Defence Against the (Digital) Dark Arts: Defining, Detecting and Measuring Unlawful Dark Patterns in the EU

Constanta Rosca[1], Bogdan Covrig[1], Catalina Goanta[1], Jerry Spanakis[1], Gunes Acar[2]

[1]Law and Tech Lab, Maastricht University
[2]imec-COSIC, KU Leuven

Abstract—Using semi-automated measurement techniques, a 2019 study by Mathur *et al.* revealed the widespread use of dark patterns on shopping websites: the study found ∼1800 instances of dark patterns on ∼11% of ∼11,000 popular shopping websites worldwide. The use of dark patterns on shopping websites to influence consumers' transactional decisions raises concerns relating to the fairness of commerce. In the EU, the Unfair Commercial Practices Directive aims to protect consumers from unfair commercial practices employed by traders. Dark patterns may also have legal implications on the traders' duties to disclose information to consumers under the Consumer Rights Directive. However, the relative newness of and the scale at which dark patterns are being deployed may impede consumer protection authorities' efforts to enforce the law. Building on the research by Mathur *et al.*, the main aim of this project is to explore ways to automatically measure trader compliance with the two legal instruments identified above using Data Mining and Natural Language Processing techniques. In doing so, we hope to develop a methodological framework that could allow consumer protection authorities to detect (potential) infringements in a way that minimizes manual analysis.

CONTEXT, LITERATURE AND RESEARCH QUESTIONS

Dark patterns are user interface design choices that coerce, steer or deceive users to make services more addictive, entice users to spend more money, or share more personal data [1], [2]. Mathur *et al.* [2] measured the prevalence of dark patterns to reveal their widespread use on shopping websites. Using semi-automated measurement techniques, the study found ∼1800 instances of dark patterns (15 types and 7 broader categories) on ∼11% of ∼11,000 popular shopping websites worldwide.

The use of dark patterns on shopping websites to influence consumers' transactional decisions raises concerns about the fairness of commerce. In the EU, a comprehensive legal instrument that aims to protect consumers from unfair commercial practices employed by traders is the Unfair Commercial Practices Directive (UCPD). This instrument can be used to determine the legality of dark patterns. Further, dark patterns may have legal implications on the traders' duties to disclose information to consumers (e.g. mandatory disclosures), many of which are outlined in the Consumer Rights Directive (CRD).

Faced with the need to establish the compliance of opaque business practices with a complex web of laws, public authorities are presently overwhelmed by the relative newness of and the sheer scale at which dark patterns are being deployed online. The lack of public interest technology that could help authorities detect violations of the law and enforce the law effectively is a research gap which interdisciplinary academia must fill.

Building on Mathur *et al.*, we explore ways to automatically measure trader compliance with the UCPD and CRD.

On the theoretical front, we first explore the legal characteristics of dark patterns. We then use legal labels based on European law to determine whether dark patterns comply with the legal requirements in the UCPD and CRD, as follows:

- Determine whether the dark patterns identified in the study by Mathur *et al.* may be deemed unlawful/potentially unlawful based on the UCPD. This instrument has an annex that lists practices which are unlawful in all circumstances. In addition, the directive has two different tests which we use to ascertain whether a business practice is potentially unlawful, on the basis of a case-by-case fairness assessment;
- Expand upon this taxonomy with other potential breaches of the UCPD and CRD which may qualify as new categories of (potentially) unlawful dark patterns.

In the empirical part of the project, we will explore how the measurement infrastructure developed by Mathur *et al.* can be complemented by Data Mining (DM) and Natural Language Processing (NLP) techniques to detect dark patterns that are (potentially) unlawful according to the UCPD and CRD. The project will look into patterns in webpage contents (using NLP) and design (using DM). The adapted methods will be used to measure the prevalence of legally questionable practices on popular European online retailers' websites according to Tranco rankings and WebShrinker categories. Additionally, we will explore the feasibility of multilingual infringement detection to mirror the multilingual EU market. A side goal of the project is to investigate third parties that offer dark patterns other than social proof notifications. Eventually, we hope to develop a methodological framework that could allow consumer protection authorities to detect (potential) infringements in a way that minimizes manual analysis.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. Narayanan, A. Mathur, M. Chetty, and M. Kshirsagar, "Dark patterns: Past, present, and future," *Commun. ACM*, vol. 63, no. 9, pp. 42–47, Sept. 2020, doi: https://doi.org/10.1145/3397884.

[2] A. Mathur et al., "Dark patterns at scale: Findings from a crawl of 11K shopping websites," in *Proc. ACM Hum.-Comput. Interact.*, vol. 3, CSCW, Nov. 2019, Article 81, doi: https://doi.org/10.1145/3359183.