

BRIEFING NOTE

Prototyping User Empowerment

Towards DSA-compliant recommender systems

CONTENT

- 1**
Executive Summary
- 3**
Towards DSA-Compliant Recommender Systems
- 5**
UX principles for empowering user choice
- 6**
Design provocations for empowering user choice
- 18**
Endnotes

Executive Summary

What would a healthy social network look and feel like, with recommender systems that show users the content they *really* want to see, rather than content based on predatory and addictive design features?

In October 2022, the European Union adopted the Digital Services Act (DSA), introducing transparency and procedural accountability rules for large social media platforms – including giants such as Facebook, Instagram, YouTube and TikTok – for the first time. When it comes to their recommender systems, Very Large Online Platforms (VLOPs) are now required to assess systemic risks of their products and services (Article 34), and propose measures to mitigate against any negative effects (Article 35). In addition, VLOPs are required to disclose the “main parameters” of their recommender systems (Article 27), provide users with at least one option that is not based on personal data profiling (Article 38), and prevent the use of dark patterns and manipulative design practices to influence user behaviour (Article 25).

Many advocates and policy makers are hopeful that the DSA will create the regulatory conditions for a healthier digital public sphere – that is, social media that act as public spaces, sources of quality information and facilitators of meaningful social connection. However, many of the risks and harms linked to recommender system design cannot be mitigated without directly addressing the underlying business model of the dominant social media platforms, which is currently designed to maximise users’ attention in order to generate profit from advertisements and sponsored content. In this respect, changes that would mitigate systemic risks as defined by the DSA are likely to be heavily resisted – and contested – by VLOPs, making independent recommendations all the more urgent and necessary.

It is in this context that a multidisciplinary group of independent researchers, civil society experts, technologists and designers came together in 2023 to explore answers to the question: ‘How can the ambitious principles enshrined in the DSA be operationalised by social media platforms?’. On August 25th 2023, we published the first brief, looking at the relationship between specific design features in recommender systems and specific harms.¹ Our hypotheses were accompanied by a list of detailed questions to VLOPs and Very Large Online Search Engines (VLOSEs),

ACKNOWLEDGEMENTS

This brief was drafted by Jesse McCrosky (Thoughtworks), Tanya O'Carroll (independent expert), Caroline Sindere (Convocation Research + Design) and Katarzyna Szymielewicz (Panoptykon Foundation). The ideas and recommendations were generated in two multidisciplinary workshops, held in May and July 2023, which brought together independent researchers, civil society experts, technologists and designers.

Additionally, we would like to acknowledge input and contributions from the following experts:

Sara Bundtzen, Research and Policy Analyst, Institute for Strategic Dialogue Germany

Marc Faddoul, Director, AI Forensics

Maximilian Gahntz, AI Policy Lead, Mozilla Foundation

Emily Gorcenski, Data Scientist, Thoughtworks

Fernando Hortal Foronda, EPD's Digital Policy Advisor

Abhishek Sharma, Senior Designer, Superbloom

Kris Shrishak, Senior Fellow, Irish Council for Civil Liberties

Margaux Vitre, User Researcher, École Normale Supérieure

Molly Wilson, Senior Designer and Researcher, Superbloom

which serve as a 'technical checklist' for risk assessments, as well as for auditing recommender systems.

In this second brief, we explore user experience (UX) and interaction design choices that would provide people with more meaningful control and choice over the recommender systems that shape the content they see. We propose nine practical UX changes that we believe can facilitate greater user agency, from content feedback features to controls over the signals used to curate their feeds, and specific 'wellbeing' features. We hope this second briefing serves as a starting point for future user research to ground UX changes related to DSA risk mitigation in a better understanding of user's needs.

This briefing concludes with recommendations for VLOPs and the European Commission.

With regards to VLOPs, we would like to see these and other design provocations user-tested, experimented with and iterated upon. This should happen in a transparent manner to ensure that conflicting design goals are navigated with respect to the DSA. Risk assessment and risk mitigation is not a one-time exercise but an ongoing process, which should engage civil society, the ethical design community and a diverse representation of users as consulted stakeholders.

The European Commission should use all of its powers under the DSA, including the power to issue delegated acts and guidelines (e.g., in accordance with Article 35), to ensure that VLOPs:

- ▶ Implement the best UX practices in their recommender systems
 - ▶ Modify their interfaces and content ranking algorithms in order to mitigate systemic risks
 - ▶ Make transparency disclosures and engage stakeholders in the ways we describe above.
-

Towards DSA-Compliant Recommender Systems

DSA PROVISIONS

There are five key provisions within the DSA, which lay out new obligations for VLOPs when it comes to how they design and deploy recommender systems across their products and services. These include:

Article 25	Prohibiting the use of dark patterns and manipulative design practices to influence user behaviour. ²
Article 27	Ensuring recommender systems' "main parameters" are described in plain language with options to modify or influence these.
Articles 34 and 35	Requiring VLOPs to first assess and then put in place reasonable, proportionate and effective mitigation measures to systemic risks identified, including testing and adapting their recommender systems.
Article 27	Ensuring that users are provided with at least one option for each of their recommender systems that is not based on personal data profiling.

MITIGATION

Considering these requirements, a user-centred interpretation of DSA-compliant recommender systems will need to introduce mitigation measures at three levels:

1 User-facing transparency & explainability	User-facing transparency and meaningful explainability of the signals being used to recommend content are important for users to make informed decisions when curating their feeds. Moreover, both are necessary first steps for civil society organisations and regulatory bodies aiming to keep platforms accountable.
2 Safe defaults for everyone	Users should not be responsible for making their experience on social media platforms safe. Safe defaults are paramount, considering that most users lack the awareness, time or skills to customise their experience. This should include top-down interventions to ranking algorithms, in order to make them less dependent on engagement, and thereby safer for all users.

3 Advanced tools for empowering user

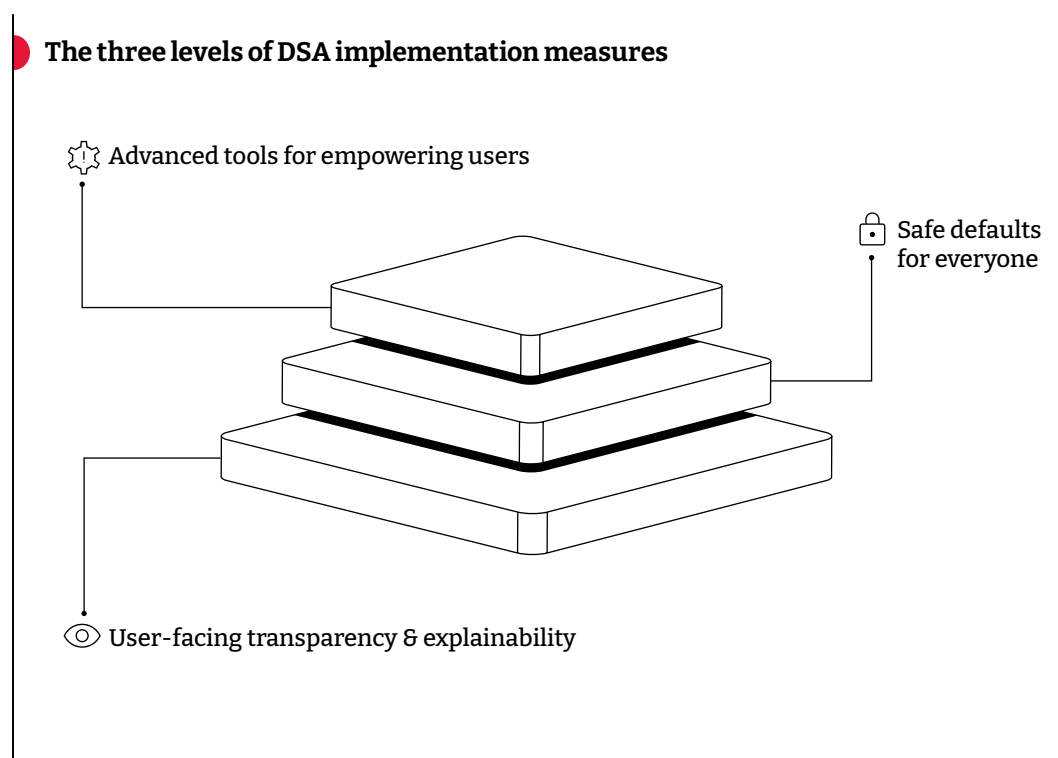
Users who want more control over their social media feeds should be encouraged. Since there is no 'one size fits all', VLOPs should offer a range of content curation tools, allowing users to control the signals that are used by recommender systems to shape their feeds. This should include providing a choice of recommendation logics, based on outcomes that users value, such as exploration or well-being.

User-facing transparency, including meaningful explanations of how recommender systems work, is an important foundation for compliance. We have outlined some of the technical data that VLOPs must make available in our first brief, *Fixing Recommender Systems: from identification of risk factors to meaningful transparency and mitigation*.

In this brief, we take a deep dive into the third element: advanced tools for empowering user choice. Much like transparency, this is an area where simple, clear and uncontroversial solutions are within reach, and it is for this reason that we start here. It is, however, important to underscore that individual choice will only partially address the many documented harms connected to recommender systems.

In many ways, the most important and urgent mitigation measures belong to the second category, which we called 'safe defaults'. Such measures shall include top-down interventions to re-engineer recommender systems away from dependence on engagement-based rankings. It is this more complex area of research that will be the subject of our future work and recommendations.

FIGURE 1



UX principles for empowering user choice

Design expertise will have a vital role to play in shaping regulatory compliance when it comes to recommender systems. Design expertise can help avoid the kind of failures that we saw when online platforms implemented earlier regulations, including the ePrivacy Directive and EU 2018 General Data Protection Regulation (GDPR) – which saw many platforms opt to introduce burdensome and non-context sensitive cookies banners, making users’ informed and free consent impossible.

In order to avoid similar failures, designers should be considered as integral stakeholders in interpreting and translating the DSA into UX design and features that will shape users’ actual experiences on social media and other platforms.

The following UX principles and recommendations should guide VLOPs’ approach to implementing the DSA as regards user control and empowerment over recommender system features:

UX PRINCIPLES AND RECOMMENDATIONS

- ▶ Users shouldn't have to understand in detail how a recommender system works in order to be able to exercise choice and control over their feeds.
- ▶ Platforms need to avoid information overload and user fatigue with unnecessary friction.
- ▶ Platforms need to ensure that new features are easy to find and access across multiple spaces within a product.
- ▶ Platforms must respect user feedback and show the user that their feedback has resulted in concrete changes.
- ▶ Users should have an option to reset their personalised feed (start clean) or try alternative feeds (e.g., a feed that is not based on data profiling, or is based on a restricted set of explicitly-provided signals).
- ▶ Users should have control over which signals (including personal data) are used to make recommendations.
- ▶ Users should have control over trade-offs that are made between competing objectives in a recommender system.
- ▶ Users should be able to customise content displayed to them using advanced options (e.g., prioritise certain sources of information or topics).
- ▶ Users should be able to protect their wellbeing (e.g., by setting granular limits for social media use or personal filters for unwanted content).
- ▶ Users should be able to use a service for multiple purposes and have their recommendations reflect their purpose for a given session (e.g., sometimes, the user may request to be entertained or kill time, while at other times they may want to see only educational content).
- ▶ Users should be provided with different entry points to access control settings and transparency features, from intuitive, high level settings for most users to advanced, granular settings for super users.
- ▶ Users should not be offered settings that provide a ‘fallacy of control’ (such as a ‘why am I seeing this?’ feature that does not provide a meaningful answer). Settings should provide real and actionable value to the user.

Design provocations for empowering user choice

We propose nine practical UX changes that we believe can facilitate greater user agency, from content feedback features to controls over the signals used to curate their feeds, to specific ‘wellbeing’ features. To meaningfully empower users, we believe these UX changes could be a starting point for further user research, and we invite VLOPs, and the Human Computer Interaction (HCI) research community to build upon these.

FEEDBACK FEATURES

Feedback features encourage users to provide granular feedback on specific pieces of content they view, including:

1. Instant explicit feedback: whether the content is desired or undesired
2. Complaint mechanisms: flagging, filtering or reporting content

CONTENT CURATION FEATURES

Content curation features encourage and allow users to convey specific expectations and preferences about the type of content they wish to see (and when):

3. ‘Choose your feed’: A non profiling-based feed
4. ‘What are you here for?’: Onboarding screen to learn about user’s intention
5. ‘What matters more to you?’: Balance sliders to control trade-offs made by the algorithm
6. ‘Control what feeds your feed’: Signals dashboard
7. ‘See how your feed has changed’: Daily/weekly report

WELLBEING FEATURES

Wellbeing features help users to protect themselves from harms such as compulsive use of social media and exposure to borderline or simply unwanted content:

8. ‘Set your timer to sign off’: Prompts and reports on screen time
9. ‘Make your feed a safe space’: Protective filters

For each design proposal, we outline its value for the user, its limitations and how it might be implemented in the real world. We also suggest how VLOPs can direct their users towards more informed choices, instead of trapping them in default settings that are not oriented towards their empowerment.

1

Instant explicit feedback: Whether the content is desired or undesired

Feedback Feature

Why we think it is important

Social media platforms typically offer a 'like' button or similar functionality for explicit feedback, which serves a dual purpose: social and curational. Because of the social purpose, users may feel pressure to 'like' content (to send a positive feedback to a content creator and their network) even when they would rather see less, not more, of it).

It is therefore important that these two functions are separated. Users should be able to show support for content creators or indicate their social allegiances independently from teaching the algorithm (recommendation engine) whether or not they want more recommendations similar to the content that has been 'liked'.

What it does

An intuitive tool for users to indicate whether the content they have been recommended is wanted or unwanted. Users who want to give more granular feedback (say why the content is unwanted) should be able to do so easily, in no more than one more click. By sending this explicit feedback, users can shape what is recommended in the future, independently of social feedback they may want to give other users (such as 'likes').

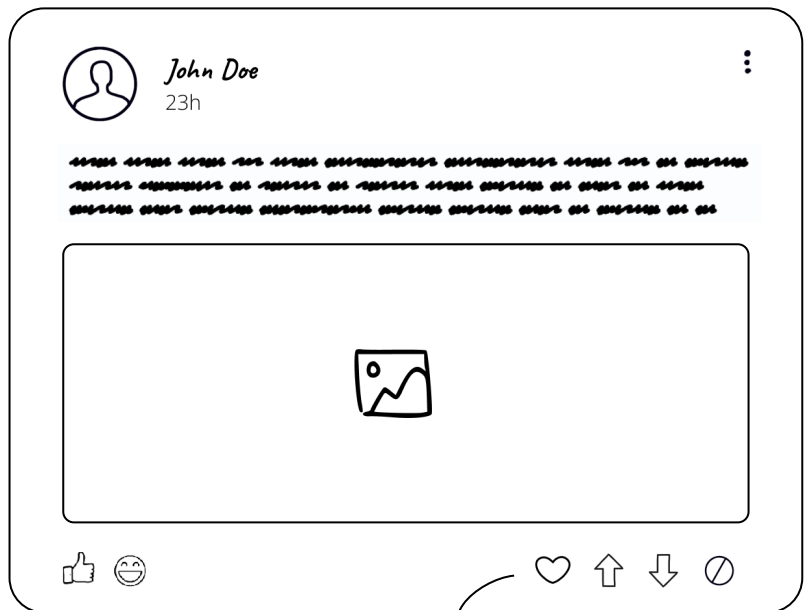
Limitations


Feedback is, by definition, reactive. It can achieve a lot but it alone cannot be used to proactively shape a user's feed. Ultimately, user feedback is just another category of signals – or type of data – fed into the recommender system. It can be taken into account or ignored by the algorithm, with very limited transparency and accountability.


How should it feel? What could it look like?


A granular feedback feature should be easily accessible, e.g. by pressing longer on a piece of content or a feedback button.


Granular feedback options:



 Remember I like content like this

 Continue showing me content like this

 Show me less content like this

 Never show me content like this

<div style="display: flex; justify-content: space-between; align-items: center;"> 2 <div style="flex-grow: 1;"> <h2 style="margin: 0;">Complaint mechanisms: Flagging, filtering or reporting content</h2> </div> Feedback Feature </div>	
<p>Why we think it is important</p> <p>Complaints serve many functions, from reporting technical bugs to dealing with everyday problems such as forgotten passwords or reporting spam. Complaints may also cover harassment cases, including ongoing harassment issues that require investigation and support from trust and safety teams.</p> <p>Currently, complaints (including content reports) are sent off into the ether with few ways for users to contact a person, elevate a complaint or gain insight into how and when their complaint will be resolved. When complaints are answered users often lack the ability to contest the result.</p>	<p>Limitations</p> <p>Needless to say, even the most effective and transparent complaint mechanism does not, in itself, solve problems which trigger complaints.</p> <p>While it helps improve a user’s overall experience, every platform needs to take real actions to solve problems reported by its users.</p>
<p>What it does</p> <p>When something goes wrong, users need a simple way to signal this issue, receive help and contest an inadequate or unfair response. This is an essential condition of reclaiming agency.</p> <p>In addition to having standard feedback features, users need a channel to send stronger feedback to the recommendation engine, such as ‘never show this type of content to me, because I find it harmful’. All of this can be achieved with a well-designed complaint mechanism.</p>	<p>How should it feel? What could it look like?</p> <p>Complaint mechanisms can span a variety of designs, including flagging, filtering or reporting content.</p> <p>What makes them effective is the next step, which should be a friendly feature allowing for a back and forth between the user and the platform.</p>

3

‘Choose your feed’: A non profiling-based feed

Content Curation Feature

Why we think it is important

The DSA requires that VLOPs offer their users at least one version of the recommender system that is not based on profiling (i.e. does not use their personal data, including behavioural observations, to shape recommendations).³ Article 38 does not say or specify what logic an alternative version of the recommender system should be. In practice, VLOPs usually provide an option to access either a chronological feed (top recommendations reflecting what has been posted most recently by accounts followed by the user) or a trending feed (top recommendations reflect what is most popular in the user’s social network). Both options have their flaws and do not allow for meaningful user choice. On top of what is required by the DSA, users should be offered a third option: a tailored feed based on their explicit preferences (such as their subscriptions) and explicit feedback (e.g., “show me more of this” or ‘show me less’).

What it does

With one click or swipe, users can switch between a default personalised feed (based on their behavioural signals and implicit feedback), an alternative feed designed by the platform (not based on profiling, such as chronological or ‘trending’ feed), or a third way: a tailored feed curated by the user (based on their explicit feedback and freely-declared information).

Limitations

The default is what matters most. As long as the law does not require that social media platforms offer their users a healthier, safer feed (not based on tracking and profiling) as default, alternatives will only attract advanced, conscious users. Nevertheless, it is essential that they exist and are developed into an increasing number of attractive options.

How should it feel? What could it look like?



4

‘What are you here for?’: Onboarding screen to learn about a user’s intention

Content Curation Feature

Why we think it is important

Users typically have different modes of interaction with the platform or, as we define it, ‘algorithmic personas’. They might use the platform to look for specific information, for idle entertainment or to relax before bed. When the user logs into the platform, the algorithm has no such context and will therefore attempt to engage the user with a wide range of content corresponding to their various personas. In some cases, this can be incongruous for the user, for instance when they are distracted with entertainment content while they are accessing the platform looking for news.

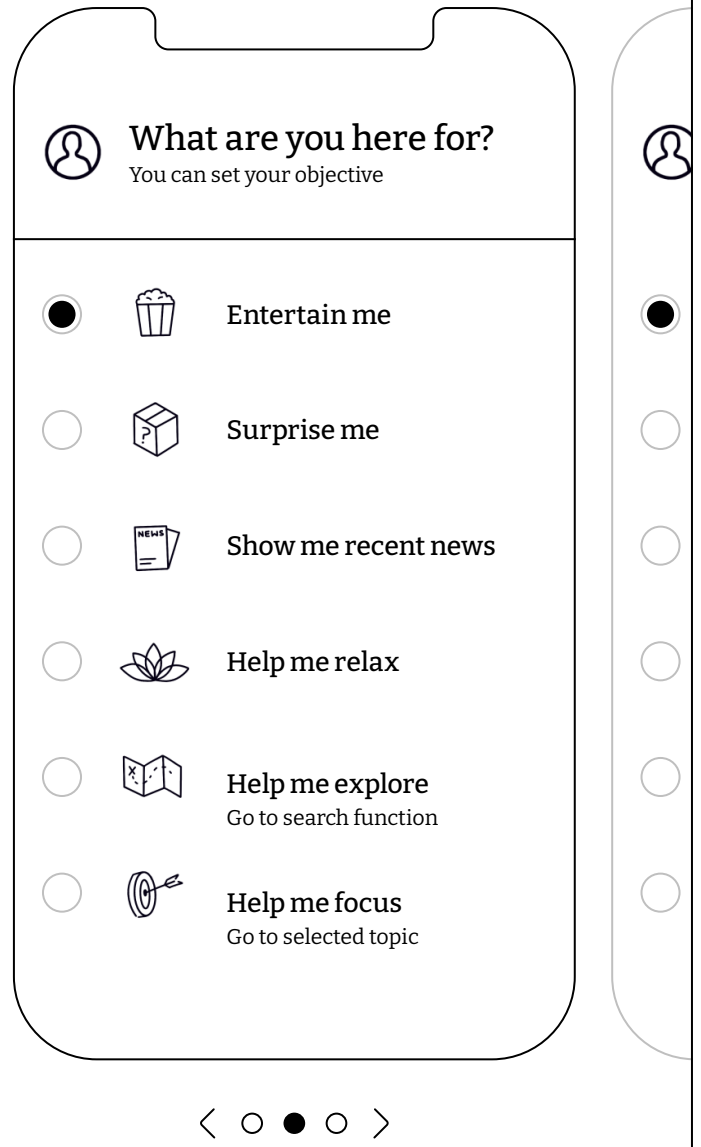
What it does

A product feature that explicitly asks users about their state of mind/intention for the specific session should increase the quality of recommendations (including for the tailored feed option we describe above). In theory, adjusting recommendations to the user’s state of mind/intention should also increase meaningful engagement. In practice, the onboarding screen adds friction by making the user stop and reflect ‘what is my intention for this session?’. In many cases, the answer to that question might be ‘nothing really’ or ‘procrastination’ (which is fine, as long as it is a choice and not a compulsive behaviour encouraged by addictive design features). This short moment of reflection may lead some users to leave the app. For users who enter the platform with a specific objective, an onboarding screen will be an opportunity to improve their experience. It sends a clear, explicit signal on user expectations to the recommender engine. Implementing this feature could therefore lead to less engagement, but of higher quality, and on content more aligned with the user’s intention. This would represent a net win for users.

Limitations

Expressing a preference by clicking one option on the onboarding screen will still be interpreted by the platform. Users rely on the platform to know (find out based on behavioural observations) what will be ‘entertaining’ or ‘relaxing’ for them.

How should it feel? What could it look like?



5

'What matters more to you?': Balance sliders to control trade-offs made by the algorithm

Content Curation Feature

Why we think it is important

When ranking recommended content, algorithms make trade-offs, balancing different types of content to provide a diverse, and thus engaging experience. For instance, should locally-relevant content be prioritised above globally popular content? Should the content be more informative or entertaining? Should the feed suggest new content or exploit known areas of interest? Most of these recommender system trade-offs are not visible to the user, and they tend to prioritise user engagement over other metrics. We believe a set of trade-offs should be visible and that recommendations should be optimised for metrics other than user engagement. Users would also have the agency to select the trade-off they prefer, depending on their current needs.

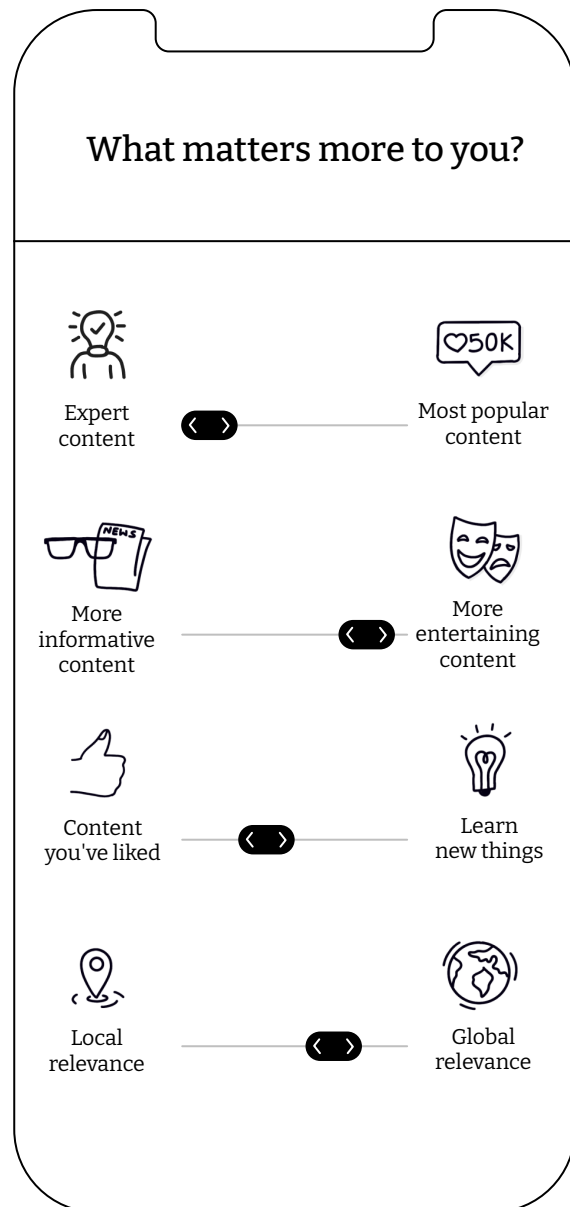
What it does

With this feature, users can control trade-offs that are made by recommender system designers to maximise their engagement and give their own instructions (or guidelines) to the algorithm. They can express their preference on scales that are understandable and relatively easy to describe algorithmically (e.g., local vs global relevance). By expressing their preference for one experience over another (e.g., 'give me more informative content' over 'give me more entertaining content') users gain more granular control on the composition of their feed. This feature is complementary to 'What are you here for?'; and, with the balance sliders, users can give instructions to the algorithm that endure beyond a particular session.

Limitations

This feature relies on the algorithm to interpret users' preference. For some trade-offs (such as local vs global relevance), categorisation will be quite objective and straightforward. In other cases we will have to rely on the platform to categorise content as 'informative' or 'entertaining'. This imbalance of power is fundamental to centralised social media, where content is hosted and curated by the platform provider. This is also why we demand meaningful transparency and independent auditing of all mitigation measures introduced by VLOPs.

How should it feel? What could it look like?



6

'Control what feeds your feed': Signals dashboard

Content Curation Feature

Why we think it is important

Recommendation systems use a wide variety of signals (often referred to as 'features' in machine learning) to make predictions about what will be the most engaging content. Signals include interactions with previously recommended content – both explicit (such as clicking 'show me more content like this') and implicit (such as watching a video or lingering on a post). Beyond these behavioural signals, recommender systems also use (inferred) location, device information and signals from the user's social network (e.g., 'your friend X commented on this post'). Users currently can't control which of these signals will be used to customise their feed.

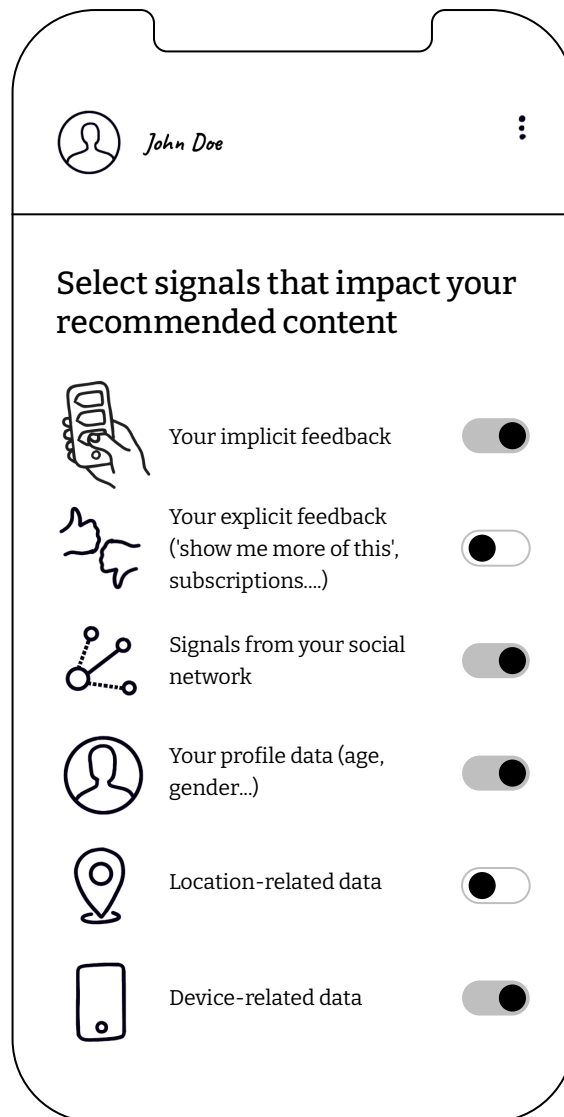
What it does

With this feature, users can learn about types of signals that feed into their recommender system and control them. In particular, they can block types of signals which they perceive as sensitive, such as device-related data or records of their behaviour on the platform. Users can experiment and learn how to change/improve their experience with the recommender system, even without fully understanding their logic.

Limitations

The architecture of machine learning systems makes it difficult to track all signals that have an impact on the final output – in this case on recommended content. Therefore, this feature does not provide an explanation ('why am I seeing this?') as to why a user is seeing specific content. It allows for experimentation, which will affect the whole feed.

How should it feel? What could it look like?



7

‘See how your feed has changed’: Daily/weekly report

Content Curation Feature

Why we think it is important

Quite a few feedback and content curation tools already exist on social media platforms, but users and researchers have questioned their effectiveness. Several reports have shown that explicit feedback (‘show me more/show me less’) and expressed preference for specific sources or topics had little or no impact on recommended content.⁴

In this context, it is not surprising that users feel discouraged from making choices and they may perceive additional prompts as ‘friction’. The best way to motivate users to make this effort is to show them how their feedback will impact their feed.

What it does

This feature invites users to explore how their choices have influenced what has been recommended to them (e.g., how their feed has changed after they blocked certain signals from what is usually fed into the system, or expressed preference for more informative content over entertainment). Through this interactive tool, users can learn how recommender systems work, and monitor whether their choices have been respected by the platform.

Limitations

This feature showcases examples of posts that experience a significant increase or decrease in their ranking scores after the preferences were updated by the user. A potentially even more insightful approach to transparency would be to enable users to trace back specific signals that have influenced a particular recommendation. However, this could be computationally prohibitive, unlike the approach we suggest here, which should be relatively straightforward to implement (as this only requires recomputing the ranking scores on a pool of post candidates, based on the updated preferences).

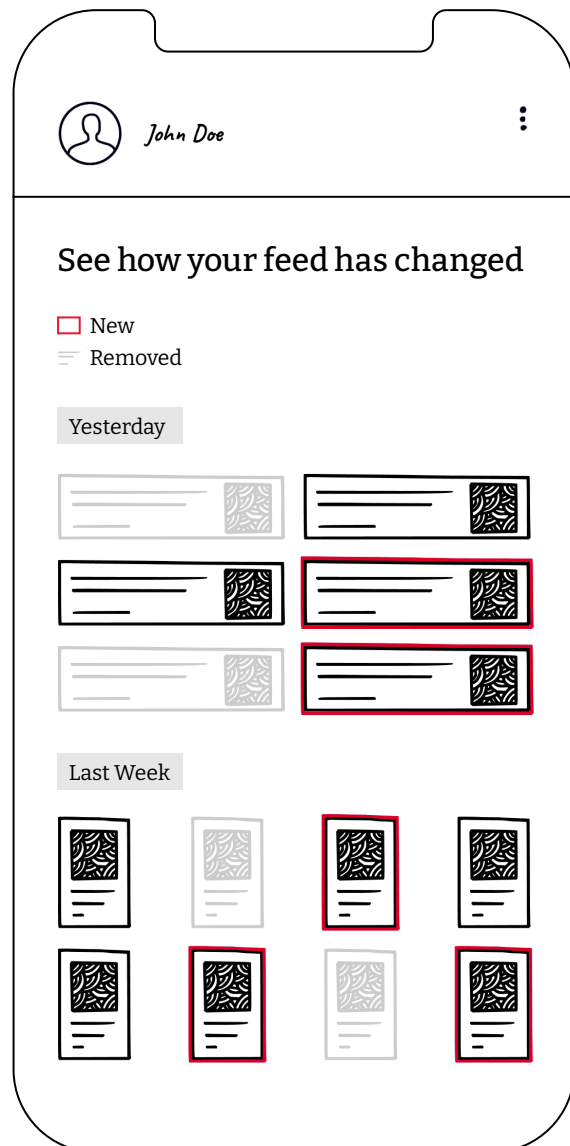
How should it feel? What could it look like?

An embedded post that appears on the feed and prompts users to check what has changed as a result of their choices, e.g.:

‘Do you like new content in your feed? Learn where it came from!’

After expressing a preference or changing a setting, the user would be shown examples of posts that have been removed or added to the feed due to their action.

From there, a user is directed to a dashboard for more insights:



8

‘Set your timer to log off’: Prompts and reports on screen time

Wellbeing Feature

Why we think it is important

Social media apps are designed to encourage frequent and extended engagement, leading their use to become habitual. Notifications pull users in, while there is some evidence that personalised recommendations may trigger a reward mechanism in users’ brains. Compulsive social media behaviour is reinforced by technological affordances that allow users to enjoy a frictionless browsing experience.

For all of these reasons, it is unrealistic and unfair to expect that users will avoid this trap and choose to discipline themselves with self-imposed measures. VLOPs should proactively offer features that cater for users’ wellbeing and curb the addictive potential of their services.

What it does

Users are encouraged to set and respect their daily time limit for social media. This feature allows users to set different targets for entertainment, work or social activity, thus reflecting different personas users have in their social media consumption.

More importantly, it changes the framing from (negative) ‘set your time limits’ to (positive) ‘set your timer’ and ‘save time for something that matters to you’. In this framing, every minute away from social media means one more minute for family life, physical activity or focused work.

Limitations

This feature can reduce harms caused by prolonged social media usage and exposure to content that triggers anger, anxiety, low self-esteem or compulsive behaviour. However, it will not adequately address the mental health harms connected to the business model of the dominant social media platforms as long as they continue to be optimised for engagement.

Catering for the digital wellbeing of individual users is important, but it should not be seen as a long-term solution in risk mitigation.

How should it feel? What could it look like?

Prompts encourage users to set their ‘timer’ or their own ‘target for free time’ (i.e. time spent outside of the platform), making this choice as granular as they need. Fun weekly reports show users how much time they have saved by observing self-imposed targets, or how much time they have lost by yielding to addictive social media features.

9

‘Make your feed a safe space’: Protective filters

Wellbeing Feature

Why we think it is important

There is no way to remove all harmful content from social media, not least because many examples of harmful content remain legal and are deeply contextual, meaning different things in different places and to different people. Policies for content moderation will therefore always be contested, socially negotiated and imperfect. A more effective way to protect individual users from exposure to what *they* perceive as harmful or unwanted (in a given moment of their life) is to let *them* curate their own feed, without implications for what others can see in their feeds.

What it does

This feature gives agency to the user, allowing them to make their feed a safer space without having to wait for a platform to moderate or remove content that they personally find harmful. Users can choose whether content flagged by the platform as ‘likely to be borderline’ should be suppressed at all costs – even if this means accidental suppression of innocuous content – or not. For example, some people may choose to suppress all sexualized content, while others will welcome it. Users can also curate their feed by filtering out content based on keywords (relating to anything they find problematic).

Limitations

This provocation is a ‘band-aid’ to increase user safety with regard to legal content, which should not be removed but still can be perceived as harmful by some users. This approach does not replace content moderation and does not reduce the need for more effective, more nuanced moderation. While this approach is user driven, it also requires users to invest their time in labelling content they find unwanted or harmful. The more granular and explicit the feedback, the better chance that the algorithm will learn and follow individual preference for avoiding specific types of content.

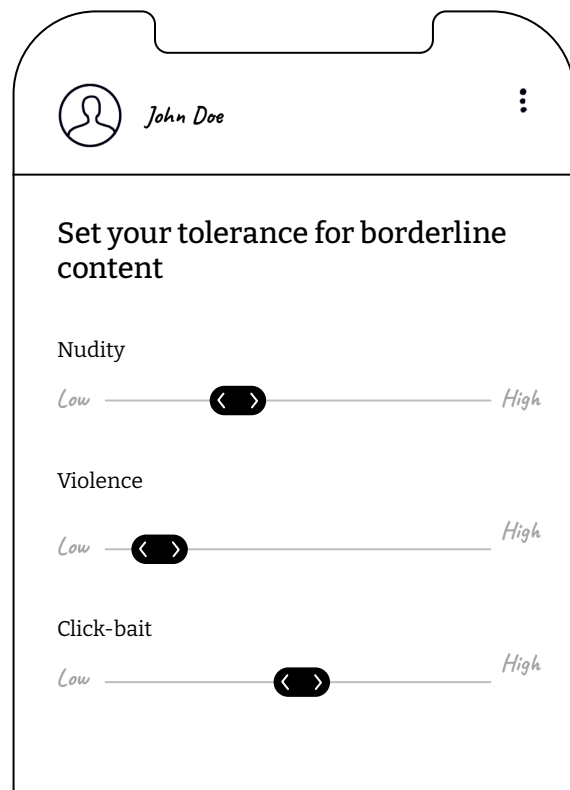
How should it feel? What could it look like?

On/off feature, which removes all content that is likely to be borderline from the feed (based on labelling done by the platform).



The option to filter out (remove from feed) content based on keywords selected by the user (anything they find problematic).

Balance sliders that allow users to set their tolerance for different classes of borderline and (subjectively) harmful content, such as sexualized content and legally-permitted nudity, verbal aggression and profane language, content that may trigger body image issues, and clickbait promoting paranoia and mistrust.



This feature can be designed differently for different classes of borderline and (subjectively) harmful content.

Conclusion and Recommendations

Designing interface features to increase user agency is a challenging task. Yet, as researchers, technologists and designers, we believe this can be done. We also acknowledge that each of the solutions we put forward as provocations for VLOPs and the design community has its flaws and limitations. However, we won't create solutions without experimentation.

We would like to see these and other design provocations user-tested, experimented with and iterated upon by online platforms. This should happen in a transparent manner, to ensure that conflicting design goals are navigated with respect to the DSA. Risk assessment and risk mitigation is not a one-time exercise but an ongoing process, which should engage civil society, the ethical design community and a diverse representation of users as consulted stakeholders. Below, we summarise our expectations towards both VLOPs, which are responsible for risk assessment and mitigation, and the European Commission, which is responsible for effective implementation of the DSA.

RECOMMENDATIONS TO VLOPS

Test these ideas and disclose the results of user testing and feedback

Civil society and researchers can propose design solutions as mitigation measures, but only VLOPs can test how they work in practice. The only way to make expert debate about alternative propositions possible is to have VLOPs reveal results of their A/B testing, user testing, usability testing, etc. A good starting point would be to begin testing and building upon the design provocations provided here and share the results and feedback.

Disclose research findings from risk assessment and mitigation

According to the DSA, VLOPs have to assess systemic risks related to the functioning of their recommender systems on an ongoing basis.⁵ It is important that this evidence is shared with civil society organisations – in particular those representing consumers – as well as researchers, academics and relevant communities of users, especially those who may be facing disproportionate exposure to risk of harm on the platform.

Research conducted by VLOPs can be the source of valuable, actionable information for these stakeholders. VLOPs should also disclose their research findings and their rationale when deciding to implement specific mitigation measures (including user research).



Provide better insights into product tooling

Designers and technologists working with VLOPs need to understand how to operationalise and interpret the DSA to make more responsible decisions. Such understanding will benefit from a dialogue with civil society experts and ethical design communities. However, these stakeholders can't provide cogent and adequate recommendations for VLOPs on how to improve their user interfaces without insights into how a particular product has been designed.

Revealing key product decisions and related information should be an element of transparency disclosures (e.g., documentation of design processes and design sprints related to specific products, design decisions and testing for a new feature, the specific user needs and related user research that shaped a product decision, etc.)

Engage civil society experts, ethical design communities and impacted groups in risk assessment and mitigation

According to the DSA, VLOPs should engage “representatives of groups potentially impacted by their services, independent experts, and civil society organisations” when conducting their risk assessments and designing their risk mitigation measures.⁶ Companies should embed such consultations into their methodologies, including surveys, focus groups, round tables etc. Much more transparency is needed in how VLOPs engage with civil society and how they implement feedback they have received from all stakeholders.⁷

Acknowledging the high societal impact of their core services, VLOPs should create space for uninvited feedback from civil society organisations and affected communities.

RECOMMENDATIONS TO THE EC

The European Commission should use all of its powers under the DSA, including the power to issue delegated acts and guidelines (e.g., in accordance with Article 35), to ensure that VLOPs:

- ▶ Implement the best UX practices in their recommender systems
- ▶ Modify their interfaces and content ranking algorithms in order to mitigate systemic risks
- ▶ Make transparency disclosures and engage stakeholders in the ways we describe above.

1. Panoptykon, Irish Council for Civil Liberties and People vs Big Tech: Fixing Recommender Systems. From identification of risk factors to meaningful transparency and mitigation, see online: https://panoptykon.org/sites/default/files/2023-08/Panoptykon_ICCL_PvsBT_Fixing-recommender-systems_Aug%202023.pdf
2. The ban complements, but does not overwrite the prohibitions already established under consumer protection and data protection rules, where a large number of dark patterns that mislead consumers are already banned in the EU.
3. Official text of the DSA: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2065>
4. Does this button work? Investigating YouTube's ineffective user controls, Mozilla Foundation (2022) See online: <https://foundation.mozilla.org/en/research/library/user-controls/report/>; I See Me Here: Mental Health Content, Community, and Algorithmic Curation on TikTok, Milton et al., (2023) See online: <https://dl.acm.org/doi/10.1145/3544548.3581489>; Algorithms of Trauma, Panoptykon Foundation and P. Sapieżyński (2021) See online: <https://en.panoptykon.org/algorithms-of-trauma>
5. Article 34 of the DSA [<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2065>]
6. Recital 90 of the DSA [<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2065>]
7. For detailed recommendations on how VLOPs should engage civil society and impacted communities in their risk assessments please refer to Towards meaningful fundamental rights impact assessments under the DSA, co-authored by Access Now and ECNL [<https://www.accessnow.org/wp-content/uploads/2023/09/DSA-FRIA-joint-policy-paper-September-2023.pdf>]