

Preventing Unwanted Experiences

HF 4400 mandates that platforms follow their users' explicit desires.

4 Main Features:

- Privacy by Default
- Responsible Amplification
- Sensible Rate Limits
- Do not encourage greater usage for children

4 Main Features:

- Privacy by Default
- Responsible Amplification
- Sensible Rate Limits
- Do not encourage greater usage for children

Plus **Transparency** of design choices and experimentation to understand new features that may elicit unwanted experiences.

Privacy by Default - The Problem

Many Cases of Manipulation and Fraud Begin with Unwanted Contact from Strangers, Facilitated by Loose Privacy Defaults and High Rate Limits, with Especially Serious Consequences for Youth

“On more than one occasion, while I was a minor I had received sexually explicit photos from men who added my account. I did not need to add them back to see the image they had sent me. No minor should ever be subjected to this....until the company makes changes, more minors will unfortunately get sent these unwanted pictures.”

- Quote from a complainant to the Minnesota Attorney General’s Office

FBI Sacramento
FBI Sacramento Press Office
(916) 746-8106

[Twitter](#) [Facebook](#) [Email](#)

January 17, 2024

Sextortion: A Growing Threat Preying Upon Our Nation’s Teens

Victims Manipulated to Create Sexually Explicit Material for Extortionists

In one study of adolescents in Utah⁴⁹

26%

reported getting involved in an unwanted conversation

17%

reported a stranger trying to meet them

Per internal research detailed in the Wall Street Journal⁵²

1/8

users under the age of 16 said they experienced unwanted sexual advances on Instagram

Privacy by Default - The Solution

Default settings:

- Account & content are not publicly visible and not allowed to be integrated into AI models.
- Messages cannot originate outside of an account's existing extended network (friends of friends). User can initiate messages / connections.
- Location cannot be revealed outside of friends.
- Content from that user cannot be forwarded to others.

Privacy by Default - The Solution

Default settings:

- Account & content are not publicly visible and not allowed to be integrated into AI models.
- Messages cannot originate outside of an account's existing extended network (friends of friends). User can initiate messages / connections.
- Location cannot be revealed outside of friends.
- Content from that user cannot be forwarded to others.

Users can change these settings anytime they want. But platforms cannot cajole them to do so.

Responsible Amplification - The Problem

Many Consumers are Having Experiences with Unwanted Disturbing, Graphic, and Sexual Content, Often Recommended by AI Powered Algorithms



“When she was 13, she started cutting herself. When asked why, she said that girls on Instagram talked about how it was exhilarating to cut yourself, so she did it.”

- Quote from a complainant to the Minnesota Attorney General’s Office



According to a US study by Common Sense Media ²⁶

12

the average age
of pornography
exposure online

15%

reported seeing
pornography at age
10 or younger

58%

of youth said they came
across pornography
online by accident

>50%

respondents said they
viewed violent porn

Responsible Amplification - The Solution

Optimize for what users

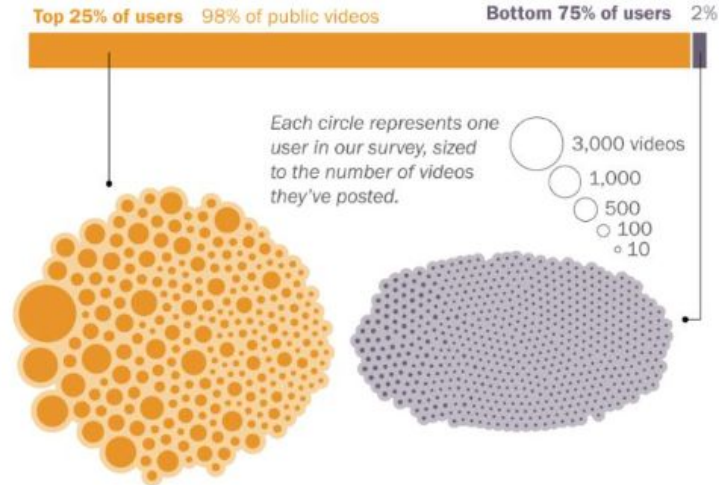
- 1) explicitly request to see
- 2) perceive to be high quality

NOT what they engage with.

Sensible Rate Limits - The Problem

The most active 25% of U.S. adult TikTok users produce 98% of public content

Share of videos from U.S. adult TikTok users produced by the ...



Note: Based on respondents who provided a valid TikTok handle.
Source: Survey of U.S. adult TikTok users conducted Aug. 7-27, 2023. Data on respondents' TikTok accounts collected using automated web scraping. Analysis includes all public videos accessible as of Nov. 7, 2023.
"How U.S. Adults Use TikTok"

PEW RESEARCH CENTER

DID YOU KNOW?



Most online misinformation comes from a small minority of "superspreaders," but social media amplifies their reach and influence.



AMERICAN PSYCHOLOGICAL ASSOCIATION

Sensible Rate Limits - The Solution

Sensible rate limits

1) new users cannot exceed median user activity for engaging with others (e.g. commenting, inviting people to groups, messaging others, making friend requests).

2) established users have their impact on others reduced as they exceed two standard deviations from median usage.

Do not encourage greater usage for kids - The Problem



"I have two teenagers, one is diagnosed with ADHD, Depression and anxiety. We have Bark to help monitor our kids usage but we only use it to shut off their phones or give them approved more time. They constantly sneak passwords and figure out ways to bypass the system....I feel they give no attention to their schoolwork or even care about it. I see their attention spans have decreased. They cannot focus on a person talking to them."

- Quote from a complainant to the Minnesota Attorney General's Office



Of Minnesota College Students...

55%

report having an issue with excessive computer/internet use⁶⁴

47%

report that excessive computer/internet use is affecting their academics⁶⁴

46%

report getting adequate sleep on three or fewer days a week⁶⁷

Do not encourage greater usage for kids - The Solution

Based on Device based identification - IF you have parental controls on a phone, social media platforms on this phone will not use platform features that increase, sustain, or extend a user's engagement with the platform beyond the user's express preferences regarding time or duration including:

- Optimizing feeds for time spent or content consumed
- Content feeds without a finite ending
- Auto-playing videos or other content
- Notifications that are not time sensitive
- Visible counts of engagement

Projected Impact:

The proposed solutions will not solve the problem of bad things happening online....

....BUT it will outlaw design choices that platforms know lead to negative outcomes and allow platforms that are responsible to compete on a level playing field with those that are not.

Misconceptions about this bill:

- It is too vague to comply with.
 - Per [this recent paper](#), many platforms already have efforts to optimize for quality and explicit user preference. More broadly, having worked at a platform for 4+ years, I can assure legislators that all these provisions are specific enough to be implementable and can point to specific work at platforms that would already comply, if rolled out more widely,
- Some terms are subjective.
 - Terms like “quality” are intentionally left to users to define, and platforms to elicit. The goal is to move from “time spent” to explicit judgments of preference and quality, as defined by users.
- Engagement limits restrict access.
 - Engagement limits are specific to how they affect other users. Users are free to read as much public content as they want. There is no right do dominate others’ experiences by posting/commenting/messaging/inviting a lot.
- It bans algorithms.
 - It does not. The bill mandates that algorithms optimize for users’ explicit preferences, rather than assuming that what people pay attention to is what they want more of.
- Privacy defaults restrict network effects.
 - Yes, this is true. But users care about their experience, not the networks (or the profit of businesses). Per Common Sense Media, more teenage girls believe that having public accounts has a negative effect (33%) than a positive effect (22%).