Addressing Deceptive Designs with a Tech-Assisted Citizen Science Approach

Carlos Böhm
Technical University of Darmstadt

Helen Bader
Technical University of Darmstadt

Ephriam Zimmer
Technical University of Darmstadt

Nina Gerber
Technical University of Darmstadt

Abstract

Deceptive design patterns are widespread in digital services, even though they are prohibited by EU regulations such as the Digital Services Act. Users frequently interact with these deceptive design patterns unconsciously, without realizing that they are being manipulated. Since deceptive design patterns constantly evolve and compete with current case law and regulatory developments, recognizing, measuring, and documenting them remains challenging. Our AIDA project aims to raise users awareness of deceptive design patterns, educate them on how to deal with such practices, and monitor their occurrence across various contexts and over time. To achieve this, we combine automated technical methods for detection and documentation with a citizen science approach. Citizen science, which trains and involves citizens in research, is already being used successfully in many research areas. In this poster, we present the project and report early insights from a study with adolescents (n = 20) investigating motivation and forms of participation in a tech-assisted citizen science project.

1 Introduction

Deceptive design patterns manipulate users to make unintended decisions, often by exploiting cognitive biases and limiting their digital sovereignty. Users of digital services encounter these patterns in various contexts, such as shopping websites, cookie banners, social media platforms, and games. [4, 13] define deceptive design patterns as "user interface design choices that benefit an online service by coercing, steering, or deceiving users into making unintended and potentially harmful decisions." Such manipulation undermines users ability to interact autonomously with digital services and curtails their digital sovereignty. These practices are particularly concerning, as they frequently induce users to disclose personal data, which jeopardizing their privacy [11, 18].

Legislators have recognized the problem of deceptive design patterns and introduced legal measures to address it, such as the Digital Services Act, which prohibits these practices within the EU. However, these laws can only be effective if deceptive design patterns are recognized and their prohibition is enforced. Due to their constantly evolving and often subtle nature, deceptive design patterns are difficult to identify for both humans and automated systems. [5, 6, 14].

Research has begun to explore how users can become more aware of deceptive design patterns. Studies show that many users have difficulty recognizing deceptive design patterns in digital interfaces, but their ability to identify them improves when they are informed about such practices or supported by visual countermeasures, such as highlighting deceptive elements and explaining their risks [3, 15].

Other projects have explored the automatic detection of deceptive design patterns. For example, prior work has proposed crawlers, browser plug-ins, and CSS-based analyses to identify deceptive design patterns on websites and in cookie banners [7, 12]. Other approaches have combined automatic detection with user reporting and warning mechanisms [2]. Together, these projects have established technical detection methods and reporting systems based on screenshots or descriptions of deceptive design patterns. However, because deceptive design patterns are highly diverse and continuously evolving, no single intervention can fully address the problem [1].

To tackle this challenge, our project combines technical detection methods, visual countermeasures, and user education within a citizen science approach. Citizen science enables people without formal scientific training to actively contribute to research and has already proven successful in fields such as astronomy and environmental protection [10, 17]. Building

USENIX Symposium on Usable Privacy and Security (SOUPS) 2024.
August 11–13, 2024, Philadelphia, PA, United States.

on these successes, we apply citizen science to privacy and security research to support the large-scale detection, reporting, and monitoring of deceptive design patterns across different online contexts and over time. **Research Goal:**

- Development of a support tool for the automatic identification of deceptive design patterns.
- A holistic and evaluated citizen science education program tailored to various target groups.
- Comprehensive documentation of deceptive design patterns across contexts and over time, including trends and legal developments.
- An expanded taxonomy that incorporates newly discovered contexts and deceptive design patterns.

2 Study: Focus group with adolescents

In this ongoing exploratory study, we conducted focus groups with adolescents (n = 20), aged 14–18, in Germany between May and June 2026. Adolescents were selected because of their vulnerability to deceptive design patterns [16] and their potential to benefit from citizen science [8, 9]. The adolescents were recruited through contacts with a local digital media education center that had not previously addressed deceptive design patterns. The focus groups lasted approximately 90 minutes and ranged in size from six to seven participants. Participants received €10 as compensation. The sessions included a short demographic questionnaire and an introduction to the topics of deceptive design patterns and citizen science. We explore motivations and barriers to participation in such a project. The group work focused on designing a citizen science project on deceptive design patterns that would appeal to adolescents, as well as identifying participants requirements for a supporting tool. Each focus group was audio recorded with the participants consent and transcribed for qualitative analysis. This study was approved by our Ethics Committee. The responses collected to date were analyzed using reflexive thematic analysis according to Braun and Clarke [19].

3 Preliminary Findings

3.1 Motivations

Above all, adolescents were primarily motivated to learn about such practices and develop skills for later life. They find it helpful to carry out these projects alongside people they know, provided that participation is voluntary and not overly time consuming. Furthermore, they are motivated by the prospect of earning rewards or recognition, as P16 described it "When a project like this uncovers something, and it ends up blowing up in a really massive way", as well as by the activity having a distinctly personal, offline character that allows for face-to-face interaction.

The main barrier to participation would be a time-intensive mandatory project, especially if its tasks were unknown in advance. Furthermore, they would not participate in any project unless they were confident that their data would be protected. Additionally, it was important to them to be able to control who participates, thereby ensuring that no unpleasant individuals could join. Finally, participants wanted to ensure their participation would not lead to negative reactions. Regarding this, P15, for example, said “I think people probably wouldn’t participate because they’re afraid of being laughed at.”

3.2 Project Ideas

In their project ideas, the participants focused primarily on raising awareness of the issue through various channels, such as TikTok videos, classroom presentations, and writing songs on the topic to be performed by celebrities to reach a wider audience. Others opted for a journalistic approach or sought to generate greater awareness through open dialog and mutual exchange. A final approach proposed engaging young people in hobbies as a means of encouraging them to get away from their smartphones. P05 described it this way:“We want to educate people about this, specifically how to completely disconnect from their phones or the internet.

4 Discussion

The results suggest that adolescents prefer participatory approaches to learning about deceptive design patterns, including peer collaboration, discussions, and creative awareness activities. This complements previous citizen science research emphasizing learning opportunities and skill development as important motivations for youth participation [8, 9]. Participants also emphasized the importance of voluntary participation, privacy protection, and a safe learning environment. Future work should build on these findings by developing and evaluating educational interventions on deceptive design patterns, such as challenges on social media, and involving adolescents in the creation of learning materials. This study will be extended to include adults and older adults to examine age-related differences in educational needs and learning preferences.

Acknowledgments

This research work has been funded by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

References

[1] K. Bongard-Blanchy, A. Rossi, S. Rivas, S. Doublet, V. Koenig, and G. Lenzini. 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 Designing Interactive Systems Conference 2021*, pages 763–776, Virtual Event, USA, 2021. ACM.

[2] A. Curley, D. O'Sullivan, D. Gordon, B. Tierney, and I. Stavrakakis. The design of a framework for the detection of web-based dark patterns. In *Proceedings of the 15th International Conference on Digital Society (ICDS 2021)*, 2021.

[3] L. Di Geronimo, L. Braz, E. Fregnan, F. Palomba, and A. Bacchelli. 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*, pages 1–14, Honolulu, HI, USA, 2020. ACM.

[4] C. M. Gray, L. Sanchez Chamorro, I. Obi, and J.-N. Duane. Mapping the landscape of dark patterns scholarship: A systematic literature review. In *Designing Interactive Systems Conference*, pages 188–193, Pittsburgh, PA, USA, 2023. ACM.

[5] C. M. Gray, C. T. Santos, N. Bielova, and T. Mildner. An ontology of dark patterns knowledge: Foundations, definitions, and a pathway for shared knowledge-building. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–22, 2024.

[6] J. Gunawan, A. Pradeep, D. Chones, W. Hartzog, and C. Wilson. A comparative study of dark patterns across web and mobile modalities. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–29, 2021.

[7] P. Hausner and M. Gertz. Dark patterns in the interaction with cookie banners. *arXiv preprint arXiv:2103.14956*, 2021.

[8] Christothea Herodotou, Maria Aristeidou, Grant Miller, Heidi Ballard, and Lucy Robinson. What do we know about young volunteers? an exploratory study of participation in zooniverse. *Citizen Science: Theory and Practice*, 5(1):1–14, 2020.

[9] Christothea Herodotou, Nashwa Ismail, Maria Aristeidou, Grant Miller, Ana I. Benavides Lahnstein, Maryam Ghadiri Khanaposhtani, Lucy D. Robinson, and Heidi L. Ballard. Online community and citizen science supports environmental science learning by young people. *Computers & Education*, 184:104515, 2022.

[10] C. J. Lintott, K. Schawinski, A. Slosar, K. Land, S. Bamford, D. Thomas, M. J. Raddick, R. C. Nichol, A. Szalay, D. Andreescu, et al. Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 389(3):1179–1189, 2008.

[11] J. Luguri and L. J. Strahilevitz. Shining a light on dark patterns. *Journal of Legal Analysis*, 13(1):43–109, 2021.

[12] A. Mathur, G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty, and A. Narayanan. Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–32, 2019.

[13] A. Mathur, M. Kshirsagar, and J. Mayer. 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*, pages 1–18, Yokohama, Japan, 2021. ACM.

[14] I. Obi, C. M. Gray, S. S. Chivukula, J.-N. Duane, J. Johns, M. Will, Z. Li, and T. Carlock. Let's talk about socio-technical angst: Tracing the history and evolution of dark patterns on twitter from 2010–2021. *arXiv preprint arXiv:2207.10563*, 2022.

[15] R. Schäfer, P. M. Preuscho, R. Röpke, S. Sahabi, and J. Borchers. Fighting malicious designs: Towards visual countermeasures against dark patterns. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–13, Honolulu, HI, USA, 2024. ACM.

[16] René Schäfer, Sarah Sahabi, Annabell Brocker, and Jan Borchers. Growing up with dark patterns: How children perceive malicious user interface designs. In *Proceedings of the 13th Nordic Conference on Human-Computer Interaction (NordiCHI '24)*, pages 1–17, New York, NY, USA, 2024. Association for Computing Machinery.

[17] B. L. Sullivan, C. L. Wood, M. J. Ili, R. E. Bonney, D. Fink, and S. Kelling. ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10):2282–2292, 2009.

[18] C. Utz, M. Degeling, S. Fahl, F. Schaub, and T. Holz. (Un)informed Consent: Studying GDPR Consent Notices in the Field. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 973–990, London, United Kingdom, 2019. ACM.

[19] Braun Virginia and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.