

1-2023

Documenting Privacy Dark Patterns: How Social Networking Sites Influence Users' Privacy Choices

Dominique Kelly
Western University, dkelly48@uwo.ca

Jacquelyn Burkell
Western University, jburkell@uwo.ca

Follow this and additional works at: <https://ir.lib.uwo.ca/fimspub>



Part of the [Communication Technology and New Media Commons](#), [Library and Information Science Commons](#), and the [Social Media Commons](#)

Citation of this paper:

Kelly, Dominique and Burkell, Jacquelyn, "Documenting Privacy Dark Patterns: How Social Networking Sites Influence Users' Privacy Choices" (2023). *FIMS Publications*. 376.
<https://ir.lib.uwo.ca/fimspub/376>

Documenting Privacy Dark Patterns: How Social Networking Sites Influence Users' Privacy Choices

Dominique Kelly and Jacquelyn Burkell

Faculty of Information and Media Studies, Western University
dkelly48@uwo.ca, jburkell@uwo.ca

Introduction

In 2010, user experience practitioner Harry Brignull coined the term *dark patterns* to describe user interface (UI) design strategies intended to ‘trick’ users into taking actions that benefit online services. Since then, research into dark patterns has grown steadily, with several studies focusing on so-called *privacy dark patterns* that steer users toward choices that reduce their online privacy. For example, a website or app might set defaults that enable widespread data sharing (Bösch et al., 2016), encourage the user to disclose their phone number in exchange for increased security (Fritsch, 2017), or complicate the process of rejecting tracking cookies in a consent pop-up (Nouwens et al., 2020). Alongside growing scholarly interest, these tactics have attracted recent attention from journalists (e.g., Lima, 2022; Morrison, 2021; Pardes, 2020) and consumer protection and regulatory bodies (e.g., Commission Nationale de l’Informatique et des Libertés, 2019; European Data Protection Board, 2022; Forbrukerrådet, 2018), prompting legislative proposals to constrain their use (Kelly, 2019; Merkel, 2021).

Privacy dark patterns hinder the ability of users to make conscious, informed decisions about the management of their personal data. When these strategies are employed by social networking sites (SNSs), they can lead users to make decisions that expose them to a number of risks and harms, including reputational damage (Ronson, 2015; Solove, 2007), cyberstalking and identity theft (Kroll & Stieglitz, 2019), and regret over posted content (Wang et al., 2011). SNSs and third parties can also take advantage of detailed user profiles assembled and accessible as a result of permissive privacy settings to influence users in ways that they might not anticipate or even be aware of. For instance, personal data can reveal person-specific vulnerabilities to be exploited through targeted and tailored advertising (Susser et al., 2019a) or personalized nudging (Christl, 2017; Yeung, 2017). Few laws specifically target the use of dark patterns, placing a burden on users to resist tactics that are both pervasive (Di Geronimo et al., 2020; Nouwens et al., 2020) and effective at influencing behaviour (Luguri & Strahilevitz, 2021; Nouwens et al., 2020).

As avid users of the internet and social media, teens are in a particularly vulnerable position. Forty-five percent of American teens ‘say they are online on a near-constant basis,’ and smartphone use is nearly ubiquitous among members of this group (Anderson & Jiang, 2018a, p. 2). Teens perceive social media platforms as key tools for ‘connecting and maintaining relationships, being creative, and learning more about the world’ and regularly post about their accomplishments, family, emotions and feelings, and dating lives (Anderson & Jiang, 2018b, p. 4). At the same time, youth tend to be naïve about commercial surveillance online (Steeves et al., 2010). Crocco et al. (2020) found that high school students ‘exhibited a surprising degree of trust’ in Facebook and Google, ‘assuming that [the companies] would do them no harm’ (p. 26). Similarly, most teens ‘do not express a high level of concern about third-party access to their data’ (Madden et al., 2013, p. 2) and only a third report that they ‘often or sometimes delete or

restrict access to things they share on social media because they are concerned it could negatively impact them later in life' (Anderson & Jiang, 2018b, p. 10).

Teens' potential lack of knowledge and concern about online privacy combined with their active use of social media leave them vulnerable to effects of dark patterns on privacy choices. Given this vulnerability, research investigating how dark patterns affect young people is surprisingly limited (e.g., Fitton & Read, 2019). Moreover, little research has examined dark patterns in the specific context of SNSs (e.g., Mildner & Savino, 2021). This work aims to address these research gaps by (1) identifying privacy dark patterns on five SNSs popular among American teens; (2) examining how these strategies are deployed in three common user procedures (registering an account, configuring account settings, and logging in and out of the account); and, (3) outlining the implications for teens' privacy on social media and the development of dark pattern countermeasures.

Literature review

Behavioural economics, nudging, and dark patterns

Since the 1950s, behavioural economists have documented numerous cognitive biases and heuristics that affect people's decision-making (e.g., Simon, 1957, 2000; Tversky & Kahneman, 1974, 1986). Thaler and Sunstein (2008) took up this body of work to demonstrate how designers can *nudge* individuals toward certain actions by changing the *choice architecture*, or the context in which they make decisions. One example of a nudge is deliberately setting an option favoured by the designer as the default. For reasons including *loss aversion* (the tendency to dislike losses more than equivalent gains) and the *status quo bias* (the tendency to remain with the status quo), people typically stick with the defaults they are given (Thaler & Sunstein, 2008). Thus, policymakers could increase organ donations by implementing an opt-out system, where the default choice is to be a donor (Johnson & Goldstein, 2003; Thaler & Sunstein, 2008).

Digital nudging refers to the deliberate use of UI design elements to influence people's behaviour in digital choice environments (Mirsch et al., 2017; Schneider et al., 2018; Weinmann et al., 2016). In the context of online privacy decision-making, one line of work has investigated how users can be nudged to protect their personal data (Acquisti et al., 2017; Kroll & Stieglitz, 2019; Warberg et al., 2019). For example, Acquisti et al. (2017) compiled design strategies supportive of users' privacy and security grounded in behavioural research, such as setting privacy-protective defaults, increasing the cost or difficulty of setting a risky configuration, and allowing users to reverse poor decisions (e.g., delete regrettable posts). On the other hand, researchers have also probed how users can be nudged to *disclose* personal data (Chang et al., 2016; Gambino et al., 2016; Sundar et al., 2020). Chang et al. (2016), for instance, found that exposure to more 'risky' or 'explicit' images on a hypothetical social media site shaped participants' perceptions of the social norms concerning information disclosure and increased the amount of information they divulged in a subsequent task.

An emerging body of research examines how UI design strategies – or *dark patterns* – can influence users to make choices or take actions that benefit online services. Like nudges, dark patterns typically operate by exploiting users' decision-making vulnerabilities (Mathur et al., 2019; Waldman et al., 2020; Susser et al., 2019a). For example, a dark pattern on an SNS could capitalize on the *framing effect* and *loss aversion* (Tversky & Kahneman, 1986) by emphasizing the perks one will forego if they choose to opt out of facial recognition technology, with eliding the potential negative consequences of opting in (Forbrukerrådet, 2018). Dark patterns are analogous to what

Thaler (2018) has termed ‘sludges’: not-so-benevolent nudges that ‘weaponize people’s mental heuristics and cognitive biases against them’ (Gunawan et al., 2021, p. 3).

In essence, dark patterns constitute a particular *kind* of nudge: one that is digital and pushes the user toward a choice that benefits the online service. And, by extension, *privacy dark patterns* can be understood as digital nudges that steer users toward choices that reduce their online privacy.

Privacy dark patterns

The invasion of users’ privacy has long been recognized as a potential end goal for dark patterns. Borrowing a term coined by Tim Jones of the Electronic Frontier Foundation, Harry Brignull included *Privacy Zuckering* in his original patterns library (darkpatterns.org) to describe UIs that trick you ‘into publicly sharing more information about yourself than you really intended to.’ Early scholarly work in this area documented techniques used to weaken users’ privacy observed ‘in the wild,’ such as unnecessarily complicated privacy settings (Bösch et al., 2016) and attempts to block users from accessing online services when TOR anonymizer technology is detected (Fritsch, 2017). Related research has identified dark patterns deployed by social media platforms to frustrate users’ attempts to delete their accounts (Kelly & Rubin, 2022; Schaffner et al., 2022) and website design choices that make it confusing or difficult for users to delete data or opt out of email communications and targeted advertising (Habib et al., 2019). More generally, researchers have documented dark patterns applicable to various contexts, including users’ privacy choices; for instance, a site might give certain options visual or interactive precedence over others (Gray et al., 2018) or deliberately increase the user’s workload (Conti & Sobiesk, 2010).

Studies have begun to establish the prevalence of privacy dark patterns in online services and their impact on user behaviour. Nouwens et al. (2020) scraped the designs of the five most commonly used consent management platforms (CMPs) on the top 10,000 websites in the UK and found that only 11.8 percent of the sites were compliant with the EU’s General Data Protection Regulation (GDPR) – meaning that they had no boxes pre-ticked, made rejection as easy as acceptance, and made consent explicit. A field experiment further revealed that removing the ‘reject all’ button from the consent pop-up’s first page (i.e., the layer describing the pop-up’s general purpose and offering consent bulk options like ‘accept all’) increased the probability of consent by 22-23 percent. Luguri and Strahilevitz (2021) experimentally determined that ‘mild’ dark patterns (e.g., defaults, false hierarchy, and confirmshaming) more than doubled the percentage of consumers who signed up for a dubious subscription service and ‘aggressive’ dark patterns (e.g., nagging, toying with the user’s emotions, and presenting trick questions) nearly quadrupled the percentage of consumers who signed up. Meanwhile, Graßl et al. (2021) found that the design strategies of obstruction (i.e., requiring multiple steps) and defaults, when deliberately used to sway users *away from* consenting to tracking cookies in consent requests, were effective at influencing users’ behaviour.

Consumer protection groups and regulatory bodies have also contributed to the privacy dark patterns literature. In 2018, the Norwegian Consumer Council (Forbrukerrådet) published a report revealing how Facebook, Google, and Windows 10 influenced users’ privacy choices in GDPR pop-ups by, for instance, obstructing the path to a privacy-friendly option or framing a privacy-invasive option in positive language. France’s National Commission on Informatics and Liberty (Commission Nationale de l’Informatique et des Libertés [CNIL]) (2019) assembled a non-exhaustive list of ‘potentially deceptive design practices’ (p. 28) that could hamper users’ attempts to protect their personal data. Most recently, the European Data Protection Board

(EDPB) (2022) presented a typology of dark patterns employed by social media platforms that ‘lead users into making unintended, unwilling and potentially harmful decisions regarding the processing of their personal data’ (p. 2) and explained how each tactic infringed on GDPR requirements.

Comparison to prior work

Empirical studies of dark patterns have been conducted in contexts that include shopping websites (Mathur et al., 2019), mobile apps (Di Geronimo et al., 2020; Gunawan et al., 2021), and online privacy notices (Nouwens et al., 2020). We contribute to the extant literature by documenting privacy dark patterns in a sample of SNSs popular among American teens in order to identify the tactics that teen users are likely to encounter during their use of social media. Methodologically, we follow the approach introduced by Di Geronimo et al. (2020), who screen-recorded user interactions with online services and then coded the recordings for the presence of dark patterns. We also adapt the approach taken by the EDPB (2022) and Gunawan et al. (2021) by examining how dark patterns manifest in specific user procedures – namely, registering an account, configuring account settings, and logging in and out.

Methods

Sample selection

Our sample consists of the five most popular SNSs among American teens in fall 2020: Snapchat, TikTok, Instagram, Twitter, and Discord (Statista, 2021). Two of the sites (Snapchat and TikTok) were accessed through a mobile device, while the remaining three (Instagram, Twitter, and Discord) were accessed through a desktop browser. The purpose of interacting with the sites through different devices was to capture both mobile and desktop UIs. The sites vary in their primary forms of content that is posted: Twitter and Discord, for example, are largely text-based, while photographs are common on Instagram and Snapchat, and videos are often shared on Snapchat and TikTok.

Data collection

We registered experimental accounts for each of the five SNSs in our sample using the same Protonmail email address. Snapchat and TikTok were downloaded to a mobile device (a Samsung Galaxy tablet), while Instagram, Twitter, and Discord were accessed through a desktop browser (Google Chrome, newly downloaded and with no extensions enabled). From March to May 2022, the first author recorded videos of her interactions with each SNS using a built-in Android screen recorder for the mobile apps and Windows 10’s screen-recording software, Windows Game Bar, for the desktop sites. Recording videos allows real-time user interactions with the sites to be captured and analyzed, and this approach has been utilized in other empirical studies on dark patterns (Di Geronimo et al., 2020; Gunawan et al., 2021; Kelly & Rubin, 2022).

For the purpose of this study, we define *privacy dark patterns* as UI design strategies intended to influence users to explicitly or implicitly make privacy-invasive choices. This definition guided our analysis and is adapted from the privacy dark patterns literature (e.g., Bösch et al., 2016; EDPB, 2022; Forbrukerrådet, 2018). Personal data refers to ‘any factual or subjective information, recorded or not, about an identifiable individual’ (Office of the Privacy Commissioner of Canada, 2020, p. 5). Brandimarte et al. (2012) distinguish between the *release* (i.e., willing disclosure) of personal data, *access* to it, and others’ *usage* of it. We consider a

choice to be privacy-friendly if it *limits* the release, access, and usage of personal data and a choice to be privacy-invasive if it *facilitates* these actions. Adapting Brandimarte et al.'s (2012) framework, Table 1 defines each action in the context of SNSs and provides examples of what constitutes a privacy-invasive or privacy-friendly choice.

Table 1

Examples of privacy-invasive and privacy-friendly choices related to the user's management (release, access, and usage) of their personal data on SNSs

Action	Definition	Privacy-invasive choice	Privacy-friendly choice
Release	Controlling how much personal data is disclosed to the site.	Entering a phone number when prompted at login.	Pressing a 'skip' button when prompted to enter a phone number at login.
Access	Controlling whether other individuals (on-site connections, other site members, or the general public) have access to the personal data released to the site.	Accepting defaults that allow one's email address to be shared publicly.	Changing defaults that allow one's email address to be shared publicly.
Usage	Controlling whether and how one's personal data is used by the site.	Consenting to personalized ads and the sale of one's personal data to third parties when prompted during account registration.	Rejecting personalized ads and the sale of one's personal data to third parties when prompted during account registration.

The following procedures were recorded to capture user-SNS interactions. A protocol was followed for each procedure to ensure consistency in user actions across the five SNSs. The first author always attempted to make the most privacy-friendly choices possible.

- **Registering an account:** The first author viewed the site's terms and conditions, filled in required fields (e.g., providing an email address and password), selected privacy-friendly options when they were made available, and logged out. When required, she also logged into her Protonmail email account and clicked a link in an email from the SNS to confirm her email address.
- **Configuring account settings:** The first author logged into her account, navigated to account settings, and reviewed all options, including those not explicitly labelled as controlling 'privacy' or 'security.' She attempted to adjust settings governing the management of personal data to be more privacy-friendly, generally by clicking toggle or radio buttons. She then logged out of her account.
- **Logging in and out of the account:** The first author logged into and out of the account three additional times on different dates. The site was accessed from a different location on one of these occasions. When options were presented to the user after login or just before logging out, the privacy-friendly option was chosen.

Data collection for each procedure was carried out on separate days. The purpose of logging into the accounts multiple times, on different dates, was to determine whether the site

presented messages or notifications to users that might not be available during their first login. For the desktop sites, the first author cleared browsing data (i.e., browsing history, cookies and other site data, and cached images and files) at the end each session. Recordings were started on each SNS's login page, before login information was entered, and were ended immediately after logging out. When an SNS required an email address confirmation, the first author ended the current recording and started a new recording on the Protonmail login page.

In total, the dataset consists of 35 video recordings that range from under one minute to nearly 20 minutes in length.

Data analysis

We analyzed the user-SNS interactions captured in the recordings, focusing on explicit and implicit decision points, for evidence of UI design strategies intended to influence users to make privacy-invasive choices. We specifically examined the use of visual and verbal UI elements (e.g., buttons, text, pop-ups, pre-selected options, and images) and the stylistic choices that accompanied them (e.g., their size, colour, contrast, and placement). Our analysis was informed by the digital nudging literature (Mirsch et al., 2017; Schneider et al., 2018; Weinmann et al., 2016), which outlines how UI design elements can guide users' choices in digital environments.

The recording files were imported into NVivo, a software program for qualitative data analysis (QSR International, 2022). We manually assigned codes to temporal segments of the recordings, based on the design strategies flagged in our initial inspection. We also noted the presence of any further privacy-invasive strategies and updated our codebook to incorporate these additions. Successive rounds of coding were performed in NVivo, and our codebook was continually refined, until no new codes emerged from the data. The final codebook contains clear inclusion criteria at the level of specific kinds of design strategies (i.e., 'privacy dark patterns') and instructions on coding procedures.

The units of analysis were the procedures (registering an account, configuring account settings, and logging in and out) for each of the five SNSs. The presence (1) or absence (0) of each privacy dark pattern within each unit of analysis was recorded.

Results and discussion

We analyzed the recording of each procedure (registering an account, configuring account settings, and logging in and out) to identify design strategies that could influence users to make privacy-invasive choices. We begin this section by examining how these strategies manifested within the three procedures studied. De-identified screenshots from our dataset are included to provide illustrative examples of the strategies. We then present a typology of privacy dark patterns that emerged from our analysis of the procedures and our identification of patterns in how sites utilized the design strategies to influence users' privacy choices. Lastly, we assess the prevalence of the privacy dark patterns within each of the three procedures.

Privacy-invasive design strategies observed in the procedures of registering an account, configuring account settings, and logging in and out

Registering an account

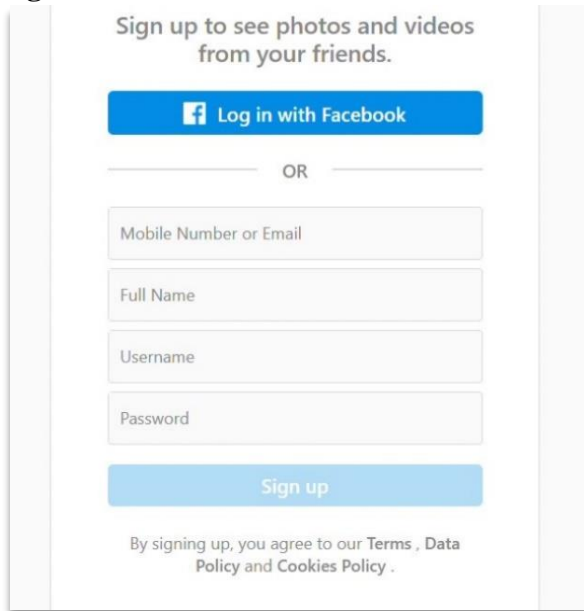
The account registration processes for all of the SNSs in our sample required the user to provide, at minimum, contact information (an email address or phone number), a password, and a username. In some cases, the user was required or given the option to input additional data, such as their full name and date of birth. Several sites presented multiple fields to be filled in but failed to distinguish between required and optional data (e.g., by failing to place asterisks beside data that was required) (see Figure 1). In the example shown in Figure 1, the ‘sign up’ button did not become clickable until the required data (phone number/email address and username) was input. The optional field (asking for the user’s full name) was situated between the required fields, with the result that the user could be given the false impression that this field as well was required in order to register an account, particularly if they entered their data in the suggested order.

Some sites also offered the user the choice of using either an email address or a phone number to sign up, but presented the phone number as the default choice, with the switch to email requiring an additional click (see Figure 2). It is generally easier to create a ‘throwaway’ or secondary email address that protects one’s privacy compared to a secondary phone number. In these situations, the user might add their phone number because it requires less effort, because they do not realize an alternative exists, or because they perceive the default to be the recommended option.

After inputting basic account data, the user was, in some instances, given the opportunity to view and opt out of certain default settings. Defaults in the account registration process included pre-ticked boxes that allowed other site members to find the user by their phone number and email address (see Figure 3) and that enabled tracking of the user’s web browsing activity for the purpose of ‘personalization.’ In the example shown in Figure 3, the site design serves to distract the user from changing defaults: a high-contrast ‘sign up’ button entices the user to complete the process, while a link leading to the pre-ticked boxes, which implement privacy-related defaults, is hidden in a block of fine print (see Figure 4). The user could easily miss this link, and therefore not review the default selections. As a result, the user could remain unaware that the account will be discoverable by their email and phone contacts.

During the account registration process, the user was also sometimes asked to make decisions regarding the management of their personal data, such as whether to accept permissions or sync contacts. The UI shown in Figure 5 demonstrates how, in one site, visual elements were used to steer users to make a privacy-invasive choice – in this case, agreeing to sync contacts. The button to consent to syncing contacts is large, colourful, and centrally located, naturally attracting the user’s attention, while the button to decline is small, faint, and hidden in the top right corner of the UI (see Figure 5).

Figure 1



Sign up to see photos and videos from your friends.

[Log in with Facebook](#)

OR

Mobile Number or Email

Full Name

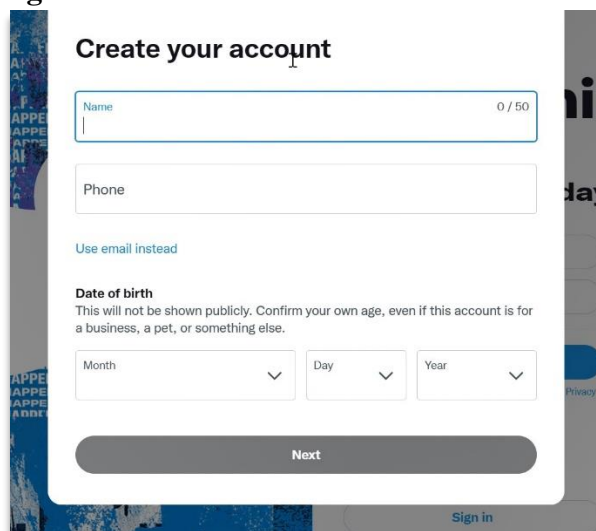
Username

Password

[Sign up](#)

By signing up, you agree to our [Terms](#), [Data Policy](#) and [Cookies Policy](#).

Figure 2



Create your account

Name 0 / 50

Phone

[Use email instead](#)

Date of birth
This will not be shown publicly. Confirm your own age, even if this account is for a business, a pet, or something else.

Month Day Year

[Next](#)

[Sign in](#)

Figure 3

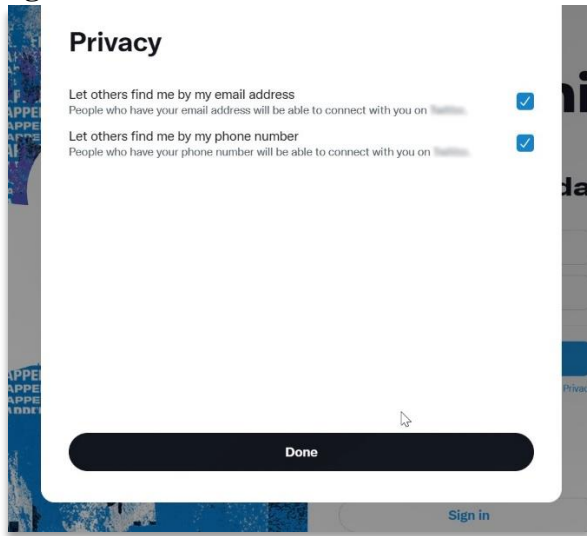


Figure 4

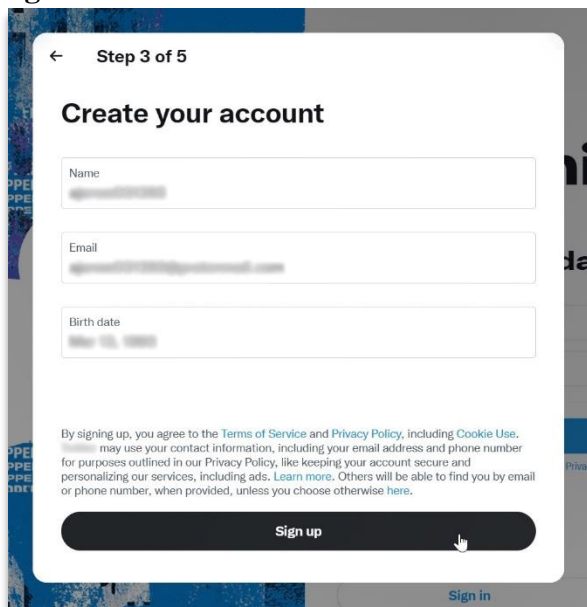
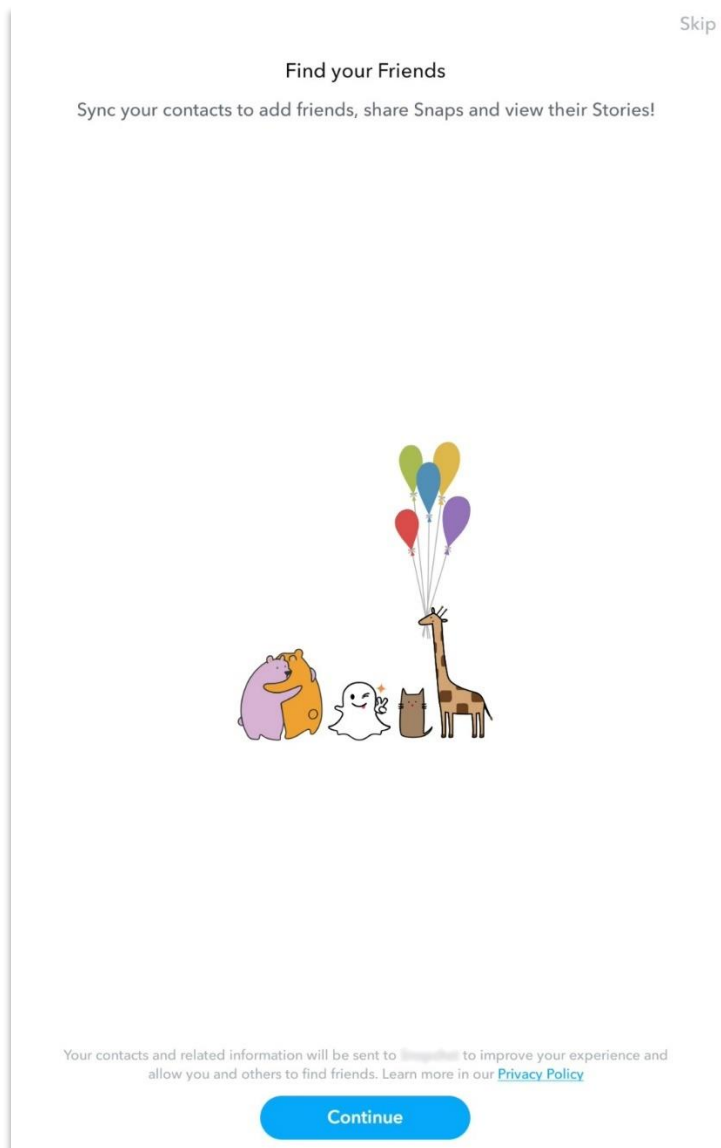


Figure 5

Configuring account settings

Instead of asking the user to indicate explicitly their preferences during account registration, sites selected certain options by default – including those governing the management of personal data. All sites in our sample set multiple privacy-invasive defaults that could be changed only through ‘account settings,’ meaning that the existence of the defaults was concealed unless the user deliberately searched for them. None of the sites suggested that the user check their account settings after registering an account to review the existing defaults and ensure that these settings aligned with their preferences. Features enabled by defaults included: marketing emails; push notifications; personalization (e.g., targeted or tailored ads); personalization based on the user’s inferred identity; the display of the user’s posts to non-connections (i.e., the general public or other site members); direct messages from non-connections; remaining discoverable to other site members by one’s phone number; and the exposure of the user’s online status and/or activities to other site members.

In some cases, the effort associated with deselecting privacy-invasive defaults was increased by requiring the user to address individually each of multiple closely-related options instead of providing a ‘reject all’ button. For example, one site set numerous defaults enabling different types of push notifications and required each to be individually deselected by clicking toggle buttons (see Figure 6), while another divided settings for personalized ads into three categories with no ‘reject all’ button available (see Figure 7). Omitting bulk options transforms what could be a one-step process to preserve privacy into a tedious, multi-step process.

Attempting to change a privacy-invasive default sometimes prompted a pop-up requiring the user to confirm their choice by clicking an additional button (see Figures 8 and 9). These confirmation pop-ups unnecessarily prolonged the opt-out process. In certain instances, the confirmation pop-ups also toyed with the user’s emotions by highlighting losses they would suffer if they proceeded, such as receiving ads ‘less relevant to [their] interests’ (see Figure 8) or losing ‘personalized recommendations and suggestions’ that could not be recovered (see Figure 9).

Some sites also gave the user the option of setting their account to ‘private’ (see Figure 10). This label was, however, misleading, since enacting the setting did not alter all relevant aspects of data management (e.g., in one case, it made the user’s posts visible only to their connections, but did not change any other privacy-invasive defaults; see Figure 11). Users who select this option might receive a false sense of privacy and leave the account settings before checking and configuring other options.

Figure 6

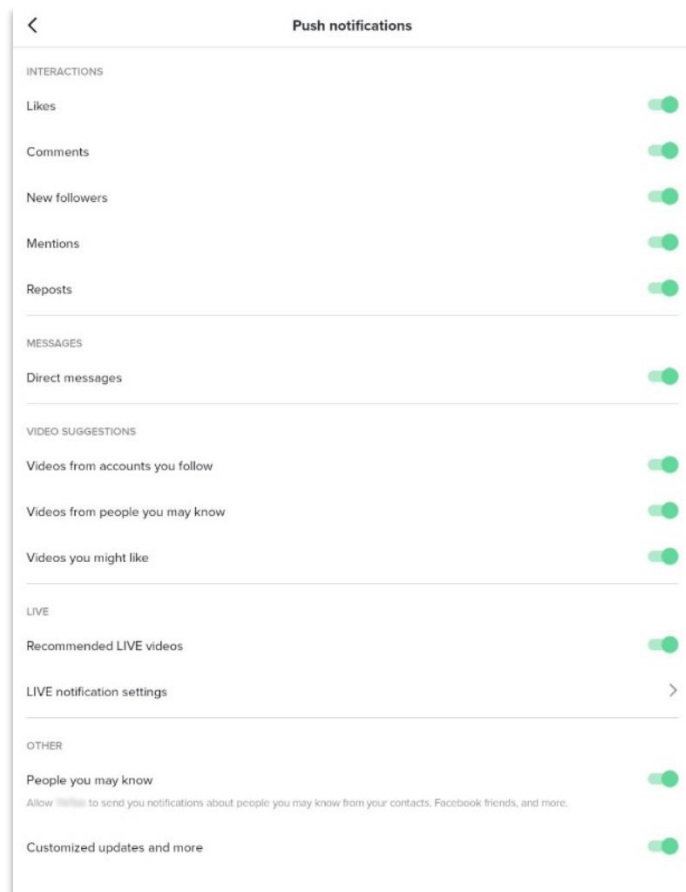


Figure 7

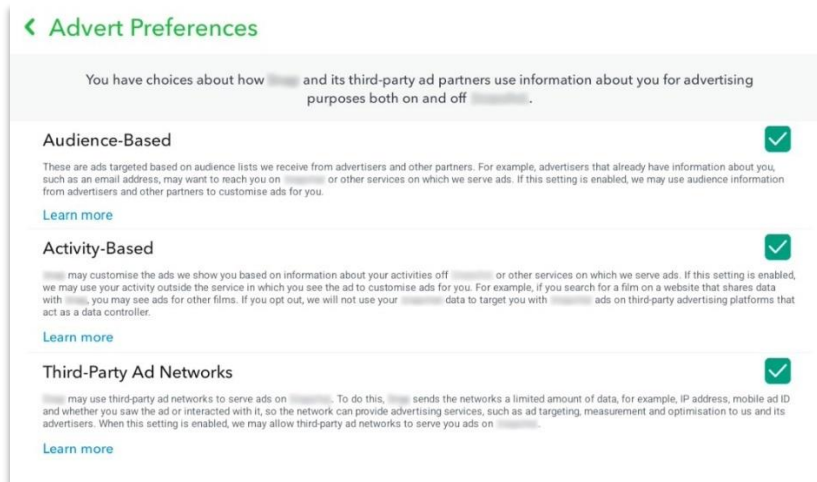


Figure 8

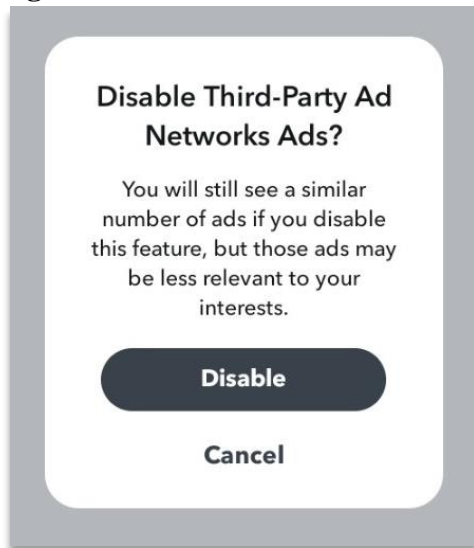


Figure 9

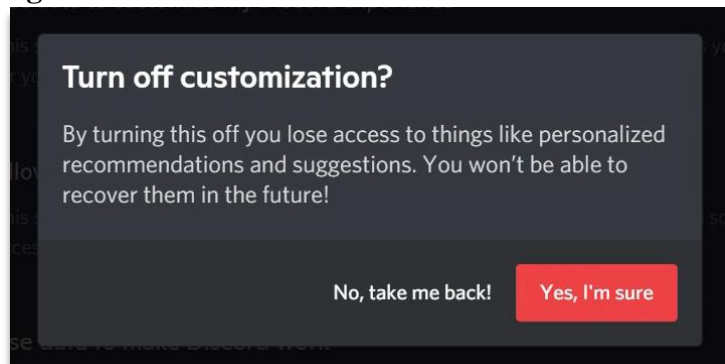


Figure 10

Account Privacy

Private Account

When your account is private, only people you approve can see your photos and videos on Instagram. Your existing followers won't be affected.

Activity Status

Show Activity Status

Allow accounts you follow and anyone you message to see when you were last active on Instagram apps. When this is turned off, you won't be able to see the activity status of other accounts.

Story Sharing

Allow Sharing

Let people share your story as messages

Comments

Figure 11

Account Privacy

Private Account

When your account is private, only people you approve can see your photos and videos on Instagram. Your existing followers won't be affected.

Activity Status

Show Activity Status

Allow accounts you follow and anyone you message to see when you were last active on Instagram apps. When this is turned off, you won't be able to see the activity status of other accounts.

Story Sharing

Allow Sharing

Let people share your story as messages

Comments

Logging in and out

All sites required the user to provide their contact information or username and their password to log into their account. At login, the user was sometimes interrupted by a pop-up that obscured the UI and needed to be manually dismissed. These pop-ups asked the user to make privacy-invasive choices, such as consenting to push notifications (see Figure 12), providing their school email address, saving their login information, or syncing their contacts from another SNS.

In some cases, the privacy-invasive choice in the interruptive pop-up was given greater salience than its privacy-friendly alternative. For example, one site made the button to turn on push notifications slightly bolder and brighter than the button to reject the invitation, drawing the user's attention toward the former button (see Figure 12). This site also phrased the choice to

reject push notifications as ‘not now’ instead of using more neutral language (see Figure 12). The user might assume (in this case, correctly) that they will continue to be pestered with requests at login, with no apparent way to permanently reject the pop-ups.

Logging out of an account and not saving login information are considered privacy-friendly in our analysis because they ensure that, if someone else uses the same device, they cannot access the user’s account. To log out of their account, the user was generally required to navigate to a main menu and click a ‘logout’ button. Sites sometimes prompted the user to confirm their choice once or even multiple times through a pop-up (see Figures 13 and 14).

Figure 12

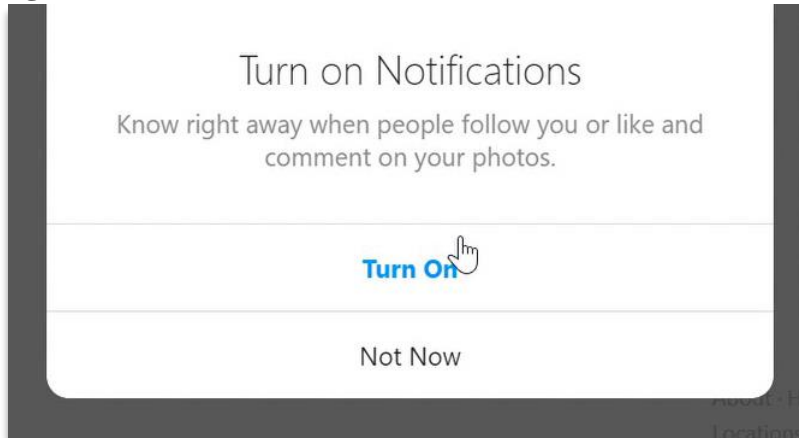


Figure 13

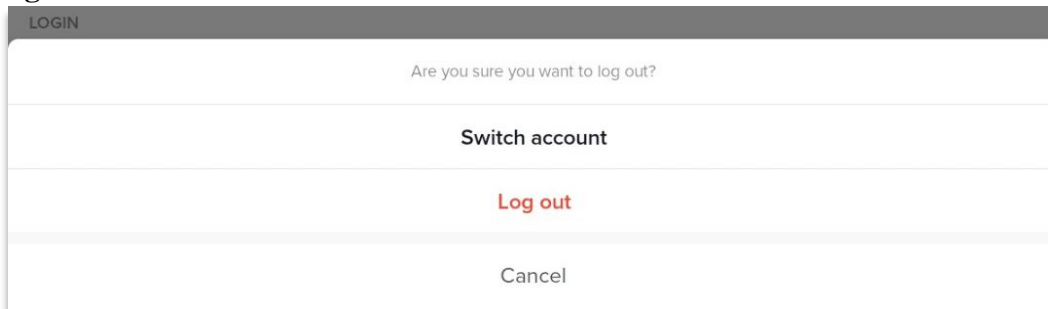
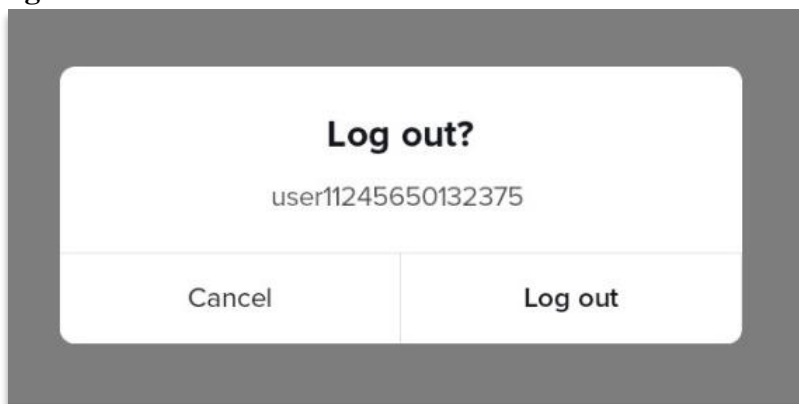


Figure 14



Typology of privacy dark patterns

In the process of analyzing the three procedures on the five SNSs, we identified patterns in how the sites utilized design strategies to influence users' privacy choices. In some cases, these strategies have been identified in previous research on dark patterns; in other cases, we identify previously undocumented dark patterns. The following typology (see Table 2) emerged from our observations. In total, 10 dark pattern subtypes were identified and thematically organized into three major types by their primary mode of influence on the user. Where applicable, we connect our dark pattern subtypes to similar dark patterns identified in prior research.

Table 2
Typology of privacy dark patterns in SNSs

Privacy dark pattern types and subtypes	Description	Similar privacy dark patterns identified in prior research
1 Obstruction	The site increases the effort that users must exert to make a privacy-friendly choice (i.e., by requiring more actions).	
<i>1.1 Defaults</i>	Privacy-invasive options are selected by default prior to user interaction, requiring the user to locate and change them.	<ul style="list-style-type: none"> • <i>Bad Defaults</i> (Bösch et al., 2016) • <i>Default Settings</i> (Forbrukerrådet, 2018) • <i>Default Sharing</i> (CNIL, 2019) • <i>Deceptive Snuggness</i> (EDPB, 2022)
<i>1.2 Confirmations</i>	Attempts to make privacy-friendly choices are accompanied by pop-ups that require the user to confirm their decision by clicking an additional button.	<ul style="list-style-type: none"> • <i>Ease</i> (Forbrukerrådet, 2018) • <i>Longer Than Necessary</i> (EDPB, 2022)
<i>1.3 Interruptions</i>	Pop-ups asking the user to make a privacy-invasive choice appear and must be manually dismissed. The requests are irrelevant to the user's current activity.	<ul style="list-style-type: none"> • <i>Repetitive Incentive</i> (CNIL, 2019) • <i>Continuous Prompting</i> (EDPB, 2022)
<i>1.4 Missing Bulk Options</i>	Three or more closely-related privacy-invasive defaults are presented together without a corresponding bulk option (e.g., a 'reject all' button).	<ul style="list-style-type: none"> • <i>Ease</i> (Forbrukerrådet, 2018) • <i>Longer Than Necessary</i> (EDPB, 2022) • <i>Privacy Zuckering</i> (Bösch et al., 2016) • <i>Obfuscating Settings</i> (CNIL, 2019)

		<ul style="list-style-type: none"> • <i>No 'Bulk' Options for Settings</i> (Gunawan et al., 2021)
2 Obfuscation	The site obscures, hides, or omits relevant information and options, with the intent of confusing or misleading the user into making a privacy-invasive choice.	
<i>2.1 Attention Manipulation</i>	The buttons for privacy-invasive choices are given greater salience than privacy-friendly choices through their size, colour, placement, and/or contrast.	<ul style="list-style-type: none"> • <i>Attention Diversion</i> (CNIL, 2019) • <i>Hidden in Plain Sight</i> (EDPB, 2022)
<i>2.2 False Requirements</i>	In a task flow, several empty fields to be filled in with the user's data appear, without any indication of which fields are required and which are optional.	
<i>2.3 False 'Private' Account</i>	The user is given the opportunity to set their account to 'private,' but enacting this setting does not alter all privacy-invasive defaults.	
<i>2.4 Concealed Settings</i>	After account registration, the site does not suggest that the user check their account settings to ensure that the current defaults align with the user's preferences.	
3 Pressure	The site actively encourages the user to make a privacy-invasive choice by presenting the privacy-invasive choice positively, and/or presenting the privacy-friendly choice negatively, through language and visuals.	
<i>3.1 Emotional Pressure</i>	The risks or costs of a privacy-friendly choice (e.g., the loss of certain features) are emphasized to evoke feelings of fear or guilt in the user.	<ul style="list-style-type: none"> • <i>Framing</i> (Forbrukerrådet, 2018) • <i>Improving the Experience</i> (CNIL, 2019) • <i>Safety Blackmail</i> (CNIL, 2019) • <i>Blaming the User</i> (CNIL, 2019) • <i>Emotional Steering</i> (EDPB, 2022)
<i>3.2 Conditional Rejections</i>	The button to reject a privacy-invasive option uses wording implying the user will be asked or required to accept the option at a later time (e.g., 'not now').	<ul style="list-style-type: none"> • <i>Continuous Prompting</i> (EDPB, 2022)

Privacy dark patterns as mutually-reinforcing strategies

To influence users' privacy choices, sites often deployed multiple dark patterns that complemented and reinforced one another. *Defaults*, for example, were frequently used in conjunction with other dark patterns to increase the likelihood that users would stick with the sites' preselected options. *Concealed Settings* make it easy for users to remain unaware, after creating an account, that various default settings are in place and can be altered. Even if users do enter their account settings and attempt to change privacy-invasive defaults, *Missing Bulk Options* and *Confirmations* prolong and complicate the process, while *Emotional Pressure* urges users to reconsider their choice at the last moment. One of the sites in our sample utilized all five strategies described above to deter users from opting out of privacy-invasive options in their account settings.

Prevalence of privacy dark patterns in the procedures of registering an account, configuring account settings, and logging in and out

The following table (see Table 3) displays the number of sites, out of our sample of five, that used each dark pattern subtype within each procedure. Notably, no single subtype stands out for its prevalence across the five sites in account registration. On the other hand, privacy-invasive *Defaults* were observed in the account settings for all five sites. Three sites further required users to confirm their choice after attempting to change at least one of those defaults (*Confirmations*), while three sites failed to offer a bulk 'reject all' option when multiple closely-related defaults were grouped together (*Missing Bulk Options*). *Concealed Settings* were observed in all five sites, as none reminded the user that they could check and configure their default settings after they had registered an account. *Confirmations* at logout were also prevalent, with four sites requiring the user to confirm their choice after clicking a 'logout' button; meanwhile, *Interruptions* at login occurred in three sites. *Attention Manipulation* was observed in pop-ups in four sites during attempts to log in or out of the user account.

Table 3

Number of sites (maximum 5) that used each privacy dark pattern subtype across the procedures of registering an account, configuring account settings, and logging in and out

Privacy dark pattern subtype	Registering an account	Configuring account settings	Logging in and out
<i>Defaults</i>	2	5	0
<i>Confirmations</i>	0	3	4
<i>Interruptions</i>	0	0	3
<i>Missing Bulk Options</i>	0	3	0
<i>Attention Manipulation</i>	2	0	4
<i>False Requirements</i>	2	0	0
<i>False "Private" Account</i>	0	2	0
<i>Concealed Settings</i>	0	5	0
<i>Emotional Pressure</i>	0	2	0
<i>Conditional Rejections</i>	0	0	2

Limitations

Our study has some limitations. Additional privacy dark pattern types and subtypes might have been encountered if: (1) the user account had been retained for a longer period of time; (2) additional user procedures were analyzed (e.g., editing the user's profile or browsing site content); (3) both the mobile and desktop UIs for each site were studied; and, (4) a larger sample size was chosen. Future studies could address these limitations and/or use our existing typology to assess the prevalence of privacy dark patterns in a larger sample of SNSs or online services. Researchers could also experimentally determine how certain privacy dark patterns impact user behaviour.

In prior work, *Forced Action* has been used to describe cases where users are required to undertake an action (Gray et al., 2018), such as immediately consenting to new terms to regain access to their account (Forbrukerrådet, 2018). While undoubtedly problematic, these tactics are beyond the scope of this study: we limited our analysis to those instances where users are afforded *choices* in the management of their personal data – and where opportunities might therefore arise for resistance.

We cannot claim to know the intentions of the designers who created the SNS UIs with total certainty. For instance, failing to distinguish between mandatory and optional fields (as in *False Requirements*) could be an instance of poor design, or an 'anti-pattern' (Gray et al., 2018), rather than an intentional attempt by the designer to confuse or mislead the user. Given that collected data have clear economic value (Acquisti et al., 2016), however, it seems plausible that design elements that nudge users toward privacy-invasive options are deliberate.

Conclusion

In this study, we examined the UI design strategies employed by SNSs to influence users to make privacy-invasive choices ('privacy dark patterns'). We content-analyzed recordings of three common user procedures (account registration, configuring account settings, and logging in and out) on five SNSs popular among teens and identified three major types of privacy dark patterns (Obstruction, Obfuscation, and Pressure) and 10 subtypes.

The naïveté that teens sometimes demonstrate regarding online privacy and surveillance (e.g., Crocco et al., 2020) might render them especially susceptible to the influence of dark patterns on privacy decision-making which will in turn expose these young users to a variety of risks and harms. For example, nudging users to select or maintain options that make much of their personal data publicly visible could invite reputational damage (Ronson, 2015; Solove, 2007), cyberstalking and identity theft (Kroll & Stieglitz, 2019), and eventual user regret over posted content (Wang et al., 2011). Even when the data are restricted from public view, disclosure of personal information increases the ability of sites to infer additional, sensitive information about users and to detect person-specific vulnerabilities (Susser et al., 2019b). Combined with settings that enable personalization of the user's experience (e.g., by delivering targeted and tailored ads), data-rich profiles could allow sites to influence users' behaviour, potentially in ways outside of the user's conscious awareness. This possibility is well-demonstrated in the 2017 report of a leaked internal Facebook document that allegedly detailed how advertisers could target ads at teens during moments when they feel insecure (Susser et al., 2019a).

To protect teens from privacy risks and harms on social media, there is a clear need for the development of dark pattern countermeasures. One approach to train users to recognize and

respond to those dark patterns that they encounter online (Bösch et al., 2016; Fritsch, 2017; Rossi & Bongard-Blanchy, 2021). Efforts in this area are closely related to the field of digital literacy, a subset of media literacy that focuses on teaching skills relevant to the use of digital technologies (Common Sense Media, 2020). Insights from our study could inform the development of digital literacy materials that show teens how to resist the privacy dark patterns frequently deployed in social media.

Nevertheless, user training only constitutes one part of an effective solution. Countermeasures that target the source of the problem – that is, the design choices deliberately made by social media companies and other online service providers – are critical and have begun to emerge in recent years. In March 2021, for example, amendments to the California Consumer Protection Act (CCPA) ‘banning the use of dark patterns to subvert or impair the process for consumers to optout of the sale of personal information’ were approved (Merkel, 2021), and in March 2022, the EDPB published guidelines for the design of social media interfaces that are free of dark patterns and therefore compliant with the GDPR. Further research could determine whether certain dark patterns – or combinations of mutually-reinforcing dark patterns – identified in our study also warrant regulation.

References

- Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economics Literature*, 54(2), 442-492.
<https://doi.org/http://dx.doi.org/10.1257/jel.54.2.442>
- Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., & Wilson, S. (2017). Nudges for privacy and security. *ACM Computing Surveys*, 50(3), 1-41.
<https://doi.org/10.1145/3054926>
- Anderson, M., & Jiang, J. (2018a, May 31). Teens, social media and technology 2018. *Pew Research*. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>
- Anderson, M., & Jiang, J. (2018b, November 28). Teens' social media habits and experiences. *Pew Research*. <https://www.pewresearch.org/internet/2018/11/28/teens-social-media-habits-and-experiences/>
- Bösch, C., Erb, B., Kargl, F., Kopp, H., & Pfattheicher, S. (2016). Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proceedings on Privacy Enhancing Technologies*, 2016(4), 237-254. <http://dx.doi.org/10.1515/popets-2016-0038>
- Brandimarte, L., Acquisti, A., & Loewenstein, G. (2012). Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*, 4(3) 340-347.
<https://doi.org/10.1177/1948550612455931>
- Brignull, H. (n.d.). Dark patterns. <https://www.darkpatterns.org/>
- Chang, D., Krupka, E. L., Adar, E., & Acquisti, A. (2016). Engineering information disclosure: Norm shaping designs. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 587-597. <http://dx.doi.org/10.1145/2858036.2858346>
- Christl, W. (2017). *How companies use personal data against people: Automated disadvantage, personalized persuasion, and the societal ramifications of the commercial use of personal information*. October. <http://crackedlabs.org/en/data-against-people>
- Commission Nationale de l'Informatique et des Libertés. (2019). *Shaping choices in the digital world*.
https://linc.cnil.fr/sites/default/files/atoms/files/cnil_ip_report_06_shaping_choices_in_the_digital_world.pdf
- Common Sense Media. (2020, June 4). *What is digital literacy?*
<https://www.commonsensemedia.org/articles/what-is-digital-literacy>
- Conti, G., & Sobiesk, E. (2010). Malicious interface design: Exploiting the user. *WWW'10: Proceedings of the 19th international conference on World wide web*, 271-280.
<https://doi.org/10.1145/1772690.1772719>
- Crocco, M. S., Segall, A., Halvorsen, A-L., Stamm, A., & Jacobsen, R. (2020). "It's not like they're selling your data to dangerous people": Internet privacy, teens, and (non-)controversial public issues. *The Journal of Social Studies Research*, 44, 21-33.
<https://doi.org/10.1016/j.jssr.2019.09.004>
- Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F., & Bacchelli, A. (2020). UI dark patterns and where to find them: A study on mobile applications and user perception. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-14.
dl.acm.org/doi/pdf/10.1145/3313831.3376600
- European Data Protection Board. (2022). *Dark patterns in social media platforms: How to*

- recognise and avoid them.* https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-32022-dark-patterns-social-media_en
- Fitton, D., & Read, J. C. (2019). Creating a framework to support the critical consideration of dark design aspects in free-to-play apps. *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, 407-418. <https://doi.org/10.1145/3311927.3323136>
- Forbrukerrådet. (2018). *Deceived by design: How tech companies use dark patterns to discourage us from exercising our rights to privacy.* fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf
- Fritsch, L. (2017). Privacy dark patterns in identity management. In L. Fritsch, H. Roßnagel, & D. Hühnlein (Eds.), *Open Identity Summit 2017: Proceedings* (pp. 93-104). <http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-63722>
- Gambino, A., Kim, J., Sundar, S. S., Ge, J., & Rosson, M. B. (2016). User disbelief in privacy paradox: Heuristics that determine disclosure. *Conference on Human Factors in Computing Systems – Proceedings*, 2837-2843. doi:10.1145/2851581.2892413
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-14. <https://doi.org/10.1145/3173574.3174108>
- Graßl, P., Schraffenberger, H., Zuiderveen Borgesius, F., & Buijzen, M. (2021). Dark and bright patterns in cookie consent requests. *Journal of Digital Social Research*, 3(1), 1-38. <https://doi.org/10.33621/jdsr.v3i1.54>
- Gunawan, J., Pradeep, A., Choffnes, D., Hartzog, W., & Wilson, C. (2021). A comparative study of dark patterns across web and mobile modalities. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW), 1-9. <https://doi.org/10.1145/3479521>
- Habib, H., Zou, Y., Jannu, A., Sridhar, N., Swoopes, C., Acquisti, A., Cranor, L. F., Sadeh, N., & Schaub, F. (2019). An empirical analysis of data deletion and opt-out choices on 150 websites. *Proceedings of the Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 387-406. <https://www.usenix.org/conference/soups2019/presentation/habib>
- Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302(5649), 1338-1339. <https://doi.org/10.1126/science.1091721>
- Kelly, D. & Rubin, V. L. (2022). Dark pattern typology: How do social networking sites deter disabling of user accounts? *12th International Conference on Social Media & Society*, July 18 -19, Toronto, Canada. <https://easychair.org/publications/preprint/GD6S>
- Kelly, M. (2019, April 9). Big Tech’s ‘dark patterns’ could be outlawed under new Senate bill. *The Verge*. <https://www.theverge.com/2019/4/9/18302199/big-tech-dark-patterns-senate-bill-detour-act-facebook-google-amazon-twitter>
- Kroll, T., & Stieglitz, S. (2019). Digital nudging and privacy: Improving decisions about self-disclosure in social networks. *Behaviour & Information Technology*, 1-19.
- Lima, C. (2022, January 5). Google is manipulating browser extensions to stifle competitors, DuckDuckGo CEO says. *The Washington Post*. <https://www.washingtonpost.com/politics/2022/01/05/google-is-manipulating-browser-extensions-stifle-competitors-duckduckgo-ceo-says/>
- 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>
- Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., Smith, A., & Beaton, M. (2013),

- May 21). Teens, social media, and privacy. *Pew Research Center and the Berkman Center for Internet and Society at Harvard University*.
<https://www.pewresearch.org/internet/2013/05/21/teens-social-media-and-privacy/>
- 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 Human-Computer Interaction*, 3(81), 1-32.
<https://doi.org/10.1145/3359183>
- Merkel, J. (2021). Dark patterns come to light in California data privacy laws. *The National Law Review*, 12(230). <https://www.natlawreview.com/article/dark-patterns-come-to-light-california-data-privacy-laws>
- Mildner, T., & Savino, G-L. (2021). Ethical user interfaces: Exploring the effects of dark patterns on Facebook. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*, 1-7.
<https://doi.org/10.1145/3411763.3451659>
- Mirsch, T., Lehrer, C., & Jung, R. (2017). Digital nudging: Altering user behavior in digital environments. *13th International Conference on Wirtschaftsinformatik*, 634-648.
- Morrison, S. (2021, April 1). Dark patterns, the tricks websites use to make you say yes, explained. *Vox*. <https://www.vox.com/recode/22351108/dark-patterns-ui-web-design-privacy>
- Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020). Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*, 1–13.
<https://doi.org/10.1145/3313831.3376321>
- Office of the Privacy Commissioner of Canada. (2020, September). *Privacy guide for businesses*.
https://www.priv.gc.ca/media/2038/guide_org_e.pdf
- Pardes, A. (2020, August 12). How Facebook and other sites manipulate your privacy choices. *The Verge*. <https://www.theverge.com/2021/3/16/22333506/california-bans-dark-patterns-opt-out-selling-data>
- QSR International. (2022). *NVivo*. <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home>
- Ronson, J. (2015). *So you've been publicly shamed*. Riverhead Books.
- Rossi, A., & Bongard-Blanchy, K. (2021). All in one stroke? Intervention spaces for dark patterns. *Conference on Human Factors in Computing System (CHI '21)*, 1-5.
- Schaffner, B., Lingareddy, N. A., & Chetty, M. (2022). Understanding account deletion and relevant dark patterns on social media. *Proceedings of the ACM Human-Computer Interaction*, 6(CSCW2), 1-43. <https://doi.org/10.1145/3555142>
- Schneider, C., Weinmann, M., & vom Brocke, J. (2018). Digital nudging: Guiding online user choices through interface design. *Communications of the ACM*, 61(7), 67-73.
- Simon, H. A. (1957). *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*. Wiley.
- Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society*, 1, 25-39.
- Solove, D. J. (2007). *The future of reputation: Gossip, rumor, and privacy on the Internet*. Yale University Press.

- Statista. (2021). Most popular social networks of teenagers in the United States from fall 2012 to fall 2020. <https://www.statista.com/statistics/250172/social-network-usage-of-us-teens-and-young-adults/>
- Steeves, V., Milford, T., & Butts, A. (2010). Summary of research on youth online privacy. *Office of the Privacy Commissioner of Canada*. https://priv.gc.ca/media/1731/yp_201003_e.pdf
- Sundar, S. S., Kim, J., Rosson, M. B., & Molina, M. D. (2020). Online privacy heuristics that predict information disclosure. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-12. <https://doi.org/10.1145/3313831.3376854>
- Susser, D., Roessler, B., & Nissenbaum, H. (2019a). Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review*, 4(1), 1-39.
- Susser, D., Roessler, B., & Nissenbaum, H. (2019b). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2), 1-22.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin Books.
- Thaler, R. (2018). Nudge, not sludge. *Science*, 361(6401), 431. <https://doi.org/10.1126/science.aau9241>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *The Journal of Business*, 59(4), S251-S278.
- 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>
- Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., & Cranor L. F. (2011). "I regretted the minute I pressed share": A qualitative study of regrets on Facebook. *Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS '11)*, 1-16. <https://doi.org/10.1145/2078827.2078841>
- Warberg, L., Acquisti, A., & Sicker, D. (2019). Can privacy nudges be tailored to individuals’ decision making and personality traits? *18th Workshop on Privacy in the Electronic Society (WPES'19)*, 175-197. <https://doi.org/10.1145/3338498.3358656>
- Weinmann, M., Schneider, C. & vom Brocke, J. (2016). Digital nudging. *Business & Information Systems Engineering*, 58(6), 433-436.
- Yeung, K. (2017). ‘Hypernudge’: Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118-136. <https://doi.org/10.1080/1369118X.2016.1186713>