4 Grand Canal Square
Dublin 2
D02 X525
Ireland

2 May 2022

Meta's comments on the European Data Protection Board draft Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them

1. Introduction

Meta welcomes the opportunity to respond to the public consultation of the European Data Protection Board (EDPB) on its draft *Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them* ("Draft Guidelines").

Dark patterns are emerging as an important issue in the context of preserving fairness, transparency, control and the integrity of the user experience, including but not limited to the data protection rights of individuals. We recognise that specific and potentially harmful effects of dark patterns are a cause of genuine concern, so it is important that this subject is addressed with evidence-based and practical guidelines that are consistent with GDPR. Over the past years, Meta has been dedicating resources to deepen understanding of this issue and we welcome the opportunity to share our expertise on this matter in the context of this consultation.

We welcome the work done by the EDPB to prepare the Draft Guidelines and note that there are many useful points that are relevant to all organisations with an online presence and that can apply to their websites, apps and other online channels. We also offer some observations on the Draft Guidelines which we hope are helpful to the EDPB as they review the responses received to the public consultation.

2. The Scope of the Draft Guidelines is Too Narrow

With regard to in-scope entities, the EDPB has chosen to focus its Draft Guidelines on “social media providers” (and, at times, “social media platforms”). Firstly, the concept of “social media platform interface” is a novel construct, is not defined, and has no established definition in EU law, creating significant confusion for data controllers.

More significantly, the focus on social media platforms in these Draft Guidelines is unjustifiably narrow and insufficient to protect data subjects from the risks of dark patterns. Specifically, the Draft Guidelines fail to address similar potential risks found in other online services that collect and process data, such as eCommerce platforms, public administration websites, search engines, media streaming services, news services, community websites, educational services. Indeed, many of these contexts have been the subject of academic investigation of dark patterns[1]. GDPR does not distinguish between different types of data controllers and processors. These guidelines should follow the same convention, in the interest of consistency, the common protection of people’s rights online, and to provide legal clarity.

Other online services are required to meet GDPR requirements and similarly engage in the same practices examined in the Draft Guidelines, including posting privacy policies, seeking consent, enabling data subject rights or allowing for account deletion. In addition, from a data subject’s perspective, there is no reason to narrowly focus on a small set of online services; transparency and control are paramount regardless of which entity collects and processes his or her data. Academic research has shown that Dark Patterns can and do occur in non-social media services. To meaningfully address risk and meet people’s expectations, regulatory guidelines on dark patterns should apply across the full range of online services.

Recommendation: The Draft Guidelines should apply broadly to all data controllers and processors, in line with existing and widely-understood concepts in the GDPR. Rather than narrowly focus on social media services, it would be helpful to identify the common set of practices which are at risk of being perceived as dark patterns across a range of domains.

3. Context is Relevant to Assessing Dark Patterns

Under GDPR, controllers are responsible and accountable for interpreting principles and provisions laid out in GDPR and implementing compliance solutions that ensure the protection of user rights and freedoms, whilst being tailored to their businesses. When doing so, controllers must take into account various factors, including the nature and subject of the processing and the potential impact on the rights and freedoms of data subjects. In addition, controllers must account for important additional context such as business imperatives and user

[1] See, for example, What Makes a Dark Pattern... Dark?, Mathur, Mayer, Kshirshsagar, Princeton University, 2021.

research. Further, when determining compliance with data protection by design and default, Article 25 GDPR specifically provides for factors such as the state of the art; costs of implementation; the nature, scope, context and purposes of processing; and the risks of varying likelihood and severity for rights and freedoms of individuals.

As the Draft Guidelines lean heavily on the principles and requirements of the GDPR, such factors should also consistently be accounted for in an assessment of dark patterns. The Draft Guidelines, however, lose sight of these factors and, specifically, the importance of context. As one example, the Draft Guidelines state that not having all data protection settings in one place can amount to the dark pattern of 'Too many options'.[2] However, there may be compelling and legitimate reasons to not centralise such settings in one place. Most online services will have many settings that technically relate to the processing of data, including those pertaining to security or content. To consolidate them all in one place might overwhelm or confuse users, or alternatively result in users never encountering such settings. In contrast, letting users discover settings in-context (*e.g.*, learning why they are seeing a particular piece of content and how they can see more or less of it), means that they are more likely to understand and meaningfully engage with settings.

Recommendation: The Draft Guidelines should reflect the context-based approach of GDPR and ensure that data controllers can weigh all relevant equities relative to context, user risk, and expectations in order to best assess and avoid the risk of dark patterns.

4. The Guidelines Should be Clear and Implementable

For the Draft Guidelines to be successfully implemented, data controllers must be able to easily understand them. However, the Draft Guidelines - despite being 64 pages and almost 30,000 words - do not meet this standard. They are presented through a list of examples, and in many instances it is unclear which of these are meant to be interpreted as requirements and which are simply recommendations. The introduction of some suggestions as “best practices” adds another layer of confusion.

Controllers need succinct, objective criteria for controllers to be able to assess their practices against dark patterns, rather than a long list of prescriptive examples. This format risks having less robustness and longevity against current and future processing activities carried out by controllers. Further, by over-relying on concrete examples at the expense of clear, broadly-applicable principles, it will be hard for controllers to apply the taxonomy to practices beyond the extremely specific examples highlighted.

[2] Draft Guidelines, Para. 118.

Recommendations: The Guidelines should be simplified to make them easy to understand and apply. Examples are helpful but the focus should be on overarching principles, supported by evidence and the relevant provisions of GDPR. To be durable, the Draft Guidelines should be applicable to different contexts and services. Most importantly, the Guidelines should make it clear when something is actually required by the GDPR (as opposed to a recommendation). The EDPB also could consider using its authority to instead publish the document as 'Best Practices' or 'Recommendations', pursuant to Article 70(1) GDPR.

5. The Draft Guidelines Should be Supported by Evidence-Based Research

To ensure the highest levels of clarity, legitimacy, and effectiveness, the Draft Guidelines should be evidence-based and research-supported. The EDPB bases many of its conclusions and findings on what it believes users will expect in terms of content and interface design, and on what choices it believes users will make. The EDPB makes a number of strong assertions about best practices in content design, information architecture, and even in approaches to device security. However, it is unclear to what extent these assumptions and conclusions are substantiated by robust research.

As one example, the Draft Guidelines imply that humour easily distracts users, and that users would not expect account deletion functionalities on the bottom page of account settings[3]. It is equally possible, however, that humour actually attracts people's attention to important information and is therefore privacy-enhancing. Unfortunately, here is no research cited to support these and many other assumptions found in the Draft Guidelines. Online services share the EDPB's objective to make privacy and data protection information accessible to users and agree that services should be designed with users' needs in mind. However, it is important that product makers have freedom to design their particular services in line with their brands and in light of their users' expectations, within constraints of fairness, honesty and clarity.

By drawing on objective user research, the Draft Guidelines can best support the core aim of providing a secure, fair, transparent, and high quality user experience. There are several areas where the Draft Guidelines fail to draw on objective research and risk creating counterproductive effects, including undermining data security and the quality and accessibility of online data controls. Of particular note:

a. Two-factor authentication should include phone authentication

The Draft Guideline states that "social media providers are required not to ask for additional data such as the phone number, when the data users have already provided during the sign-up process are sufficient. For example, to ensure account security,

[3] Draft Guidelines, Example 51.

enhanced authentication is possible without the phone number by simply sending a code to users' email accounts or by several other means"[4]. Prohibiting the use of phone numbers for authentication would significantly undermine people's online security. Authentication based on mobile phones is an accessible method for many users. In addition it is highly effective; users must physically possess the phone in order to authenticate. For these reasons, phone authentication has been acknowledged and recommended by Data Protection Authorities,[5] and Meta offers 2fac through phone numbers as well as through security keys and authentication apps.

Best practice should seek to maximise the use and application of 2fac, regardless of the method. Data controllers are of course still subject to GDPR requirements relating to securing personal data (Article 32) and data minimisation, to ensure that authentication does not come at the expense of data protection principles.

b. In-context transparency and controls are as important as centralised approaches

The Draft Guidelines describe the dark pattern "Overloading - Privacy Maze", in which: i) information on data subject rights is available on multiple pages, ii) the privacy policy informs on all rights, but does not redirect to a place where these can be exercised, and iii) the user's account does not contain information on all rights. The Draft Guidelines also stress that it should always be easy to find the overview of all rights, and that the controller should provide further information in case some of those rights are not available to users.

Although centralising information about, and access to, data subject rights is important, such an approach should not come at the expense of efforts that companies undertake to provide in-product/contextual access to information on data subject rights. In other words, a centralised approach is not exclusive from an in-product/contextual approach. In fact, such contextual efforts, rather than constituting a dark pattern, actually increase transparency and consolidate a privacy framework which gives users consistent trust and assurance that they can find relevant information and corresponding functionalities regarding their data subject rights. In-context and centralised mechanisms should be mutually supporting, driven by evidence based approaches to understanding users' expectations and needs[6].

c. Enabling data subject rights depends on user guidance and education

[4] Draft Guidelines, Para. 30.

[5] See, for example, the CNIL's blog post on 2 factor authentication.

[6] See, for example, TTC Labs research about ensuring visibility of privacy settings.

The Draft Guidelines set a low threshold for what actions would obstruct the exercise of data subject rights[7] . This approach fails to acknowledge data controllers' experience that, despite regulatory guidance having been issued, data subject rights remain a topic that is still not fully understood by data subjects. In particular, there is a common misunderstanding from data subjects that their rights are absolute in terms of scope and conditions for exercise. Also, some data subjects often formulate their requests so broadly or abstractly that they may not be well understood nor well addressed by data controllers (or indeed covered by GDPR rights).

In this context, adding certain steps and specific questions as part of enabling data subject rights can be beneficial to manage data subjects' expectations (in terms of applicability of the right they want to exercise), to ensure their request is appropriately and timely addressed, and to help educate them about their rights under GDPR. In fact, helping data subjects navigate their rights by breaking the process into easy-to-understand steps effectively contributes to the controller's GDPR obligation to *facilitate* the exercise of data subject rights, and should not be deemed a dark pattern. Efforts from controllers to continuously guide and inform data subjects about this topic should be encouraged rather than qualified as dark patterns.

d. Account deletion and deactivation

As recognized by the EDPB, the deletion of a social media account determines the deletion of personal data linked to that account.

The Draft Guidelines refer to a situation where a controller would inform its users at account deletion that "you'll lose everything forever", "you won't be able to reactivate your account", "your friends will forget you".[8] In the guidelines related to these examples, the EDPB states that emotional steering such as "threatening users that they will be left alone if they delete their accounts" infringes Article 12(2) and article 5 GDPR. However, social media providers have an obligation to inform users adequately of the (data protection) consequences of their decisions, e.g., that they will no longer be able to access their data if they delete their accounts. As such, a factual statement such as "you won't be able to reactivate your account" should not amount to a dark pattern. To be clear, this does not mean that deleting an account should be made difficult.

Relatedly, offering the option to 'pause' rather than delete the account should be seen as a safety net, empowering users to control how they engage with the service.

[7] Draft Guidelines, Paras. 147-148.

[8] Draft Guidelines, Example 52.

Additional options include avoiding messages or notifications, or not appearing visible to other users.

Users also should be able to reactivate a deactivated account with the expectation that their settings, data, and account are restored to the state they were in when the account was "paused". This is because users are expressly choosing not to remove their accounts but to render them temporarily inert, a well understood and legitimate user need that should be respected.

It is clearly in the interest of users for controllers to be transparent about the objective consequences of account deletion, especially given that people's content and network are often personally and professionally valuable to them. The EDPB should avoid undermining and discouraging data controllers' efforts in this respect. Unclear guidance could undermine meaningful, people-centric account controls, like deactivation and reactivation, or reduce transparency about the consequences of account deletion in a manner that could cause highly negative experiences such as the loss of valued photographs or connections.

Recommendations: Reformulate the relevant sections of the Draft Guidelines to ensure that guidelines that relate to security are driven by evidence and are informed by security best practices. The Draft Guidelines should be based on and point to user research when discussing user behaviour, identifying dark patterns, and recommending best practices. Particular care should be undertaken not to undermine provisions for in- context transparency and controls through a singular focus on centralised controls.

6. Meta's Approach

Meta shares the EDPB's interest in and concern about potential dark patterns. Addressing dark patterns means responding to bad practices but also developing and promoting good ones.

In 2018, Meta launched Trust Transparency and Control Labs (TTC Labs), a collaborative cross-industry privacy design lab. TTC Labs brings together global experts to apply a people- centric approach to develop and share best practices for good design for privacy, with a particular focus on trust, transparency and control of data.

To date, TTC Labs have brought together over 400 companies, 150 policy and academic organisations and 60+ design agencies globally to focus on shared privacy challenges facing industry and policy communities. These challenges include transparency and consumer awareness, data transparency for young people, informed consent and notification, and designing for AI explainability. Through the TTC Labs platform, the team publishes prototype designs, reports and design guides for people-centric privacy design.

a. Sharing privacy research

Last year, the team piloted a new program to share synthesised research conducted at Meta on topics such as users' privacy concerns, how to design effective privacy settings, how tech companies can change their product development processes to better support user privacy, and more. We publish this research to share the evidence that drives some of Meta's product practices, so that others within industry can benefit from our research insights in an effort to improve privacy practices ecosystem-wide. Several research articles relate to issues described in the Draft Guidelines. Two relevant examples:

I. In "How to make privacy settings easier to find using better names and organisation," our research team identified two important strategies that help make privacy settings easier to find: (1) Present privacy settings in short lists that are grouped based on users' mental models for privacy topics; and (2) Use descriptive names for privacy settings that avoid the generic word "privacy".

II. In "Consumer Journeys in Becoming Privacy Conscious", privacy researcher Jim Hudson, Ph.D. interviewed consumers who had recently attempted to make conscious changes to their messaging behaviour in response to a perceived privacy threat, to better understand how consumers respond to privacy-threatening situations. The Privacy Beliefs and Judgments (PB&J) framework is a new approach to measuring privacy concerns that can be adapted to measure privacy concerns for a variety of topics and products.

b. A research-led approach to identifying dark patterns concerns

Driven by an evidence-based approach, the TTC Labs team at Meta have been working with internal product, legal, the DPO office, policy, and design teams to research the concept of dark patterns, explore practical approaches to addressing them, and develop a working framework to help inform the company's approach.

The concept of "dark patterns" can cover a broad range of areas and disciplines, so the team first decided on a scope that would be broad enough to be meaningful but narrow enough to be useful. This work focused on privacy dark patterns that encourage or promote unnecessary data collection and processing. It also focused on product practices that can be tractable for use and interrogation by product teams and internal privacy stakeholders. Dark patterns can appear in many other domains, notably including ecommerce, wellbeing and even in "real-life contexts", as recognised by scholars.

The team sought to understand and build the definition of the problem based on external perceptions of dark patterns, drawing on a wide range of available research, publications and guidance. The team reviewed the external literature, including academic, legal, government, civil society, and industry publications. This process involved synthesising a range of relevant publications, with a focus on definitions and types of dark patterns to guide product practices. This synthesis was validated with practitioners, adjusting the categorization and framing to ensure the research insights are relevant and practical to designers, engineers, content designers, product managers and others who build digital services.

During this work, the team uncovered more than 50 definitions and descriptions for the term 'dark patterns' with the initial goal to identify a definition that helps to answer these questions: What are the concretely identifiable parts that allow us to identify something as a privacy dark pattern? What is a dark pattern vs. what is just bad design? How can this definition be used and misused? How can we develop a framework that is durable enough to apply across different contexts, can be built upon and can also be applied practically.

There is no single definition that adequately addresses the complexity of identifying, mitigating, and responding to dark patterns for every use case, but there are identifiers and markers that can aid in doing so. From this, the team developed a taxonomy to codify privacy dark patterns by developing a framework which will allow practitioners to name, identify, and then mitigate dark patterns concerns.

c. **Developing a privacy dark pattern taxonomy**

The team developed a dark pattern taxonomy that describes 6 product and design practices that are at risk of being perceived as dark patterns. In summary, the key areas include:

- **Interface Interference:** Interface Interference comprises design practices that distract users from certain options and/or manipulate their understanding through style, copywriting and other interface design techniques. *Example: A prompt with prominent 'allow' button and a greyed-out 'skip' option that is difficult to find.*
- **Hard to Opt-out:** Hard to Opt-out comprises product practices that make opt-out experiences more difficult than they inherently need to be, in order to dissuade people from revoking or adjusting their data permissions. *Example: An overly long process with additional friction for people to change the visibility of their contact details.*

- **Nagging:** Nagging comprises user experiences where requests to turn on or adjust data permission are presented unnecessarily and/or repeatedly and in contexts that are disconnected from user actions. *Example: A service that repeatedly reminds people to complete their profile when they log in.*

- **Forced Timing and Action:** Forced Timing and Action comprises situations in which users are forced to make a privacy decision on the spot and/or on a short and arbitrary deadline, thereby accelerating the user's decision-making process and pressuring the user to act. *Example: A privacy notice that arbitrarily creates urgency or fear to share data with a 3rd party service.*

- **Privacy Intrusive Defaults:** Privacy Intrusive Defaults comprises products where data permission defaults promote widespread collection and/or use of unnecessary personal data. *Example: A browser that collects payment information by default.*

- **Rewards and Punishment:** Rewards and Punishment comprises product practices where products incentivize data sharing, associating user choices that share data with arbitrarily assigned rewards and choices that restrict data with punishments. *Example: A prompt that punishes people for not sharing their location data by removing safety features.*

Over the last year, the team has shared their work on the dark pattern taxonomy with Data Protection Authorities, with more than 30 industry members as well as with academic stakeholders, seeking feedback, validation and working to further strengthen Meta's approach. This work to address dark pattern concerns is continuing and there is significant further development both in terms of development of mitigation practices as well as implementation underway. In addition, we will continue to revise and develop this work in line with evolving industry practices. We hope that the EDPB's work will complement and provide support to industry- lead and applied approaches like this taxonomy as it continues to develop, while Meta will continue its work to share valuable research findings and insights, as well as propagating good design practices through TTC Labs and similar initiatives.

7. Conclusion

We welcome the EDPB's interest in this important issue. We strongly urge that guidance on dark patterns – a common issue that is a concern for users online – is applied across the full range of relevant digital services. The goal should be to provide a consistently fair and high quality experience for people online. The range of issues this guidance addresses are common across other services.

We agree with many of the concerns and recommendations in the Draft Guidelines and are committed to continue our work to mitigate the risks of dark patterns. Via TTC Labs, we have invested in cross- industry efforts to innovate on good design through collaborative approaches, as well as initiatives to share relevant primary research and have proactively driven internal work to identify and mitigate dark patterns via our own dark patterns taxonomy.

The Draft Guidelines would benefit from improved structure, as well as providing evidence to support some of the strong assertions that are made in a number of specific areas. There is even risk of counterproductive effects where the Draft Guidelines make unsupported suppositions that are inconsistent with best practices for design or the integrity and consistency of the user experience. Finally, there is ambiguity in language and expression which may make it difficult for data controllers to understand and implement the requirements.

We appreciate the opportunity to comment on the Draft Guidelines and support the broad aims of the EDPB to provide clarity, consistency and fairness in user interfaces. To best achieve this goal, revisions of the Draft Guidelines should seek greater clarity, a refined scope which is based meaningfully on evidence and clear, and unambiguous articulation of the relevant objective principles which should be applied.