Mapping the Landscape of Dark Patterns Scholarship: A Systematic Literature Review

Colin M. Gray
gray42@purdue.edu
Purdue University
West Lafayette, Indiana, USA

Lorena Sánchez Chamorro
lorena.sanchezchamorro@uni.lu
University of Luxembourg
Esch-sur-Alzette, Luxembourg

Ike Obi
obii@purdue.edu
Purdue University
West Lafayette, Indiana, USA

Ja-Nae Duane
jduane@bentley.edu
Bentley University, Departments of Information and Process Management
Waltham, Massachusetts, USA

ABSTRACT

Dark patterns are increasingly ubiquitous in digital services and regulation, describing instances where designers use deceptive, manipulative, or coercive tactics to encourage end users to make decisions that are not in their best interest. Research regarding dark patterns has also increased significantly over the past several years. In this systematic review, we evaluate literature (n=79) from 2014 to 2022 that has empirically described dark patterns in order to identify the presence, impact, or user experience of these patterns as they appear in digital systems. Based on our analysis, we identify key areas of current interest in evaluating dark patterns' context, presence, and impact; describe common disciplinary perspectives and framing concepts; characterize dominant methodologies; and outline opportunities for further methodological support and scholarship to empower scholars, designers, and regulators.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Security and privacy** → **Social aspects of security and privacy**.

KEYWORDS

dark patterns, systematic review, methodology

ACM Reference Format:
Colin M. Gray, Lorena Sánchez Chamorro, Ike Obi, and Ja-Nae Duane. 2023. Mapping the Landscape of Dark Patterns Scholarship: A Systematic Literature Review. In *Designing Interactive Systems Conference (DIS Companion '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3563703.3596635

1 INTRODUCTION

Tactics of technology manipulation—often described through the concept of "dark patterns" [3]—are increasingly ubiquitous in digital services, and regulators are beginning to act in banning the most aggressive practices under consumer protection and data privacy law [1, 2]. Almost since the beginning of this term being coined in 2010, HCI scholars and practitioners have been central to this discourse, an area that is only rising in prominence and volume of research publications. As Lukoff [8] reported in a 2021 CHI workshop on "dark patterns," scholarship on the topic is quickly rising within the HCI community and beyond, from a small trickle in the mid-2010s to more than two dozen publications per year starting in 2021. Obi et al. [11] similarly reported a rise in conversations about dark patterns on Twitter over this time period, demonstrating that not only is the issue of dark patterns of increasing interest and concern, but this discourse also involves a broader range of stakeholders over time—beginning with designers and technologists and now including social scientists, computer scientists, journalists, regulators, and law scholars.

While scholarship on the topic of dark patterns is quickly increasing, it is unclear which types of deceptive design practices need to be better understood, which types of patterns produce the most harm, and what kinds of studies regulators and legal professionals need to effectively identify, characterize, and sanction the use of dark patterns in technology systems. In this work-in-progress paper, we examine prior scholarship on dark patterns—including the breadth of methods and contexts employed and common framings and disciplinary motivations for studies—allowing us to identify opportunities for new research that both extends the state of the art and produces action in the form of practitioner guidance and regulatory sanctions. Additionally, our description of existing methods, contexts, framings, and contributions used in studies provides a pathway for scholars new to the space to explore avenues for contribution while also allowing existing researchers to better understand this emerging area of study. Our contributions in this PWIP work are two-fold. First, **we describe a work-in-progress analysis of common framings of dark patterns scholarship**, including dominant methodologies, contexts, and disciplines. Second, **we offer provocations for future scholarship opportunities, gaps, and areas of tension**.

2 METHOD

The methodology of this systematic review was guided by recommendations from the PRISMA report [10]. We used the following procedure to identify and screen literature to include in our analysis

DIS Companion '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9898-5/23/07.
https://doi.org/10.1145/3563703.3596635

(summarized in Figure 1). All stages were carried out by an research team comprising six members that had previously engaged in dark patterns scholarship including investigators located in the United States, Luxembourg, and India.

2.1 Literature Collection and Screening

We conducted a search on Google Scholar and the ACM Digital Library using the search string "dark patterns" (with quotation marks) on September 13, 2022. This search returned 6,810 results in Google Scholar, and 183 results in the ACM DL (sorted by "relevance" in both databases) with no filters used to limit the search results. We downloaded the first 249 results from Google Scholar[1] and all 183 results were downloaded from the ACM DL, all as BibTeX. We then loaded all BibTeX entries into Rayyan.ai[2], a tool for collaboratively conducting systematic literature reviews. Based on our review of the resulting 432 titles and abstracts, 64 duplicate records were identified and removed. In addition, 25 records that were not written in the English language were removed. Finally, we attempted to download the full text of the remaining 343 reports and were able to retrieve all but two reports.

We then screened the remaining reports for eligibility. The following four inclusion criteria were: 1) the record had to be written in English; 2) the record had to mention "dark patterns" explicitly in the full text; 3) the record had to be published in a journal, conference proceedings, government technical report, or similar archival venue; and 4) the record had to include at least one empirical component. We excluded reports based on the following factors: reports that used an ineligible publication type (e.g. popular press article, only an abstract, workshop paper, student thesis, preprint; n=89); reports unrelated to computing or dark patterns (e.g., medical or hard science publications; n=17); reports that did not use dark patterns as a primary analytic or conceptual framing (e.g., only referencing the term in an introduction or conclusion; n=131); and reports which were not empirical in their framing (e.g., no new data collected or analyzed; only focused on argumentation; n=25). These criteria resulted in the exclusion of 262 reports from this review. This procedure resulted in 79 reports that were eligible for the analysis phase of our review.

2.2 Literature Analysis

To analyze the included literature, we used inductive qualitative content analysis [6] to answer the following questions: 1) What is the dominant context being addressed?; 2) What method(s) are used to empirically investigate instances of dark patterns?; 3) What framing concept(s) are used to motivate the study of dark patterns; 4) What is the publication type?; and 5) What appears to be the primary field for the intellectual contribution? To achieve this goal, four researchers on the team read the abstracts and relevant portions of the reports to ensure inclusion and code the selected records. The coding dimensions varied depended on the record and its contents. For example, a paper may contain more than one study on dark patterns and include varied evaluation methods and contexts. We began with a preliminary round of codes generated by the principal investigator based on an initial review of more than half of the dataset with sub-codes for each area of evaluation. This analysis was then supplemented by a review of five randomly distributed reports in the dataset by four other researchers to confirm the relevance and breadth of the preliminary codes, resulting in the addition of multiple sub-codes. Initial analysis work informed our creation of a codebook (Table 1), which we then used to code the entire dataset of 79 reports. One researcher coded all reports and then confirmed with at least one additional researcher. The principal investigator also provided consistency checks of codes across the entire dataset, engaging in conversation with the other researchers to reach full agreement on all application of codes. All categories of codes were applied non-exclusively except for the "framing" codes which we applied exclusively. Our final code application and codebook are included as supplemental material to extend our work.

[1]We may have missed additional relevant literature beyond the first 249 results but after the first 150 results, few results showed evidence of citations and the vast majority of records appeared to not be relevant to the search. Future work could address this broader swath of results and indicate if we missed any important literature.
[2]https://www.rayyan.ai

3 FINDINGS

We describe the outcomes of our analysis in relation to the context(s) addressed in the reports, the use of methods to support scientific inquiry, the framing and related contribution of the studies, publication type, and relevant disciplinary perspective of the publication venue. We report on the sub-codes we identified and focus on trends and gaps that we observed, supporting provocations for future work in Section 4.

3.1 Context(s) Being Studied

There was a broad diversity of contexts of study described in numerous ways, which we characterized as *specific functionality or domains* and *genres of sites or services*. The current landscape of research on dark patterns focuses disproportionately on genres where interaction occurs (e.g., social media, games, e-commerce; n=67) compared to precise indications of interface functionality, domains of use, or framings of interaction that include the use of dark patterns (e.g., consent banners, subscriptions, social robots; n=22).

- *Genres:* Our dataset included studies focused on many different genres of sites and services. Genres included *games* (n=7), *advertising* (n=2), *travel services* (n=2), *social media* (n=15), *e-commerce* (n=6), and *mobile apps* (n=7). Many studies also included a mixture of many different kinds of web services or genres which we coded as *digital services* (n=27).
- *Specific Functionality:* Studies were also framed based on the presence or impact of dark patterns in specific moments of interaction (e.g., consent banners), interactive flows (e.g., subscriptions), and interaction domains (e.g., social robots, VUIs, physical computing). By far, the most common functionality-driven approach to studying dark patterns was *consent banners* (n=16). *Subscriptions* were far less commonly studied with only one contribution. Emerging domains of interest in relation to dark patterns were also infrequently studied, yet likely to be important in future dark patterns research, including studies involving *voice user interfaces* (VUIs; n=1), *social robots* (n=2), and physical computing interactions (n=2)

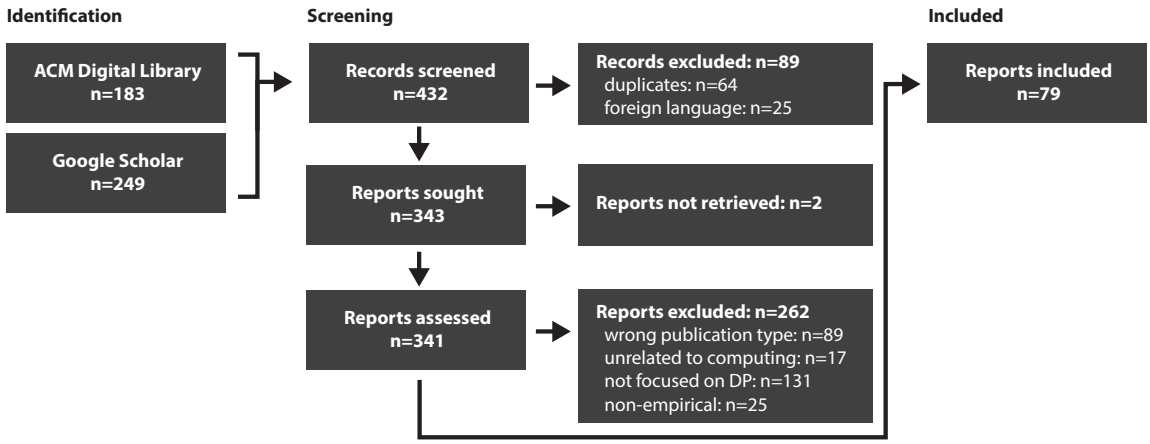

Figure 1: Systematic review flow diagram describing our identification and screening process.

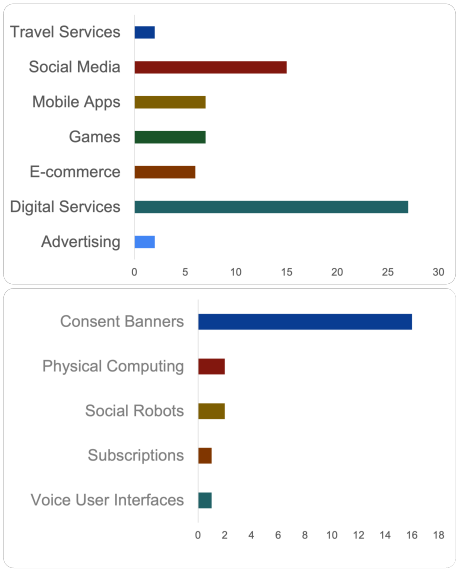

Figure 2: Distribution of papers by genre (top) and specific functionality (bottom)

through lenses such as proxemics and brain-computer interaction.

3.2 Method(s) Used to Study Dark Patterns

The most common method used to study dark patterns was *content analysis (n=36)*, through which scholars evaluated interface elements and identified whether they contained dark patterns or not. In some studies, content analysis was central to the study, but in other reports, content analysis was used to identify the presence of dark patterns that is then used to frame an experiment. The second most common method was *experimental* (n=14), in which the study included control and treatment groups to study the effects of particular design elements on user behavior. This method was particularly common in the context of consent banners, with half of the reports focusing on consent banner elements' impact on user behavior. The third most common method was *surveys* (n=10), used to gather users' behaviour, attitudes or opinions. *Literature review* of previous studies (n=8), the use of *case study* methodology to deeply interrogate the presence and/or impact of dark patterns (n=8), *design*-focused methods that involved the creation of new artifacts (n=7), *observations* (n=6), *interviews* (n=6), and *web measurement* (n=5) represent the remaining methods used in the reports we analyzed. The majority of studies attend to the existence of dark patterns through evaluation of the constituent elements in user interfaces, with less attention to the effects of dark patterns on users. Although these approaches might help policymakers and agencies in their consumer protection investigations, it may not contribute to a better understanding of the impact of dark patterns on users.

3.3 Framing of the Contribution

Reports included differing framings representing different methodological goals and limitations, which reflected the primary intent of the research contribution. A *descriptive* framing, using examples to illustrate the power, impact, or attributes of dark patterns, was the most prominent in our dataset (n=19). Notably, reports with this framing were often addressing issues in new contexts or through new approaches, and many of these reports utilized case studies or content analysis to support their contribution. A *problematizing* framing, identifying limitations of the taxonomies and finding gaps in the literature was the second most common approach (n=16). These reports also tended to use qualitative approaches, such as interviews, case study methodology, or content analysis to extend the literature base and identify areas for further study at scale. Reports with an *experimental/causal* framing sought to identify generalizable causal mechanisms relating to the presence of dark patterns, representing the third most common framing (n=15). Unsurprisingly, the vast majority of these studies used an experimental

Table 1: Codebook used by the research team to conduct confirmatory coding on the final dataset (n=79).

Context	
Genres of sites/services	
Advertising	any aspect of advertising platforms or advertisements
Digital Services	any website or service not described by any of the other types
E-Commerce	websites or apps used to purchase goods or services
Games	mobile or desktop games or game platforms
Mobile Apps	any apps on a smartphone or tablet
Social Media	any social media service (e.g., Facebook, Instagram, Tiktok, YouTube)
Travel Services	websites used to book air travel, lodging, or similar services
Specific functionality or domain	
Consent Banners	consent banner components as required by the GDPR or other regulatory statute
Physical Computing	elements of interaction in physical space and/or brain-body interfaces
Social Robots	elements of interaction with a social robot
Subscriptions	the process to sign up for or cancel a service
Voice User Interfaces	elements of interaction with a VUI
Methods	
Experimental	user behavior is characterized based on a comparison of control and experimental groups
Observation	user behavior is collected and tracked in real time, virtually or in person
Interview/Focus Group	a structured or semi-structured interview or focus group is conducted
Survey	user behavior or responses are collected through an online survey
Content Analysis	interface elements are evaluated qualitatively or/and quantitatively
Web Measurement	web sources are scraped and code evaluated, with some combination of manual and automatic inspection of the DOM or other code elements
Case Study	evaluation of a specific existing or speculative artifact that is used to support broader argumentation
Literature Review	systematic or otherwise disciplined review and synthesis of previous literature
Design	creation of new design alternatives to support downstream studies or as a means of speculative engagement
Diary Study	collection of data regarding technology usage or experiences over time
Framing	
Evaluative	leveraging existing taxonomies to identify whether something is an example of a dark pattern
Descriptive	using examples to illustrate power, impact, or attributes
Detection-Focused	creating and deploying an automated detection technique
Taxonomy-Building	defining new types of dark patterns or consolidating existing patterns
Problematizing	identifying limitations of taxonomies, identifying gaps in current literature
Experimental/Causal	identifying generalizable causal mechanisms

methodology to support their claims, with some reports also using observational or content analysis approaches. Reports with an *evaluative* framing (n=14) leveraged existing taxonomies of dark patterns to identify the presence of dark patterns in a new context with notable examples in social media, mobile apps, and consent banners. Reports with a *taxonomy-building* framing (n=12) focused on defining new types of dark patterns or consolidating existing patterns, with about half tending to explore a range of new contexts or lenses, such as privacy, games, social media, e-commerce, and physical computing. Reports with a *detection-focused* framing (n=3) were the least common, describing the creation and deployment an automated detection technique. Altogether, the diversity of framing contributions within the dark patterns literature revealed particular dimensions of dark patterns scholarship that could be used to inspire future research studies within and across these framing categories.

3.4 Contribution Type and Field

Conference publications were the most common, with 59 articles published at various conference venues, while journal publications (n=19) and a workshop paper (n=1) were less common. The articles published at these venues came from diverse disciplinary perspectives. Although we coded only at the report level and did not investigate individual authors, HCI venues were the most dominant at 40 articles. Other contributing disciplinary venues included Game Studies (n=5), Privacy and Security (n=4), Computing (n=20), Communication (n=1), Human Factors (n=1), Design (n=1), Marketing (n=1), Law (n=1), Tourism (n=1), Business (n=1), Rhetoric (n=1), and Others (n=5).These disciplinary perspectives implicitly surface the range of fields and knowledge bases that have been potentially impacted and interested in studying the manifestations of dark patterns. These articles were submitted to diverse—and generally high-prestige venues—including HCI venues with high impact potential (e.g., CHI (n=17), DIS (n=4), CSCW (n=6)) and other

computing venues that show the growing reach of dark patterns scholarship across international and non-US/EU populations (e.g., contributions in Brazil and India) and in relation to other computing perspectives (e.g., privacy and security; computing education; information systems; health informatics). However, because few studies reported the populations that they studied, it is difficult to assess specific trends—although it is clear that future research should embrace greater population diversity to address regional and culturally-specific manifestations of dark patterns.

4 PROVOCATIONS TO GROW AND SUPPORT A FUTURE LANDSCAPE OF DARK PATTERNS SCHOLARSHIP

Building on our analysis of the current empirical landscape of dark patterns scholarship, we are able to show not only the breadth of study designs, contexts, and framings, but also emerging consensus on best practices to continue to grow scholarship in this important area. In this section, we provide a set of provocations for future work to extend guidance on dark patterns. These provocations outline some emergent challenges and areas of scholarly consensus, and indicate some qualities of future scholarship that could increase the translation of efforts among scholars, design and technology practitioners, regulators, and legal professionals in the battle against dark patterns.

Dark Patterns Scholars Should Build Alignment and Community Norms. The diversity of the scholarship we analyzed—across differing methodological traditions, contexts, and fields of study—is a strength, revealing the harms of dark patterns in many of the systems that define our technologically-mediated existence. However, this diversity also comes with potential risks if we do not find areas of alignment and empirical norms. We prefer to think of these risks as *opportunities* rather than explicit gaps, since considerable and converging scholarship has already laid the groundwork for the empirical challenges that come next. First, as observed by Mathur et al. [9] and others, there is a lack of shared vocabulary which could be constructively addressed through the creation of a shared ontology to align scholarship across type and discipline. First steps towards this goal have been recently published by Gray, Santos, and Bielova [4], but substantial work remains. Second, we observed methodological pluralism across the landscape of scholarship which is a sign of strength, but within specific context and methodology categories, validity or reliability threats were present in some studies. For instance, there was variable quality in arguments derived from literature review-focused studies, and experimental studies also included differing levels of rigor depending on how they operationalized dark patterns (and at what level of granularity). To further support ecological validity, future work should also identify the kinds of demographic characteristics that should be collected for replication and shared knowledge building, alongside an evaluation of the ecological validity of various methods to address specific detection challenges. Third, we found distinct silos of different study framings but few connections across these silos. For instance, are there better ways for scholars to connect evaluation, problematization, and detection-oriented work? If we can solve this challenge, it could increase the translational effectiveness of the work, including new opportunities for transdisciplinary engagement and social impact via legal and regulatory channels that can leverage knowledge built through empirical studies.

Dark Patterns Scholars Should Outline Methodological Strengths and Weaknesses. We observed clear "types" of papers that also point to dominant methodological traditions. These types include: *exploratory work* to identify what dark patterns look like in specific places, naming their presence; *experimental work* to characterize the differential impacts of dark patterns building on existing typologies and exploratory work; and *detection-oriented work* that describes the presence of characterizable dark patterns at scale. Within each "type," there are common combinations of methods used, but there are also areas where more specific methodological guidance will be important to support the next wave of scholarship. For instance, researchers must be able to identify and characterize the presence of a dark pattern, which is almost always achieved through content analysis or other similar expert inspection technique. Specific standards of reporting will be important to allow this inspection and characterization to be falsifiable and encourage a shared language regarding the presence of dark patterns in specific, situated instances. It is also important for studies across these three types to engage in conversation with each other, including key methodological limitations. For instance, detection work has only been productive for very specific patterns and types of functionality (e.g., consent banners, e-commerce) and future work may need to include more explicit human-in-the-loop techniques to overcome the lack of detectability of many pattern types. Similarly, experimental work should leverage and acknowledge the ecological complexity of patterns in-the-wild (often addressed in exploratory work) and not seek to diminish or reduce that complexity in ways that undermine ecological validity, while also seeking to describe the power of dark patterns on users.

Dark Patterns Scholarship Should Offer Translational Opportunities. Although our focus was on empirical studies, we acknowledge the need to involve many stakeholders that benefit from this knowledge. As dark patterns are increasingly under legal scrutiny, how do we effectively translate and activate dark patterns knowledge across academic research, practitioner, regulator, and legal professional communities? One area of engagement that we left unexplored in this analysis could include emergent conversations in the legal community regarding dark patterns; due to differences in knowledge generation approaches, this argumentation-focused work (e.g., [7]) was excluded from our analysis in this paper but could provide cues towards transdisciplinary cooperation at the intersection of HCI and law (e.g., [5]). Other practices that could increase translation across communities could include leveraging shared vocabulary to indicate which specific type(s) of dark patterns are being studied, in what contexts or with what type(s) of users, and with what assumptions of the underlying definition of dark patterns. Our work could also identify common elements of reports based on the framing goal of the contribution to increase opportunities for meta-analyses of key areas of interest.

ACKNOWLEDGMENTS

This work is funded in part by the National Science Foundation under Grant No. 1909714, and by the Luxembourg National Research

Fund (FNR, Grant No. IS/14717072 Deceptive Patterns Online (Decepticon)). We also acknowledge the efforts of graduate researcher Ziqing Li and independent scholar Aditya Jain for their efforts in the analysis process.

REFERENCES

[1] 2022. *Behavioural study on unfair commercial practices in the digital environment : dark patterns and manipulative personalisation : final report.* Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/606365bc-d58b-11ec-a95f-01aa75ed71a1/language-en/format-PDF/source-257599418

[2] 2022. *Bringing Dark Patterns to Light Staff Report.* Technical Report. Federal Trade Commission. https://www.ftc.gov/reports/bringing-dark-patterns-light

[3] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. dl.acm.org, New York, NY, USA, 534:1–534:14. https://doi.org/10.1145/3173574.3174108

[4] Colin M Gray, Cristiana Santos, and Nataliia Bielova. 2023. Towards a Preliminary Ontology of Dark Patterns Knowledge. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. https://doi.org/10.1145/3544549.3585676

[5] Johanna Gunawan, Cristiana Santos, and Irene Kamara. 2022. Redress for Dark Patterns Privacy Harms? A Case Study on Consent Interactions. In *Proceedings of the 2022 Symposium on Computer Science and Law (CSLAW '22)*. https://doi.org/10.1145/3511265.3550448

[6] Hsiu-Fang Hsieh and Sarah E Shannon. 2005. Three approaches to qualitative content analysis. *Qualitative health research* 15, 9 (Nov. 2005), 1277–1288. https://doi.org/10.1177/1049732305276687

[7] M R Leiser and M Caruana. 2021. Dark Patterns: Light to be found in Europe's Consumer Protection Regime. *Journal of European Consumer and Market Law* 10, 6 (2021), 237–251. https://kluwerlawonline.com/journalarticle/Journal+of+European+Consumer+and+Market+Law/10.6/EuCML2021047

[8] Kai Lukoff, Ulrik Lyngs, Himanshu Zade, J Vera Liao, James Choi, Kaiyue Fan, Sean A Munson, and Alexis Hiniker. 2021. How the Design of YouTube Influences User Sense of Agency. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21, Article 368)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3411764.3445467

[9] Arunesh Mathur, Mihir Kshirsagar, and Jonathan Mayer. 2021. What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3411764.3445610

[10] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, and PRISMA Group. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine* 151, 4 (Aug. 2009), 264–9, W64. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

[11] Ikechukwu Obi, Colin M Gray, Shruthi Sai Chivukula, Ja-Nae Duane, Janna Johns, Matthew Will, Ziqing Li, and Thomas Carlock. 2022. Let's Talk About Socio-Technical Angst: Tracing the History and Evolution of Dark Patterns on Twitter from 2010-2021. (July 2022). arXiv:2207.10563 [cs.SI] http://arxiv.org/abs/2207.10563