawan, David Choffnes, Pardis Emami-Naeini, Tadayoshi Kohno, and Franziska Roesner. 2022. Exploring Deceptive Design Patterns in Voice Interfaces. In *2022 European Symposium on Usable Security* (Karlsruhe, Germany) *(EuroUSEC 2022)*. Association for Computing Machinery, New York, NY, USA, 64–78. https://doi.org/10.1145/3549015.3554213

[22] Brennan Schaffner, Neha A. Lingareddy, and Marshini Chetty. 2022. Understanding Account Deletion and Relevant Dark Patterns on Social Media. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 417 (nov 2022), 43 pages. https://doi.org/10.1145/3555142

[23] Michael Toth, Nataliia Bielova, and Vincent Roca. 2022. On dark patterns and manipulation of website publishers by CMPs. In *Proceedings on Privacy Enhancing Technologies Symposium*, Vol. 2022. 478–497. https://doi.org/10.56553/popets-2022-0082