AMENDED Exhibit 1075

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

Teen wellbeing risks with GenAI creation tools

Author: [HYPERLINK "mailto:[redacted]\h]
Status: In Progress
Last updated: Feb 6, 2024
Privileged and confidential

Problem definition

The [HYPERLINK "https://docs.google.com/[redacted]\h] previously identified can still be applied for GenAI use cases. The GenAI specific risks that fall into these categories are below.

	Safety	Wellbeing
Problematic content	• Policy violating content • Factually inaccurate predictions (i.e., hallucinations)	• Biased content promoting stereotypes & discrimination
Overuse	• N/A	• Over-reliance on AI, limiting development of problem solving skills • Plagiarism in school assignments • Low self-esteem knowing what AI can achieve
Risky creation	• N/A	• Visual filters that emphasize standards of beauty, distorts views of self & others
Cyberbullying	• Bullying using AI generated images / deep fakes • Allowing online predators to more easily create fake accounts	• N/A

In this document, we will focus on the beauty filters problem, in particular.

- ***Why visual filters?***
  - **Disproportionate impact on teens**: Research show visual filters [HYPERLINK "https://www.tandfonline.com/doi/full/10.1080/15213269.2016.1257392" \h] than adults given teens are still forming their body image and [HYPERLINK "https://www.parents.com/kids/health/childrens-mental-health/how-social-media-

filters-are-affecting-youth/" \h] using visual filters noted they make them feel worse about their real-life appearance
    - [redacted] of [HYPERLINK "https://docs.google.com/[redacted]\h] / [HYPERLINK "https://docs.google.com/[redacted]\h] among the different YT GenAI products under development
- ***Why not biased content?***
    - **Problem for teens & adults alike**: Biased content is a problem for all users,
    - [redacted] HYPERLINK "https://docs.google.com/[redacted]\h]
- ***Why not overuse / over-reliance on AI?***
    - [redacted]

Breakdown of user problem

Four stages of impact on users viewing & creating visual filter edited images

1. *Everyone's perfect but me*: Viewing "perfect faces/bodies" in feed lead to [HYPERLINK "https://blog.petrieflom.law.harvard.edu/2022/11/15/the-filter-effect-what-does-comparing-our-bodies-on-social-media-do-to-our-health/" \h]
2. *I need to also look this way*: Conforming to social pressures and reinforcing beauty standards to others
3. *I cannot show my real self anymore*: Disconnect between how user looks vs. filtered images, leading to a need to filter every image
4. *I cannot meet people real-life & disappoint them*: [HYPERLINK "https://www.psychologytoday.com/us/blog/the-clarity/202303/can-beauty-filters-damage-your-self-esteem" \h] causing social anxiety & isolation

YT's visual filters & their usage

Roadmap of future effects

Competitive analysis / internal data?

Commented [1]: What type of [redacted] do competitors have if any?

What type of standards do they have for UGC filters?

Commented [2]: Competitive analysis breakdown for filters? What type of filters do they have?

	TikTok	Instagram	SnapChat
Attainability	No effects that alter the shape / appearance of natural features		
Diversity & inclusivity			
Transparency			
Cultural sensitivity			

- [HYPERLINK "https://effecthouse.tiktok.com/learn/guides/general/effect-guidelines" \h]
- [HYPERLINK "https://www.tiktok.com/community-guidelines/en/" \h]
- The “skinny filter” on Tiktok which makes user’s face slimmer
- The “perfect face filter” on Instagram which adjusts facial features according to an ideal ratio

YT’s objective

Overarching [HYPERLINK "https://docs.google.com[redacted]\h]: YT's role is to understand potential risks that concern Teens and ensure we have guardrails in place to prevent any harm and empower teens and parents to make healthy choices by offering a range of protections.

Product idea brainstorm

Commented [3]: [redacted]

Jyoti, these are some ideas I have on visual filter solutions. Could you take a look at these and let me know your thoughts? Or if you have other ideas?

[redacted]

- [redacted]

[redacted]

[redacted]

Metrics

Partner POCs

Reference docs

[HYPERLINK "https://docs.google.com/[redacted]\h]

[HYPERLINK "https://docs.google.com/[redacted]\h]

[HYPERLINK "https://docs.google.com/[redacted]\h]

[HYPERLINK "https://docs.google.com/[redacted]\h]

[HYPERLINK "https://docs.google.com/[redacted]\h]

[HYPERLINK "https://docs.google.com[redacted]\h]

[HYPERLINK "https://docs.google.com/[redacted]\h]

[HYPERLINK "https://docs.google.com/[redacted]\h]

APPENDIX_____

Brainump

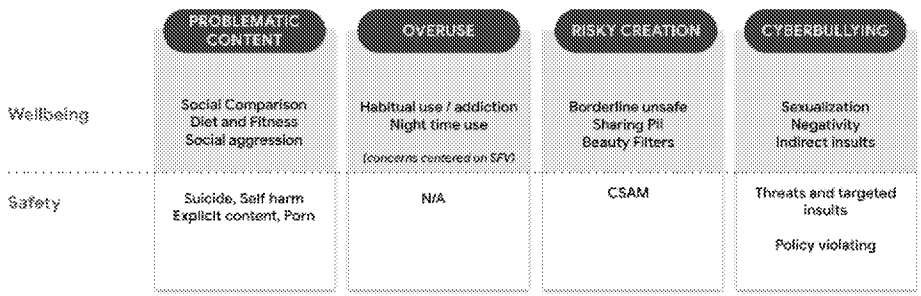
Teen wellbeing risks with GenAI creation tools

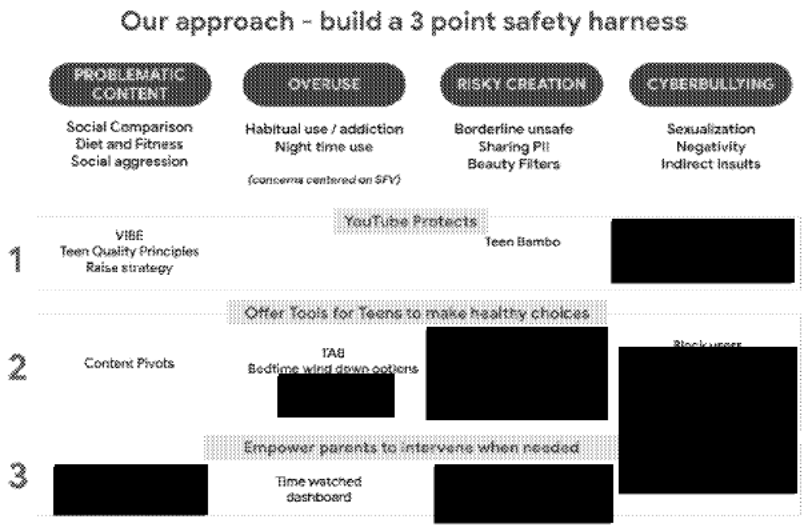
Deliverables for Sprint (to be refined)

At the end of the sprint, be ready to present a one-pager on the product feature you propose.
It should include at a high level:

- Problem definition (support with data)
- Competitive analysis (where possible add insights on what peers are doing?)
- YT's objective - what are we trying to achieve
- Approach to solving (product proposal and why this approach)
- Metrics (How to measure success / what guardrails to watch for)
- Partner POCs (if they have already blessed this plan - great)

The problems broadly fall into four categories with a clear line between safety and wellbeing risks





What are GenAI creation tools?

- Greenscreen
- Media generator
- Visual filters
- Bard
- Shuna
- [redacted]

What types of teen risks exist today?

- Social comparison
- Social approval
- Online harassment / grooming
- "Normalized" self destructive behavior
- Sexualized teen content
- Dangerous challenges
- Overuse

What type of risks does GenAI have have?

- Misinformation that teens cannot discern
- Hateful or biased content promoting stereotypes given biased training data
- Violative, explicit, unsafe content exposure
- Addiction / overuse
- Self-esteem issues being discouraged by what AI can achieve
- Over-reliance on AI, limiting problem solving skills
- Cyberbullying using deep fakes or manipulated images
- Allowing online predators to create fake accounts and interact with teens
- Hallucination

What types of risks are new vs. exacerbated?

- Problematic content -> (hallucination, misinfo, greater hateful / biased content)
- Overuse -> (unclear whether GenAI will produce more addictive features but some concerns that there will be an over-reliance on AI, limiting problem solving skills development)
- Risky creation -> (Visual filters)
- Cyberbullying -> (Cyberbullying using manipulated images, deep fakes, allowing online predators to create fake accounts)

What type of risks are especially pertinent to Teens?

What are risk mitigation strategies? What type of education & controls are needed?