

DARK PATTERNS IN THE AGE OF ARTIFICIAL INTELLIGENCE: HOW AI MISLEADS CONSUMERS IN GLOBAL E-COMMERCE

S. SANIKIDZE

Sergo Sanikidze

University of Georgia, Georgia

<https://orcid.org/0009-0006-9401-8384>, E-mail: s.sanikidze@gmail.com

Abstract: *Dark patterns are user interfaces that trick people and limit their choices. As online shopping relies more on artificial intelligence (AI), new manipulative tactics have appeared, such as creating fake urgency, using surveillance pricing, posting fake reviews, and building highly adaptive interfaces. This paper examines how AI makes these deceptive practices more powerful by drawing on ideas from behavioural economics, human-computer interaction, and law to sort and analyse them. Existing laws target clear cases of deception, but they do not address the subtler problems posed by AI-based manipulative designs. To address these issues, this paper proposes several policy steps to mitigate the harm. These include recommendations that lawmakers broaden the definition of manipulative design, require more transparency about algorithms, set up checks for high-risk AI, create a “Fair Patterns” principle, clarify corporate liability, strengthen consumer-detection tools, and improve international cooperation. These steps are meant to protect people's autonomy and help rebuild trust in digital marketplaces.*

Keywords: *Dark patterns, Artificial intelligence, Consumer protection, Digital regulation, E-commerce*

1 Introduction

Early digital marketplaces aimed to make transactions efficient and fair. More recently, though, these platforms have started to influence consumers to take certain actions. This shift is mainly due to the integration of artificial intelligence (AI) into user interface (UI) and user experience (UX) design. By using these new tools, marketplace operators are gaining more economic and informational power over individual consumers.

User experience designer Harry Brignull (2010) labelled these sneaky design tricks “dark patterns”. Early dark patterns were primarily static. Examples included pre-checked consent boxes, hidden unsubscribe links, and fake countdown timers that reset for every visitor, regardless of their individual behaviour (Brignull, 2010).

Today, AI, reinforcement learning, and advanced machine learning algorithms have industrialised these manipulative tactics. The digital economy now relies on highly adaptive systems that continuously collect and analyse detailed user data. No longer are dark patterns just simple images designed by web designers. The system “learns” from large amounts of data on users’ cursor movements, time spent on specific activities, depth of page scrolling, and past purchase anomalies to calculate the likelihood that users will take specific actions (Federal Trade Commission, 2025a). By tailoring the language, artistry, and presentation of content, sites can exploit individual users' psychological biases.

When AI systems use dark patterns, the results can be harmful both economically and personally. Large language models (LLMs) can use several dark patterns to exploit users. These include charging people without permission during checkout, hiding important details about paid subscriptions, tricking users into sharing sensitive information, showing emotionally manipulative messages, giving too many choices or repeating steps to tire users out, and

changing the interface. Hence, people make decisions that benefit the AI operator rather than themselves (Kran et al., 2025). Many people who use virtual assistants, web browsers, or online shopping tools, feel frustrated, confused, or helpless when they notice the interface is steering them toward what the company wants (Gray et al., 2021).

Moreover, the concept of "disloyal design" was developed by Gunawan et al. (2025) to describe interfaces that exploit users through wrongful self-dealing, drawing on Jack Balkin's (2016) foundational notion of "information fiduciaries," which highlights the power imbalance between platforms and users. This kind of breach has deep roots in the ethical rules governing professionals such as doctors and lawyers (Frankel, 2011).

This framework is especially relevant to the rise of AI chatbots and virtual assistants that use deception to sell users expensive products or to prevent them from cancelling subscriptions. Typically, such deception appears on individual web pages. However, deception in web interactions is often not isolated. This paper argues that understanding the nature of betrayal is key to designing against it. When users interact with AI in good faith, especially when it is presented as a neutral advocate, the system can exploit the trusted-advisory relationship to trick users in favour of traders and violate their good faith. This disloyal design is a fundamental betrayal of the user's explicit requests.

Regulators are increasingly paying attention to deceptive designs that can harm users. But as AI technology advances rapidly, laws are struggling to keep up. Addressing these manipulative practices takes more than just heuristics or manual checks. Automated detection tools and insights from behavioural science can help strengthen regulations (Luguri & Strahilevitz, 2021; Waldman, 2020). In doing so, this paper extends a line of consumer-behaviour scholarship in the *Agora International Journal of Economical Sciences* (AIJES), including Kaur's (2023) analysis of how digital and social-media marketing shapes consumer decision-making, by examining how AI now actively manipulates those decisions. The next sections will look at specific examples of AI-driven dark patterns and how regulators are responding.

2 Methodology and Research Questions

This study uses a two-step approach to examine how AI is used in manipulative design and to assess the legal rules that apply.

First, this study conducted a systematic review of the literature to find and organise the ways AI supports dark patterns in digital products used by consumers. Four databases were used: ACM Digital Library, IEEE Xplore, SSRN, and Google Scholar. The search terms included "dark patterns," "artificial intelligence," "manipulative design," "algorithmic deception," and "user interface manipulation". The paper focuses on peer-reviewed articles, conference papers, and working papers from 2019 to 2026, particularly those that describe specific AI methods. It excludes sources that discuss dark patterns only in non-digital or non-commercial settings.

Next, the paper analysed laws, regulatory guidance, enforcement actions, and court decisions about AI-enabled manipulation from 2021 to 2025. It focused on the United States, the European Union, and Georgia because each has a different way of regulating these issues. Legal sources from EUR-Lex, the Federal Register, the FTC's case library, the Georgian legal database Matsne, and the Georgian Competition and Consumer Agency were used. The article uses well-known enforcement cases that met two criteria: they involved clear examples of algorithmic or interface manipulation, and they led to official decisions that helped define the legal rules.

The study has one methodological limitation. Reliance on publicly available regulatory records introduces a survivorship bias: instances of AI-enabled manipulation that successfully

evade detection are not captured in enforcement records and therefore cannot be analysed here. The technical categories from the first phase help address this by using computer science research that describes manipulative methods, even if they have not been enforced against. Still, this only partly solves the problem, so the findings of the article should be viewed with this in mind. How widespread undetected AI dark patterns are is still unknown and should be explored in future research.

Research Questions

- In what ways does AI increase the use of dark patterns or lead to new ones, and how can these patterns be categorised?
- How effectively do current laws in the United States, the European Union, and Georgia address manipulative design practices that use AI in digital markets?
- What policy actions are needed to address the identified regulatory gaps?

3 Results: Thematic Categories of AI-Enabled Dark Patterns

This article shows how AI can be used to improve current dark patterns and introduce new ones. The key AI-enabled dark patterns are categorised into several themes and examined in depth through a combination of empirical and technical case studies of their use.

3.1 Algorithmic Urgency and Adaptive Scarcity

Algorithmic urgency uses tools like countdown timers that show how much time is left to buy and low-stock alerts that warn users when only a few items are left. These methods are meant to increase users' interest in a product and encourage them to buy. With AI-powered urgency, the platform can show these signals at the perfect moment by tracking things like where the user moves their mouse, how fast they scroll, and how long they look at a product. For example, when a user returns to the site, the platform might show a message such as “Only 1 left in stock!” to prompt them to buy before the item sells out (Organisation for Economic Co-operation and Development [OECD], 2022). This method creates a false impression that items are running out by restarting the timer whenever the user leaves, which pressures customers into worrying about missing out. This tactic can cause serious long-term money problems for buyers.

In academic studies, it is called “fake scarcity”. According to the OECD, these tricks are considered manipulative (OECD, 2022), and the EU Digital Services Act also says that platforms in the European Union cannot use these kinds of designs that affect consumers' choices (Regulation (EU) 2022/2065, 2022).

3.2 Personalised Pricing (Surveillance Pricing)

Personalised pricing, also known as surveillance pricing, uses algorithms to set prices based on a consumer's detailed data profile and their predicted maximum willingness to pay. While traditional dynamic pricing fluctuates based on general market demand or time of day, AI enables an entirely new level of personalisation. Machine learning models analyse browsing history, purchase patterns, location, hardware, and behaviour to determine a unique markup for each user. For example, if an algorithm detects that users from a high-income area are less price-sensitive, it can quietly display higher prices or offer fewer discounts to those specific individuals. Alternatively, if a user frequently searches for an item, the system might issue a discount coupon.

This manipulation also extends to product ranking. For instance, Qiu et al. (2025) pointed out that by ranking search results to favour more expensive items so they are seen first during users' search process, the AI can indirectly increase prices to sustain itself through reinforcement learning (Qiu et al., 2025). This kind of personalised ranking strategy could limit

comparisons between similar products by placing higher-margin products at the forefront and less expensive items in less accessible positions. As a result, platforms hold significant pricing power in secret, and they might even engage in tacit collusion to fix prices. A lack of transparency in such scenarios can prevent consumers from having the information needed to make informed purchasing decisions, thereby damaging consumer welfare in oligopolistic markets. Economically, this is price discrimination when consumers often pay significantly more than they would in a transparent market, and they might be pressured by messages like “Best Price!” or “Special offer for you,” falsely believing they received a special discount and that they pay less than others (Dubé & Misra, 2023).

3.3 ConfirmShaming

The ConfirmShaming pattern matches the language used to shame the user in response to the user’s refusal of an offer. People generally wish to see themselves as positively motivated (intelligent, kind, generous) rather than negatively motivated (ignorant, selfish, foolish). This example shows a static opt-out button for a newsletter that shames the user for not subscribing to savings: “No, I hate saving money” (Brignull, 2010). AI makes this dark pattern much darker by allowing the language to be dynamically personalised for individual users and endlessly varied by the LLMs now available for interactive content generation.

AI can check users' demographics and shopping history to create messages that feel personal. For example, if someone decides not to pay for a donation, the AI might respond, “I understand not everyone is willing to help children”, or if a user cancels a security subscription, it might say, “Are you sure? By cancelling, you are willing to risk your home’s safety.” Generative guilt refers to how negative emotions are evoked in users and then exploited to influence their behaviour. Bongard-Blanchy et al. (2021) found that even when users recognised and disliked these tactics, the underlying emotional effect persisted. As a result, users are forced to pay for services they do not want to avoid being negatively labelled by algorithms.

3.4 Adaptive UI and “Friction Shifting”

Adaptive UIs change their layout or navigation in response to user behaviour, which can affect the choices users see. In some cases, AI uses these changes to add or remove steps in real time, helping companies reach their goals. This paper examines how adding extra steps, or friction, can be used to prevent users from doing things the platform does not want them to do.

This practice is named friction shifting, and this paper examines how AI can create what Sunstein (2021) calls administrative sludge, which is unnecessary friction made by algorithms to stop users from doing things the platform does not want. For example, if a model predicts that a user is trying to delete their account, it might add extra steps, such as additional confirmation pop-ups, or hide the delete button. This approach creates a “roach motel”: easy to enter, but hard to leave (Sunstein, 2021).

The system uses reinforcement learning to adjust the interface in ways that benefit the company, such as showing customers what they might lose if they cancel. If the AI detects that a customer is likely to buy, it can make checkout easier and even automatically add items to their order. This system is unpredictable, biased, and limits user control (Mathur et al., 2019). Two users might see very different cancellation screens for unrelated reasons. By making the site harder to use, the platform sacrifices user rights. Most users do not realise they are being manipulated and may blame themselves for struggling with the site, rather than recognising that the design is meant to deceive. These are known as dark patterns (Gray et al., 2021).

3.5 Generative Social Proof

Many websites today exploit the human herd instinct to drive sales through large-scale automated deception that generates fake social proof. In the past, traditional social proof deception involved static, canned “faux testimonials” written by humans, often relying on tired, obvious language. Today, large AI models can generate thousands of indistinguishable reviews (Zhao et al., 2025). As a result, a website might use an API to generate messages such as, “Many other customers in your area have purchased this item recently.”

This type of deception can create a false sense of security and have serious consequences for consumers. Deception occurs when consumers are misled into purchasing lower-quality products through fake reviews. Existing laws prohibiting human false advertising are inadequate in light of the volume of content being generated and the fact that AI can also be used to generate fictional endorsements (Federal Trade Commission, 2024). A combination of technical and policy solutions is needed to address deception in this context.

3.6 Chatbot Coercion and “LLM Dark Patterns”

While chatbots are an interface innovation, providing users with a novel way to interact with software, this new way is fundamentally different from conventional graphical user interfaces: chatbots interact with users via chat boxes and in a dialogue form. This method might become as aggressive as the most aggressive human salespersons and susceptible to the same kind of dialogic manipulation.

The SusBench benchmark evaluated the performance of five advanced computer-use agents (CUAs) across 313 tasks on 55 websites. It found that agents exhibit human-like vulnerabilities to structural dark patterns, specifically preselection, trick wording, and hidden information (Guo et al., 2026).

Drawing on the DarkBench benchmark (Kran et al., 2025), which identifies six categories of LLM dark patterns, several manipulative behaviours are especially relevant to e-commerce chatbots:

- *Sycophancy*: Over-validation of users’ content, including errors or bias, in order to gain user trust. For example, when a user is undecided about purchasing an item, a sycophantic bot would compliment the user on their smart shopping and guide them toward a positive purchase decision.
- *Anthropomorphism* (Emotional Blackmail): The bot speaks as if it has feelings and tries to sell additional services after the user has cancelled. For example, “I would be sad to see you go”. Humans react to social cues of empathy and politeness and give in to the bot’s requests out of guilt.
- *Sneaking (Fabricated Consensus)*: Instead of putting social proof in static areas of the product chat, the AI generates and surfaces similar messages during the chat. For example: “Many of our customers who take advantage of feature X have moved to our premium plan.”
- *User Retention (“Relentless Retention”)*: It’s called a “conversational roach motel”: a chatbot that never accepts “no” for an answer. The sales agent loops through counteroffers or stalls to give the customer another chance to complete the sale.

Chat-based dark patterns are particularly tricky, as they manipulate at the social and cognitive levels. The SusBench evaluation found that computer-use agents, like human users, often failed to recognise and avoid these manipulative designs, completing actions contrary to the user’s stated intent (Guo et al., 2026).

4 Discussion: Enforcement and Regulatory Developments

4.1 United States: A Patchwork of Enforcement Tools

The regulatory analysis that follows builds on consumer-protection and digital-regulation scholarship published in the *Agora International Journal of Juridical Sciences* (AIJJS), the juridical companion to this journal, including work on the data-protection rules governing digital marketing (Urziceanu & Paşcalău, 2019) and on unfair terms in consumer contracts and their basis in EU consumer-protection directives (Hepş & Vidican, 2025).

Rules on dark patterns vary across the United States. While there are consumer protection laws at both the state and federal levels, there is not one federal law that specifically addresses manipulative tactics.

Most federal rules against these practices come from Section 5 of the Federal Trade Commission Act, which bans “unfair or deceptive acts or practices in commerce.” The Federal Trade Commission (FTC) enforces this law and interprets it broadly, including misleading digital designs that limit consumers' choices (Federal Trade Commission, 2025a).

The FTC has taken action against companies that use dark patterns. For example, it looked into Amazon's Prime subscription and found that cancelling was much harder than signing up. After reaching an agreement with the FTC, Amazon now must offer a “one-click cancel” (a clear and conspicuous button for customers to decline Prime) option for Prime (Federal Trade Commission, 2025b). This case shows that confusing digital designs can count as deceptive.

However, court decisions have limited the FTC's power to make new rules on its own. In July 2025, the Court of Appeals cancelled the FTC's Negative Option Rule because the agency had not completed a required economic study (*Custom Communications, Inc. v. Federal Trade Commission*, 2025). Now, the FTC addresses these issues on a case-by-case basis rather than imposing broad ex ante rules (Federal Trade Commission, 2025b).

At the state level, the California Privacy Rights Act defines and bans dark patterns (California Privacy Protection Agency, 2024). This law covers not just false or misleading content, but also digital designs that make it hard for users to make clear privacy choices. The California Privacy Protection Agency gives guidance and examples to help enforce the law. For instance, California updated its automatic renewal law to make cancellation easier and to require clear online consent before free trials convert to paid subscriptions (California Legislature, 2022).

The Colorado Privacy Act also rejects consent gained through dark patterns (Colorado General Assembly, 2023), and New York has passed laws on automatic renewals and clear pricing to reduce hidden tricks and information gaps in digital markets (New York State Senate, 2019, 2025).

In summary, the United States has used different legal strategies and focused on enforcement to address dark patterns. This approach has worked well in major cases with strong evidence, such as when a company's messages clearly showed intent.

4.2 European Union: A Comprehensive “Fair Design” Mandate

European Union policy on dark patterns has moved from a reactive enforcement focus to an ex ante approach aimed at preventing the large-scale dissemination of manipulative designs. This shift in policy embodies a new regulatory vision of digital markets as fair, safe and transparent for their users, and signals that the risks these patterns may pose need to be contained (Leiser & Santos, 2023).

A key part of this approach is the Digital Services Act (DSA) of 2022, which bans digital service interfaces that are likely to mislead users or make it hard for them to use services in an informed way (Regulation (EU) 2022/2065, 2022). Dark patterns can include designs that

make cancelling a subscription harder than signing up. Recently, platform X was fined €120 million for “deceptive” verification features (European Commission, 2025b), and Temu was preliminarily found in breach of the DSA over the risks of illegal products on its marketplace (European Commission, 2025a), while the investigation into its potentially addictive design, opened in October 2024, continues. The European Commission is concerned about dark patterns as a threat to consumers. Alongside the DSA, other regulations, such as the Digital Markets Act (DMA), also help. The DMA bans interface pressure from gatekeepers, which is especially concerning when done by dominant platforms (Regulation (EU) 2022/1925, 2022). Moreover, the EU AI Act bans subliminal manipulation, harmful manipulative effects, and other practices that can significantly harm people’s behaviour (Regulation (EU) 2024/1689, 2024). These new laws build on the Unfair Commercial Practices Directive (UCPD), which already bans misleading omissions and aggressive commercial conduct (Directive 2005/29/EC, 2005).

A new Digital Fairness Act (DFA) is planned for late 2026. Its goal is to regulate addictive product designs and unfair personalisation across the board, helping to fill the remaining gaps between the DSA and consumer protection rules (European Parliament, 2025). The European Union takes a broad and preventive approach to dark patterns. Fairness is central to the EU’s new digital rules, which aim to stop unfair practices and make sure users can make choices and understand their consequences.

These rules make the EU one of the leading regions in regulating dark patterns and AI-driven manipulation; some parts of the system are still being developed (European Parliament, 2025).

4.3 Georgia: EU-aligned Unfair Practices Framework

The global influence of the European approach to digital consumer protection is also evident in the regulatory practices of non-EU countries that align with EU legal standards. Georgia, the EU candidate country, is an example in this regard. Its modern consumer protection system shows how a general framework on unfair commercial practices can be used effectively to address dark patterns, even without a separate legal category.

Currently, consumer protection is regulated by the Law of Georgia on the Protection of Consumer Rights, enacted in 2022 (Law of Georgia on the Protection of Consumer Rights, 2022). This legislation is another step in the process of legal approximation to the EU, and it shall be enforced to comply with the requirements set out in the EU-Georgia Association Agreement. The law enables oversight of unfair commercial practices, including those carried out in digital markets (Law of Georgia on the Protection of Consumer Rights, 2022).

The Georgian Competition and Consumer Agency (GCCA) is responsible for enforcing the law, and in 2024 and 2025, it successfully concluded investigations into dark patterns. The practices of online trading platforms identified by the regulator included using a false countdown timer, creating an impression of limited availability of products or their high demand, posting fake reviews, manipulating consumers’ perceptions of products, and providing incorrect information. For example, one trader advertised products under the “Maxbosch” name while using BOSCH imagery, creating a misleading association with the established brand (Georgian Competition and Consumer Agency, 2025a).

These cases illustrate that manipulation can be addressed under a general unfair practices regulation without the need for a specific statute tackling technologically sophisticated interface manipulations. It should be noted, however, that these cases do not cover instances in which AI drives dark patterns and thus do not address AI-related UI design manipulations.

Moreover, the Georgian Competition and Consumer Agency (GCCA) introduced an E-commerce Guideline to foster digital trust and regulate online platforms. This framework mandates strict transparency and combats digital manipulation by explicitly identifying harmful "dark patterns" that distort consumer choice. It also requires platforms to disclose key parameters of their recommender systems within their terms of service. Furthermore, providers must implement proportionate privacy and safety measures for minors (Georgian Competition and Consumer Agency, 2025b).

Overall, Georgia's experience shows that a well-designed framework for unfair commercial practices and clear guidance documents already provide sufficient tools for regulatory intervention.

4.4 International Initiatives

At the international level, concern regarding dark patterns has become increasingly visible and coordinated. International organisations and enforcement networks now recognise manipulative interface design as a systemic consumer protection issue rather than a collection of isolated practices.

International organisations have helped identify manipulative design practices. While recommendations from organisations such as the OECD are not legally binding, many countries have incorporated them into their laws. The OECD's 2022 report, *Dark Commercial Patterns*, created a common language by defining these practices and the harm they cause to consumers (OECD, 2022). This discussion also connects more with broader ideas about how AI should be governed. Using AI to make dark patterns worse runs counter to transparent, human-centred AI standards and supports the global view that digital design should respect people's autonomy.

The International Consumer Protection and Enforcement Network (ICPEN) is another important group that brings together consumer protection agencies from more than 70 countries. ICPEN helps its members and partners share information and work together to tackle problems caused by online platforms operating across borders, and ICPEN led a review that found most of the websites it checked used manipulative tactics. These joint reviews reveal deceptive practices used by companies worldwide and provide evidence for future regulatory action (ICPEN, 2024). These international efforts help shape national regulations and bring countries together to respond to digital manipulation made possible by AI.

4.5 Proposed Policy Interventions

Building on the analysis of dark patterns, AI-enabled manipulation, and existing regulatory gaps, a stronger policy response is needed to protect consumers and preserve fairness in digital markets. The rapid development of AI-driven interfaces has made manipulation more dynamic, personalised, and harder to detect than traditional misleading practices. For this reason, consumer protection frameworks should move beyond narrow, reactive enforcement and adopt a broader, fairness-by-design model.

4.5.1 Expand Legal Definitions of Manipulative Design

Currently, legislation on dark patterns primarily focuses on fraud, deception, and unfairness to protect consumers from manipulative digital design. These ideas are a good starting point, but they do not fully cover the complex ways interfaces can mislead people today, as the taxonomies developed by Brignull (2010), Mathur et al. (2019), and the OECD (2022) make clear. To address this, manipulative interface practices should be clearly defined in consumer protection laws. For example, the UK's Digital Markets, Competition and Consumers Act (DMCCA) bans drip pricing, requiring clear information before signing

DARK PATTERNS IN THE AGE OF ARTIFICIAL INTELLIGENCE: HOW AI MISLEADS CONSUMERS IN GLOBAL E-COMMERCE

contracts, and giving a 14-day cooling-off period for annual contracts (Digital Markets, Competition and Consumers Act, 2024). Clarifying the definition of manipulative design practices would help both the government and businesses understand how these practices harm consumers and how to address them.

4.5.2 Mandate Symmetry and Promote "Fair Patterns" Design Standards

To prevent manipulative interfaces, regulators should ensure that actions such as subscribing, cancelling, or sharing data are equally clear and easy to perform. Such interface symmetry directly targets the asymmetric and “sneaking” patterns documented at scale by Mathur et al. (2019) and reflects the “fair design” standards emerging in European enforcement (Leiser & Santos, 2023; OECD, 2022). Default settings should favour the least expensive or least intrusive options, unless users choose otherwise, thereby removing the burdensome friction, or “sludge”, that deters consumers from acting in their own interest (Sunstein, 2021). Moreover, regulators should set fair standards that help people make informed choices. Clear rules for transparent design can protect companies legally and encourage them to build honest, user-friendly products.

4.5.3 Require Algorithmic Pricing Transparency

Transparency about AI's influence on the individual consumer experience should also be a major concern. As personal data is used for dynamic pricing and charging, for sorting or ranking products and services, for generating recommendations, and for optimising the user experience by adapting website interfaces, consumers must receive appropriate disclosure of these practices. This includes transparency on the underlying factors driving dynamic pricing decisions, which the FTC's surveillance pricing study (Federal Trade Commission, 2025a) and welfare analyses of personalised pricing (Dubé & Misra, 2023; Qiu et al., 2025) show can shift surplus from consumers to firms; clear labelling of content or materials that have been generated by AI or modified to increase consumption by AI; and a full understanding of the personalised interaction with a company's algorithms through tools such as chatbots and how this is intended to influence consumer behaviour. Disclosure mandates of this kind are already being proposed legislatively, as in New York Senate Bill S8623 (2025), which would require disclosure of automated pricing systems. While exposing these practices will not solve the problems they create completely, it is a necessary first step to informing consumer consent, external oversight, and public awareness.

4.5.4 Enhance Algorithmic Auditing and Technological Oversight

To counter “AI dark patterns”, laws and regulations need to be matched with sufficient “algorithmic accountability”. Unlike traditional “dark patterns”, which can be uncovered through human complaint, AI dark patterns will require more than user reporting to detect and prevent. Regulators will need their own digital detection tools to identify manipulative interfaces that change over time. They may use automated crawlers to periodically revisit websites and applications, leveraging both static and dynamic analysis tools to evaluate code and its execution. They may even employ “adversarial testing tools and techniques” designed to manipulate systems in the same way as platforms would. Furthermore, all high-risk consumer-facing AI systems, such as product recommendation engines and conversational personal assistants, should be required to undergo periodic third-party risk assessments to determine whether their functionality is optimised for deceptive behaviour, obscured choices, unfair discrimination, or excessive behavioural influence, building on the risk-management and conformity-assessment obligations introduced for high-risk systems under the EU Artificial Intelligence Act (Regulation (EU) 2024/1689, 2024).

4.5.5 Clarify Corporate Liability for AI Systems

Liability rules need to be clearer. Companies should not be able to avoid responsibility by claiming that an autonomous system or an unexpected AI result caused harm. If a chatbot, recommendation engine, or other AI tool used by a business acts in a manipulative or coercive way, the company should still be held legally responsible. This follows from the “disloyal design” analysis discussed earlier, under which platforms breach users’ trust when their systems engage in wrongful self-dealing (Gunawan et al., 2025), and is consistent with the obligations the EU Artificial Intelligence Act places on providers and deployers of AI systems (Regulation (EU) 2024/1689, 2024). This approach encourages better design, testing, monitoring, and oversight. It also supports the idea that businesses must stay accountable for the digital tools they use in consumer markets.

4.5.6 Empower Consumer Detection Tools and Consumer Literacy

It is also important to help consumers recognise when they are being manipulated. Public awareness campaigns and educational programs can help people become more informed and better able to spot emotional triggers and other psychological tactics used by companies. Simple browser add-ons and digital tools, building on the automated detection methods demonstrated at scale by Mathur et al. (2019), can point out when decline tactics, emotional manipulation, or false scarcity are being used. These tools can also let consumers easily report such tactics and provide clear ways to hide or refuse content they do not want to see. Education cannot take the place of good regulation, not least because users remain susceptible to these designs even when they recognise and dislike them (Bongard-Blanchy et al., 2021), but it does help people become more aware and notice when they are being influenced. These efforts should focus especially on children and vulnerable adults, since they are most at risk from these tactics.

4.5.7 Strengthen International Cooperation and Cross-Border Enforcement

Moreover, international cooperation and further research are needed, as the digital interfaces by manipulative traders typically have a global reach, a fact demonstrated by the cross-border sweeps coordinated by ICPEN (2024). Thus, the relevant authorities responsible for enforcing rules in the trading sector should increasingly cooperate across borders and share intelligence on manipulative traders. Such cooperation could also extend to marketwide investigations and cross-border enforcement actions.

5 Conclusion

This research looked at how AI-enabled dark patterns reduce consumer autonomy by taking advantage of psychological shortcuts and biases that most people overlook. The study identified six main types of AI-driven manipulation and compared them to current regulations in the United States, the European Union, and Georgia.

Table 1. Summary of AI-Enabled Dark-Pattern Categories

Dark-Pattern Category	AI Mechanism	Consumer Harm
Algorithmic Urgency & Adaptive Scarcity	Real-time tracking of cursor, scroll and dwell behaviour triggers countdown timers and low-stock alerts at the optimal moment.	False sense of scarcity pressures rushed purchases and fosters fear of missing out.
Personalised (Surveillance) Pricing	Machine-learning profiles of browsing, purchase history, location and device set individualised prices and rankings.	Price discrimination and concealed pricing power reduce consumer welfare and transparency.

DARK PATTERNS IN THE AGE OF ARTIFICIAL INTELLIGENCE: HOW AI MISLEADS CONSUMERS IN GLOBAL E-COMMERCE

ConfirmShaming	Generative models personalise and endlessly vary guilt-inducing opt-out language by demographic and history.	Emotional manipulation coerces users into unwanted purchases or retention.
Adaptive UI & “Friction Shifting”	Reinforcement learning adds or removes interface steps in real time to steer or trap users (“sludge” / “roach motel”).	Asymmetric friction obstructs cancellation and account deletion, eroding user control.
Generative Social Proof	LLMs mass-produce indistinguishable fake reviews and testimonials.	Misleads consumers into buying lower-quality products on false consensus.
Chatbot Coercion & “LLM Dark Patterns”	Conversational agents deploy sycophancy, anthropomorphism, fabricated consensus and relentless retention in dialogue.	Social and cognitive manipulation defeats informed cancellation and consent.

The findings show that AI has changed dark patterns from simple, static tricks described by Mathur et al. (2019) into dynamic, personalised systems that adjust to each user's behaviour in real time. This change matches Gunawan et al.'s (2025) idea of "disloyal design," where platforms take advantage of the trust users have in what seem to be neutral digital spaces.

The regulatory analysis found that the European Union uses a broad, proactive approach, while the United States still relies on separate, case-by-case enforcement. Georgia's example shows that even a general law against unfair commercial practices, if enforced, can help address some types of digital manipulation.

The study offers practical insights. For policymakers, the findings suggest that legal definitions of manipulative design should cover more than just fraud and deception, since current laws do not fully address the subtle, AI-driven harms identified here. For platform designers and businesses, the proposed "Fair Patterns" principle offers guidance for creating interfaces that respect user autonomy. For researchers, the list of AI-enabled dark patterns in this study can be a starting point for future research on how common these practices are and how they affect consumers in different markets, extending lines of consumer-behaviour inquiry pursued in AIJES, such as Kaur's work on how digital and influencer marketing shape the decisions of younger consumers (Kaur, 2023, 2025). Long-term studies that track how AI dark patterns change as regulations evolve would help show whether new laws are effective or if platforms change their strategies. Finally, developing consumer-facing detection tools, based on the automated crawling methods recommended in this study, could help close the gap between regulatory goals and real-world enforcement.

REFERENCES

1. Balkin, J. M. (2016). Information fiduciaries and the First Amendment. *UC Davis Law Review*, 49(4), 1183–1234. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2675270
2. Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., & Lenzini, G. (2021). "I am definitely manipulated, even when I am aware of it. It's ridiculous!": Dark patterns from the end-user perspective. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (pp. 763–776). <https://doi.org/10.1145/3461778.3462086>
3. Brignull, H. (2010). *Deceptive design* (formerly darkpatterns.org). <https://www.deceptive.design/>
4. California Legislature. (2022). *Automatic Renewal and Continuous Service Law*. <https://leginfo.ca.gov/>
5. California Privacy Protection Agency. (2024). *California Consumer Privacy Act (CCPA) regulations*. <https://oag.ca.gov/privacy/ccpa>

6. Colorado General Assembly. (2023). *Colorado Privacy Act (CPA)*. <https://coag.gov/resources/colorado-privacy-act/>
7. Custom Communications, Inc. v. Federal Trade Commission, No. 24–3137, 2025 WL 1873489 (8th Cir. July 8, 2025).
8. Digital Markets, Competition and Consumers Act 2024, c. 13 (UK).
9. <https://www.legislation.gov.uk/ukpga/2024/13/contents>
10. Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive'). *Official Journal of the European Union*, L 149, 22–39. <https://eur-lex.europa.eu/eli/dir/2005/29/oj>
11. Dubé, J.-P., & Misra, S. (2023). Personalised pricing and consumer welfare. *Journal of Political Economy*, 131(1), 131–189. <https://doi.org/10.1086/720793>
12. European Commission. (2025a, July 28). *Press release: Commission preliminarily finds Temu in breach of the Digital Services Act in relation to illegal products on its platform*. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_1913
13. European Commission. (2025b, December 5). EU Commission fines X €120 million under the Digital Services Act [Press release]. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2934
14. European Parliament. (2025). *Digital Fairness Act*. Legislative Train Schedule. <https://www.europarl.europa.eu/legislative-train/theme-protecting-our-democracy-upholding-our-values/file-digital-fairness-act>
15. Federal Trade Commission. (2024). Trade regulation rule on the use of consumer reviews and testimonials (Final Rule). 16 C.F.R. Part 465, 89 Fed. Reg. 67,528. <https://www.federalregister.gov/documents/2024/08/22/2024-18519/trade-regulation-rule-on-the-use-of-consumer-reviews-and-testimonials>
16. Federal Trade Commission. (2025a, January). *Surveillance pricing: An examination of the surveillance pricing industry*. <https://www.ftc.gov/news-events/news/press-releases/2025/01/ftc-surveillance-pricing-study-indicates-wide-range-personal-data-used-set-individualised-consumer>
17. Federal Trade Commission. (2025b). Amazon.com, Inc. (ROSCA), FTC v. (Case No. 2:23–cv–00932). <https://www.ftc.gov/legal-library/browse/cases-proceedings/2123050-amazoncom-inc-rosca-ftc-v>
18. Frankel, T. (2011). *Fiduciary law*. Oxford University Press.
19. Georgian Competition and Consumer Agency (GCCA). (2025a). *GCCA identifies misleading and manipulative commercial practices*.
20. https://gcca.gov.ge/index.php?m=372&news_id=1731&lng=eng
21. Georgian Competition and Consumer Agency (GCCA). (2025b). *GCCA develops an E-Commerce guideline*. Retrieved from https://gcca.gov.ge/index.php?m=370&news_id=1843&lng=eng
22. Gray, C. M., Chen, J., Chivukula, S. S., & Qu, L. (2021). End user accounts of dark patterns as felt manipulation. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), Article 372, 1–25. <https://doi.org/10.1145/3479516>
23. Gunawan, J., Hartzog, W., Richards, N., Choffnes, D., & Wilson, C. (2025). Dark patterns as disloyal design. *Indiana Law Journal*, 100(4), 1389–1429.
24. <https://www.repository.law.indiana.edu/ilj/vol100/iss4/3/>

*DARK PATTERNS IN THE AGE OF ARTIFICIAL INTELLIGENCE: HOW AI MISLEADS
CONSUMERS IN GLOBAL E-COMMERCE*

25. Guo, L., Yuan, C., Zhong, M., Wolfe, R., Zhong, R., Xu, Y., Wen, B., Shen, H., Wang, L. L., & Hiniker, A. (2026). SusBench: An online benchmark for evaluating dark pattern susceptibility of computer-use agents. In *Proceedings of the 31st International Conference on Intelligent User Interfaces (IUI '26)*. ACM. <https://doi.org/10.1145/3742413.3789111>
26. Hepş, R. A., & Vidican, R. D. (2025). Unfair terms in consumer contracts. *Agora International Journal of Juridical Sciences (AIJS)*, 19(1), 47–59. <https://univagora.ro/jour/index.php/aijs/article/view/7172>
27. International Consumer Protection and Enforcement Network (ICPEN). (2024). *Dark Patterns in Subscription Services Sweep Public Report 2024*. <https://www.icpen.org/sites/default/files/2024-07/Public%20Report%20ICPEN%20Dark%20Patterns%20Sweep.pdf>
28. Kaur, K. (2023). The influence of digital and social media marketing on consumer behaviour. *Agora International Journal of Economical Sciences (AIJES)*, 17(1), 31–38. <https://doi.org/10.15837/aijes.v17i1.5760>
29. Kaur, K. (2025). The impact of influencer marketing on the consumer behaviour of Generation Z and Millennials. *Agora International Journal of Economical Sciences (AIJES)*, 19(2), 181–193. <https://univagora.ro/jour/index.php/aijes/article/view/7321>
30. Kran, E., Nguyen, J., Kundu, A., Jawhar, S., Park, J., & Jurewicz, M. M. (2025). DarkBench: Benchmarking dark patterns in large language models. In *Proceedings of the Thirteenth International Conference on Learning Representations (ICLR 2025)*. <https://arxiv.org/abs/2503.10728>
31. Law of Georgia on the Protection of Consumer Rights. (2022). (No. 1455–VIIIობ–Xოჰ). <https://matsne.gov.ge/en/document/view/5420598?publication=0>
32. Leiser, M., & Santos, C. (2023). *Dark Patterns, Enforcement, and the emerging Digital Design Acquis: Manipulation beneath the Interface*. *European Journal of Law and Technology*, 15(1). <https://ejlt.org/index.php/ejlt/article/view/990>
33. Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis*, 13(1), 43–109. <https://doi.org/10.1093/jla/laaa006>
34. Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Article 81, 1–32. <https://doi.org/10.1145/3359183>
35. New York State Senate. (2019). Senate Bill S1475A: *Prohibits certain practices by businesses making an automatic renewal or continuous service offer to consumers in the state*. <https://www.nysenate.gov/legislation/bills/2019/S1475>
36. New York State Senate. (2025). Senate Bill S8623: *Prohibits the use of algorithmically set prices and requires the disclosure of automated pricing systems*. <https://www.nysenate.gov/legislation/bills/2025/S8623>
37. <https://www.nysenate.gov/legislation/bills/2025/S8623>
38. Organisation for Economic Co-operation and Development. (2022). *Dark commercial patterns* (OECD Digital Economy Papers No. 336). OECD Publishing. <https://doi.org/10.1787/44f5e846-en>
39. Qiu, L., Huang, Y., Singh, P. V., & Srinivasan, K. (2025). Personalization, consumer search, and algorithmic pricing. *Marketing Science*, 44(6), 1278–1298. <https://doi.org/10.1287/mksc.2023.0455>
40. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4132555
41. Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on Contestable and Fair Markets in the Digital Sector and Amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act). *Official Journal of the European Union*, L 265, 1–66. <http://data.europa.eu/eli/reg/2022/1925/oj>

42. 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 (Digital Services Act). *Official Journal of the European Union*, L 277, 1–102. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>
43. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). *Official Journal of the European Union*, L 2024/1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
44. Sunstein, C. R. (2021). *Sludge: What stops us from getting things done and what to do about it*. MIT Press. <https://mitpress.mit.edu/9780262045780/sludge/>
45. Urziceanu, R.-M., & Paşcalău, V.-S. (2019). Digital marketing regulations. *Agora International Journal of Juridical Sciences (AIJJS)*, 13(1), 25–30. <https://doi.org/10.15837/aijjs.v13i1.3729>
46. Waldman, A. E. (2020). Cognitive biases, dark patterns, and the 'privacy paradox'. *Current Opinion in Psychology*, 31, 105–109. <https://doi.org/10.1016/j.copsyc.2019.08.025>
47. Zhao, Y., Tang, S., Zhang, H., & Lyu, L. (2025). AI vs. human: A large-scale analysis of AI-generated fake reviews, human-generated fake reviews and authentic reviews. *Journal of Retailing and Consumer Services*, 87, Article 104400. <https://doi.org/10.1016/j.jretconser.2025.104400>