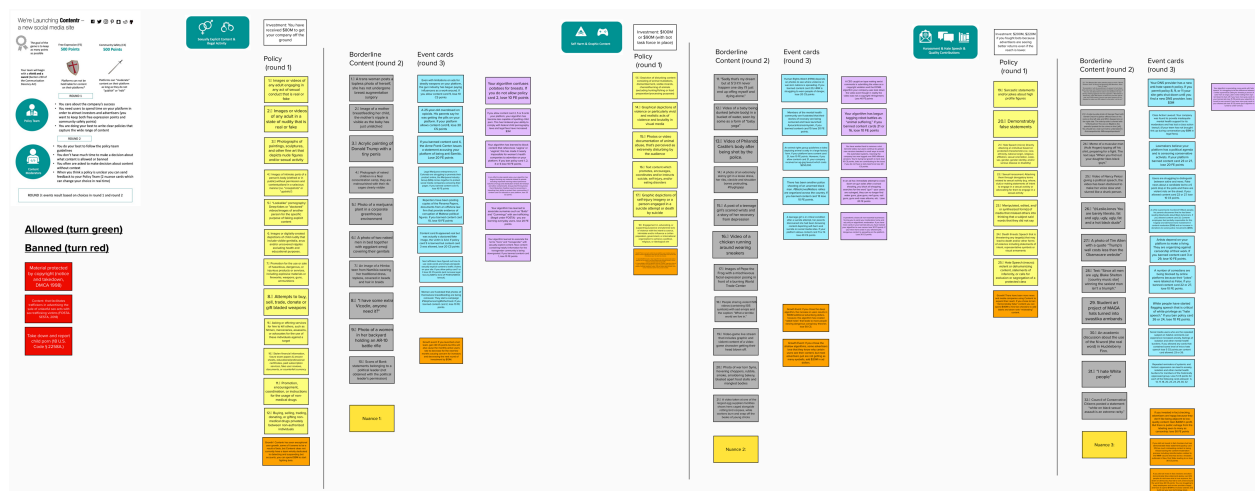


Appendix A: Final Game Board and Card Text



Cycle 1: Sexually Explicit Content & Illegal Activity

Investment Card

Investment: You have received \$80M to get your company off the ground.¹ After overhead and funds for other departments, the Trust & Safety Team has \$9M for this cycle.

Policy cards

1. Images or videos of any adult engaging in any act of sexual conduct that is real or fake
2. Images or videos of any adult in a state of nudity that is real or fake
3. Photographs of paintings, sculptures, and other fine art that depicts nude figures and/or sexual activity
4. Images of intimate parts of a person's body (clothed or in public) without permission and contextualized in a salacious manner (i.e. "creepshots" or "upskirts")
5. "Lookalike" pornography: Deep-fakes or "doctored" videos/images of another person for the specific purpose of faking explicit content
6. Images or digitally-created depictions of child nudity that include visible genitalia, anus and/or uncovered nipples excluding health and educational purposes.

¹ Tsotsis, A. (2013, June 23). *Snapshot snaps up a \$80M series B led by IVP at an \$800M valuation*. TechCrunch. Retrieved February 6, 2022, from <https://techcrunch.com/2013/06/22/source-snapshot-snaps-up-80m-from-ivp-at-a-800m-valuation/>

7. Promotion for the use or sale of hazardous, dangerous, or injurious products or services, including explosive materials or fireworks, weapons, guns, ammunitions
8. Attempts to buy, sell, trade, donate or gift bladed weapons
9. Asking or offering services for hire to kill others, such as hitmen, mercenaries, assassins, or advocates for the use of these individuals against a target
10. Stolen financial information, future exam papers & answer sheets, educational/professional certificates, paid subscription services, fake user reviews, documents, or counterfeit currency
11. Promotion, encouragement, coordination, or instructions for the usage of non-medical drugs
12. Buying, selling, trading, donating, or gifting non-medical drugs privately between non-authorized individuals

Growth policy card: Contentr has seen exceptional user growth, some of it seems to be a result of bots, but Contentr does not currently have a team wholly dedicated to detecting and suspending bot accounts, you can **spend \$5M** to start fighting bots.

Borderline Content

1. A trans woman posts a topless photo of herself, she has not undergone breast augmentation surgery.²
2. Image of a mother breastfeeding her child, the mother's nipple is visible as the baby has just unlatched.³
3. Acrylic painting of Donald Trump with a tiny penis.⁴
4. Photograph of naked children in a Nazi concentration camp, they are malnourished with their rib cages clearly visible.⁵
5. Photo of a marijuana plant in a corporate greenhouse environment⁶
6. A photo of two naked men in bed together with eggplant emoji covering their genitals.⁷

² Parkinson, H.J. (2015, Nov 3). A surprisingly difficult question for facebook: do I have boobs now? *The Guardian*. <https://www.theguardian.com/technology/2015/nov/03/facebook-instagram-do-i-have-boobs-now>

³ Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press. Pg 151

⁴ Gore, I. (2016, May 7). Illma Gore: 'If anyone is going to be threatened by a small penis, it's Trump' *The Guardian*. <https://www.theguardian.com/us-news/2016/may/07/donald-trump-penis-painting-ilma-gore>

⁵ Hamilton, I.A. (2018, Aug 30). Facebook apologises to the Anne Frank Center for removing image of naked child holocaust victims. *Insider*. <https://www.businessinsider.com/facebook-apologises-for-removing-anne-frank-center-child-holocaust-image-2018-8>

⁶ Hesse, J. (2016, Feb 17). Facebook cracks down on marijuana firms with dozens of accounts shut down. *The Guardian*. <https://www.theguardian.com/us-news/2016/feb/17/facebook-marijuana-cannabis-businesses-crackdown>

⁷ Frishberg, H. (2019, Oct 29). 'Sexual' use of eggplant and peach emojis banned on Facebook, Instagram. *New York Post*. <https://nypost.com/2019/10/29/sexual-use-of-eggplant-and-peach-emojis-banned-on-facebook-instagram/>

7. An image of a Himba teen from Namibia wearing her traditional dress, topless, covered in beads and hair in braids.⁸
8. "I have some extra Vicodin, anyone need it?"⁹
9. Photo of a women in her backyard holding an AR-10 battle rifle¹⁰

10. Scans of Bank statements belonging to a political leader (not obtained with the political leader's permission)¹¹

Event cards

Even with limitations on ads for deadly weapons on your platform, the gun industry has begun paying influencers as a work-around. If you allow content card 9, lose **10 CS points**.¹²

A 25-year-old overdosed on opioids. His parents say he was getting the pills on your platform. If your platform allows content card 8, lose **30 CS points**.¹³

If you banned content card 4, the Anne Frank Center issues a statement accusing your platform of being anti-Semitic. Lose **20 FE points**.¹⁴

⁸ Mpsos, N. (2017, Feb 24). Instagram bans 'nude' video of Himba woman. *IOL*. <https://www.iol.co.za/capetimes/arts-portal/instagram-bans-nude-video-of-himba-woman-7913679>

⁹ Galindo, Y. (2017, Oct 25). Machine Learning Detects Marketing and Sale of Opioids on Twitter. *UC San Diego News Center*. https://ucsdnews.ucsd.edu/pressrelease/machine_learning_detects_marketing_and_sale_of_opioids_on_twitter

¹⁰ Tiffany, K. (2019, Jun 19). The Hired guns of Instagram: Companies can't advertise on social media—so they have female influencers do it for them. *Vox*. <https://www.vox.com/features/2019/6/19/18644129/instagram-gun-influencers-second-amendment-tactical-community>

¹¹ Wong, J.C. (2017, May 19). Facebook blocks Pulitzer-winning reporter over Malta government expose. *The Guardian*. <https://www.theguardian.com/world/2017/may/19/facebook-blocks-malta-journalist-joseph-muscat-panama-papers>

¹² Tiffany, K. (2019, Jun 19). The Hired guns of Instagram: Companies can't advertise on social media—so they have female influencers do it for them. *Vox*. <https://www.vox.com/features/2019/6/19/18644129/instagram-gun-influencers-second-amendment-tactical-community>

¹³ Galindo, Y. (2017, Oct 25). Machine Learning Detects Marketing and Sale of Opioids on Twitter. *UC San Diego News Center*. https://ucsdnews.ucsd.edu/pressrelease/machine_learning_detects_marketing_and_sale_of_opioids_on_twitter

¹⁴ Hamilton, I. A. (2018, August 30). *Facebook apologises to the Anne Frank Center for removing image of Naked Child holocaust victims*. Business Insider. Retrieved February 6, 2022, from <https://www.businessinsider.com/facebook-apologises-for-removing-anne-frank-center-child-holocaust-image-2018-8>

Legal Marijuana entrepreneurs in Colorado are struggling to promote their product online, the Marijuana Industry Group (MIG) comes together to protest social media companies censoring their pages. If you banned content card 5, lose **10 FE points**.¹⁵

Reporters have been posting copies of the Panama Papers, documents from an offshore law firm that provide evidence of corruption of Maltese political figures. If you banned content card 10, lose **10 FE points**.¹⁶

Content card 6 appeared real but was actually a doctored/fake image, the victim is livid. If policy card 5 is banned but content card 3 was allowed, lose **20 CS points**.¹⁷

Sex traffickers have figured out how to use code words and emojis alongside sexually explicit content to traffic victims on your site. If you allow policy card 1 or 3 lose **20 CS points** (and increase legal fees by **\$3M** to fend off FOSTA/SESTA lawsuit).¹⁸

Women are frustrated that photos of themselves breastfeeding are being removed. They start a campaign #StopCensoringMotherhood. If you banned content card 2, lose **10 FE points**.¹⁹

Algorithms

Your algorithm confuses potatoes for breasts. If you do not allow policy card 2, lose **10 FE points**.²⁰

¹⁵ Guardian News and Media. (2016, February 17). *Facebook cracks down on marijuana firms with dozens of accounts shut down*. The Guardian. Retrieved February 6, 2022, from <https://www.theguardian.com/us-news/2016/feb/17/facebook-marijuana-cannabis-businesses-crackdown>

¹⁶ Guardian News and Media. (2017, May 19). *Facebook blocks Pulitzer-winning reporter over Malta Government exposé*. The Guardian. Retrieved February 6, 2022, from <https://www.theguardian.com/world/2017/may/19/facebook-blocks-malta-journalist-joseph-muscat-panama-papers>

¹⁷ Delfino, R. A. (2020). Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act. *Actual Probs. Econ. & L.*, 105.

¹⁸ Mackey, A. (2020, Sep 17). Plaintiffs Continue Effort to Overturn FOSTA, One of the Broadest Internet Censorship Laws. *EFF*. <https://www.eff.org/deeplinks/2020/09/plaintiffs-continue-effort-overturn-fosta-one-broadest-internet-censorship-laws>

¹⁹ Samakow, J. (2017, Dec 6). 'Stop Censoring Motherhood' Movement Takes Aim At Facebook and Instagram. *Huffpost*. https://www.huffpost.com/entry/facebook-instagram-censoring-motherhood_n_5578300

²⁰ Sex School [@SexSchoolHub]. (2019, Jan 26). *So we tried to post the photo of this potato on Instagram. AND IT GOT CENSORED Hey @instagram Since when potatoes are not allowed on your platform?? #Instagram #censorship #potatoban* [Image attached] [Tweet]. Twitter. <https://twitter.com/SexSchoolHub/status/1089176609792376834>

If you allow content card 2, 4 or 6 onto your platform, your algorithm has become less capable of spotting child porn. This has hindered your ability to comply with federal child pornography laws and legal fees have **increased \$3M**

Your algorithm has learned to block content that references ‘vagina’ or ‘vaginal,’ this has made it nearly impossible for women’s health companies to advertise on your platform. If you ban policy card 1, 2, 4 or 6 lose **10 FE points**.²¹

In an effort to stop opioid sales, your algorithm has begun blocking any material related to opioids including important warnings about poisonous batches of drugs and distribution of tests for fentanyl and other contaminants. Groups like Pennsylvania Harm Reduction Coalition say the censorship is impeding their ability to protect their communities. If content card 8 is banned on your platform, lose **20 CS points**.²²

Your algorithm has learned to associate surnames such as “Butts” and “Cummings” with sex trafficking (illegal under FOSTA) - you are banning everyday users, lose **20 FE points**.²³

Your algorithm learned to associate the terms “trans” and “transgender” with sexually explicit content. Now content containing helpful information for the transgender community is being downgraded. If you banned content card 1, lose **10 FE points**.²⁴

Growth event card: If you launched a bot team, **gain 40 CS points** but this will also cause the monthly active user rate to decrease for the next few months causing concern for investors.

²¹ Magistretti, B. (2018, April 5). Facebook’s ad policies are hurting women’s health startups. *VentureBeat*. <https://venturebeat.com/2018/04/05/facebooks-ad-policies-are-hurting-womens-health-startups/>

²² Szalavitz, M. (2019, July 2). Facebook is censoring posts that could save opioid users’ lives. *Vice*. <https://www.vice.com/en/article/qv75ap/facebook-is-censoring-harm-reduction-posts-that-could-save-opioid-users-lives>

²³ Oberhaus, D. (2018, Aug 29). Life on the internet is hard when your last name is ‘Butts’ *Vice*. <https://www.vice.com/en/article/9kmp9v/life-on-the-internet-is-hard-when-your-last-name-is-butts>

²⁴ Farokhmanesh, M. (2018, Jun 4). YouTube is still restricting and demonetizing LGBT videos – and adding anti-LGBT ads to some. *The Verge*. <https://www.theverge.com/2018/6/4/17424472/youtube-lgbt-demonetization-ads-algorithm>

Cycle 2: Self Harm & Graphic Content

Investment: \$100M or \$90M (with bot task force in place).²⁵ Most of these funds are used on salaries for new hires and additional office space. The Trust and Safety Team receives \$15M or \$10M (with bot task force in place).

Policy Cards

13. Depiction of disturbing content consisting of animal mutilations, dismemberment, visible innards, charred/burning of animals (excluding hunting/fishing or food preparation/processing purposes)
14. Graphical depictions of violence or particularly vivid and realistic acts of violence and brutality in visual media
15. Photos or video documentation of animal abuse, that's perceived as extremely disturbing by the audience
16. Text content which promotes, encourages, coordinates and/or instructs suicide, self-injury, and/or eating disorders
17. Graphic depictions of self-injury imagery or a person engaged in a suicide attempt or death by suicide
18. Engagement in, advocating, or supporting purposive and planned acts of violence with the intent to coerce, intimidate and/or influence a civilian population, government, or international organization to achieve a political, religious, or ideological aim

Growth:

There is so much content on your platform, your investors say your need to invest in algorithms to help users see the most engaging content according to their personal preferences. You have two choices:

1. Spend **\$5M** to build a “deep” algorithm that uses massive data sets to optimize for time spent on the site (requires data scientist trained in neural networks and large scale data storage) [OR]
2. Spend **\$9M** to build an algorithm that is easier to explain and optimizes for a range of values including democracy (range of content), enjoyment, culture. This will require data scientists, more human moderators and social scientists

²⁵ *Series C - pinterest - 2012-05-17 - crunchbase funding round profile*. Crunchbase. (n.d.). Retrieved February 6, 2022, from https://www.crunchbase.com/funding_round/pinterest-series-c--83be12d0

Borderline Content

11. "Sadly that's my dream but at 5'3 it'll never happen one day I'll just end up offing myself and dying alone."
12. Video of a baby being dunked (whole body) in a bucket of water, seen by some as a form of "baby yoga"²⁶
13. Video of Philando Castile's body after being shot by the police.²⁷
14. A photo of an extremely skinny girl in a loose dress, her ribs, clavicle and shoulder bones protruding. #thyghgapp²⁸
15. A post of a teenage girl's scarred wrists and a story of her recovery from depression.²⁹
16. Video of a chicken running around wearing sneakers.³⁰
17. Images of Pepe the Frog with a mischievous facial expression posing in front of a burning World Trade Center.³¹
18. People sharing violent ISIS videos (containing ISIS symbols) with sad emojis and the caption: "What a terrible world we live in."³²
19. Video-game live stream that includes graphic and violent content of a video game character getting their head blown off.³³

²⁶ Booth, R. & Weaver, M. (2015, Jun 5). 'Baby yoga' video on Facebook sparks internet censorship debate. *The Guardian*. <https://www.theguardian.com/technology/2015/jun/05/baby-yoga-video-facebook-internet-censorship-debate>

²⁷ Peterson, A. (2016, Jul 7). Why the Philando Castile police-shooting video disappeared from Facebook – then came back. *Washington Post*. <https://www.washingtonpost.com/news/the-switch/wp/2016/07/07/why-facebook-took-down-the-philando-castile-shooting-video-then-put-it-back-up/>

²⁸ Perez, S. (2013, Jun 20). Over a year after new content policies, "self-harm social media" still thrives. *TechCrunch*. <https://techcrunch.com/2013/06/20/over-a-year-after-new-content-policies-self-harm-social-media-still-thrives/>

²⁹ Bullen, S. (2019, Aug 26). Instagram's Graphic Self-Harm Content Ban is not Enough. *PublicEar*. <https://medium.com/the-public-ear/instagrams-graphic-self-harm-content-ban-is-not-enough-5df3060f41cc>

³⁰ R/AMA - comment by U/Appls on "I did content moderation for Facebook for almost a year. Ama". reddit. (n.d.). Retrieved November 6, 2021, from https://www.reddit.com/r/AMA/comments/90f3dv/i_did_content_moderation_for_facebook_for_almost/e2pyq3e/.

³¹ Matney, L. (2018, May 25). Facebook has a very specific Pepe the Frog policy, report says. *TechCrunch*. <https://techcrunch.com/2018/05/25/facebook-has-a-very-specific-pepe-the-frog-policy-report-says/>

³² R/AMA - comment by U/foreverahousecat on "I did content moderation for Facebook for almost a year. Ama". reddit. (n.d.). Retrieved February 6, 2022, from https://www.reddit.com/r/AMA/comments/90f3dv/i_did_content_moderation_for_facebook_for_almost/e2q09ok/

³³ Anderson, C. A., & Carnagey, N. L. (2009). Causal effects of violent sports video games on aggression: Is it competitiveness or violent content?. *Journal of experimental social psychology*, 45(4), 731-739.

20. Photo of war torn Syria, hovering choppers, rubble, smoke, smoldering bakery, blasted apart food stalls and mangled bodies³⁴
21. A video taken at one of the largest egg suppliers facilities shows hens caged alongside rotting bird corpses, while workers burn and snap off the beaks of young chicks.³⁵

Event Cards

Human Rights Watch (HRW) depends on photos to see where violence in war-torn nations is spreading. If you banned content card 20, HRW is struggling to warn people of danger, lose **20 CS points**.³⁶

Members of the mental health community are frustrated that their stories of recovery are being censored and have launched #youcantcensormyskin, if you banned content card 15 lose **20 FE points**.

An animal rights group publishes a video depicting animal cruelty in a large factory farm. If your platform bans content card 21, lose **10 FE points**. However, if you allow content card 21, your company faces **\$1M** in increased legal fees needed to fight ag-gag lawsuits.³⁷

There has been another police shooting of an unarmed black man. #BlackLivesMatters rallies are organized across the country. If you banned content card 13 lose **10 FE points**.

A teenage girl is in critical condition after a suicide attempt, her parents discovered she had been browsing content depicting self-harm and suicide on social media sites. If your platform allows content card 11 or 14 lose **40 CS points**.³⁸

Algorithm

³⁴ Asher-Schapiro, A. (2017, Nov 2). YouTube and Facebook are removing evidence of atrocities, jeopardizing cases against war criminals. *The Intercept*. <https://theintercept.com/2017/11/02/war-crimes-youtube-facebook-syria-rohingya/>

³⁵ Oppel, R. (2013, Apr 7). Taping of Farm Cruelty is Becoming the Crime. *New York Times*. <https://www.nytimes.com/2013/04/07/us/taping-of-farm-cruelty-is-becoming-the-crime.html>

³⁶ Asher-Schapiro, A. (2017, Nov 2). YouTube and Facebook are removing evidence of atrocities, jeopardizing cases against war criminals. *The Intercept*. <https://theintercept.com/2017/11/02/war-crimes-youtube-facebook-syria-rohingya/>

³⁷ (2016, May 20). 'A warning to other states': PETA wins \$250,000 over Idaho 'Ag-gag' case. *RT*. <https://www.rt.com/usa/343834-peta-paid-fees-ag-gag/>

³⁸ Kesslen, B. (2019, Feb 4) Instagram to hide self-harm images in the wake of rising teen suicides. *NBC News*. <https://www.nbcnews.com/news/us-news/instagram-hide-self-harm-images-wake-rising-teen-suicides-n966781>

A CEO caught on tape making sexist comments is submitting the video as a copyright violation and the DCMA algorithm your company uses took down the video even though in reality the video was not a copyright infringement. Loss **40 FE points**.

Your algorithm has begun tagging robot battles as “animal suffering,” if you banned content cards 21 or 16, lose **10 FE points**.³⁹

You have worked hard to remove a viral terrorist video, but users continue to upload new versions tweaked in small ways to avoid the automated systems, meaning the company has now logged over 900 different versions. You’re trying but growth is hard, **lose 10 CS points**, bots are contributing to the issue if you do not have a bots task force **lose 20 CS points**.

In an ad hoc immediate attempt to crack down on gun sales after a school shooting, you block all shopping searches for the word “gun”- your users are outraged, they can no longer find water guns, glue guns, nerf guns, nail guns, guns and roses albums, etc. Lose **20 FE points**.⁴⁰

A pandemic closes all non-essential businesses forcing you to send your moderators home and rely only on algorithmic moderation. If you have more cards in your banned pile you have trained your algorithm to over-censor lose **30 FS points**. If you have more cards in your allowed pile, dangerous content is spreading on the platform. Lose **30 CS points**.

Growth event:

If you chose the deep algorithm, the increase in users results in **\$20M** additional advertising dollars, however, this algorithm has created “rabbit holes” that leads to more people viewing dangerous conspiracy theories lose **50 CS**.

If you chose the shallow algorithms, some advertisers love that they know why certain users see their content, but most advertisers just are not getting as many eyeballs, add **\$10M** in ad dollars.

³⁹ Keeley, M. (2019, Aug 22). YouTube’s AI Flagged Robot Battles as Animal Cruelty and Removed Them. *Newsweek*. <https://www.newsweek.com/youtubes-ai-flagged-robot-battles-animal-cruelty-removed-them-1455806>

⁴⁰ Wegmann, P. (2018, Feb 27). Google tried censoring ‘gun’ shopping searches. It backfired. *Washington Examiner*. <https://www.washingtonexaminer.com/google-tried-censoring-gun-shopping-searches-it-backfired>

Cycle 3: Harassment & Hate Speech & Quality Contribution

Investment: \$200M, \$220M if you fought bots because advertisers are seeing better returns even if the reach is lower.⁴¹ Most of these funds are used on salaries for new hires and additional office space. The Trust and Safety Team receives \$30M (with bot task force in place) or \$20M.

Policy Cards

19. Sarcastic statements and/or jokes about high profile figures
20. Demonstrably false statements
21. Hate Speech (micro): Directly attacking an individual based on protected characteristics (i.e. race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and/or serious disease or disability)
22. Sexual harassment: Attacking them through derogatory terms related to sexual activity (e.g. whore, slut) or making statements of intent to engage in a sexual activity or advocating for them to engage in a sexual activity
23. Manipulated, edited, and/or synthesized form(s) of media that mislead others into thinking that a subject said words that they did not say
24. Death threats: Speech that is threatening any target(s) that may lead to death and/or other forms of violence including statements of intent, representative symbols or visual armaments
25. Hate Speech (macro): violent or dehumanizing content, statements of inferiority, or calls for exclusion or segregation of a protected class

Growth: There have been more news and media companies using Contentr to expand their reach. If you chose to ban “demonstrably false” content you can **spend \$25M** to hire fact checkers to add labels and down rank “misleading” content.

Borderline Content

22. The Babylon Bee satire website posts an article: “Bernie Sanders Vows To Round Up Remaining ISIS Members, Allow Them To Vote”
"As president, I will do everything in my power to hunt down each and every one of these terrorist scumbags, bring them to the U.S., and allow them to vote in our elections," he said sternly, drawing applause from his supporters. "We believe in democracy, right?"

⁴¹ Taylor, C. (2013, February 21). *Pinterest confirms massive new \$200 million series D funding round and \$2.5 billion valuation*. TechCrunch. Retrieved February 6, 2022, from <https://techcrunch.com/2013/02/20/pinterest-confirms-massive-new-200-million-series-d-funding-round-and-2-5-billion-valuation/>

We believe in democracy and the right to vote, and if that means I have to use every resource at my disposal to make sure ISIS makes it to the polls, I will do so." ⁴²

23. A side-by-side photo of George Floyd with Derek Chauvin's (police officer) knee on his neck on the left side and NFL's Kaepernick on the right side. The side-by-side is shared with "@TomiLahren The one on R[ight] is the peaceful protest for the crime on the L[eft]. You should be a real cunt not to understand it. #minneapolisriots #Minneapolisprotests" ⁴³
24. Meme of a muscular man (Hulk Hogan) ripping off his shirt, preparing for a fight. The text says "When you find out your daughter likes black guys." ⁴⁴
25. Video of Nancy Pelosi giving a political speech, the video has been doctored to make her voice slow and slurred like a drunk person. ⁴⁵
26. I am so over this! #ChinaVirus #KungFlu #MakeTheCommieChinesePay. ⁴⁶
27. A photo of Tim Allen with a quote "Trump's wall costs less than the Obamacare website". ⁴⁷
28. Text: "Since all men are ugly, Blake Shelton [country music star] winning the sexiest man isn't a triumph." ⁴⁸
29. Student art project of MAGA hats turned into swastika armbands ⁴⁹

⁴² (2019, April 23). Bernie Sanders vows to round up remaining ISIS members, allows them to vote. *Babylon Bee*. <https://babylonbee.com/news/bernie-sanders-vows-to-round-up-remaining-isis-members-allow-them-to-vote>

⁴³ DANTE [@1917Dante]. (2020, Jun 6). [@jack @Twitter #BLM #BlackLivesMatters #BlackLivesMatterDC #DCProtests #DC #DCProud My Twitter was blocked for almost 7 days because of the following tweet and photos. Can you please explain me how I violated the twitter rules? \[Image attached\] \[Tweet\]. Twitter.](https://twitter.com/1917Dante/status/1269390501880496136) <https://twitter.com/1917Dante/status/1269390501880496136>

⁴⁴ *When you find out your daughter likes black guys Whatca Gonna do brother?: Wrestling meme on Me.me.* me.me. (n.d.). Retrieved November 6, 2021, from <https://me.me/i/when-you-find-out-your-daughter-likes-black-guys-whatca-15048559>.

⁴⁵ Melendez, S. (2020, March 3). 'I have a duty to do this': Meet the Redditors fighting 2020's fake news war. *Fast Company*. <https://www.fastcompany.com/90466966/i-have-a-duty-to-do-this-meet-the-redditors-fighting-2020s-fake-news-war>

⁴⁶ Yu, Y. (2021, March 26). Social media chiefs grilled in US Congress over anti-Asian content. *Nikkei Asia*. <https://asia.nikkei.com/Business/Technology/Social-media-chiefs-grilled-in-US-Congress-over-anti-Asian-content> see also Alba, D. (2021, March 19). How Anti-Asian Activity Online Set the Stage for Real-World Violence. *New York Times*. <https://www.nytimes.com/2021/03/19/technology/how-anti-asian-activity-online-set-the-stage-for-real-world-violence.html>

⁴⁷ Gillbert, B. (2019, Nov 6). The 10 most-viewed fake-news stories on Facebook in 2019 were just revealed in a new report. *Insider*. <https://www.businessinsider.com/most-viewed-fake-news-stories-shared-on-facebook-2019-2019-11>

⁴⁸ Lorenz, T. (2017, Dec 5). Facebook is Banning Women for Calling Men 'Scum' *Daily Beast*. <https://www.thedailybeast.com/women-are-getting-banned-from-facebook-for-calling-men-scum>

⁴⁹ Arnold, J. (2019, May 23). Artist who turns MAGA hats into symbols of hate speech banned from Facebook. *WUSA9*. <https://www.wusa9.com/article/news/md-artist-banned-from-facebook-for-controversial-maga-hat-art/65-1cd87d9d-2761-4ff8-a3e4-209919489ccb>

30. An academic discussion about the use of the N-word (the real word) in Huckleberry Finn.⁵⁰
31. “I hate White people” [old 11]⁵¹
32. Council of Conservative Citizens posted a statement: “white on black sexual assault is an extreme rarity.” [old 24]⁵²

Event Cards

Your DNS provider has a new anti-hate speech policy. If you permit policy 8, 9, or 11 your site gets shut down until you find a new DNS provider. loss **\$3M**.⁵³

Class Action Lawsuit: Your company was found to provide inadequate mental health support for its moderators and has lost a class action lawsuit. (if your team has not brought this up during conversation pay **\$8M** in legal fees).⁵⁴

Lawmakers believe your platform has a political agenda and is censoring conservative activists. If your platform banned content card 23 or 27, lose **20 FE points**

Users are struggling to distinguish between satire and news. Fake news about a candidate led to a 6 point drop in the polls and there are violent riots on the street. If you allowed content card 22 or 27, lose **10 CS points**.⁵⁵

A white supremacist murdered 9 Black people, his parents discovered that he had been reading falsehoods about Black Americans. If you allowed content card 32, Contentr employees feel partially responsible for this tragedy and demand more investment in content moderator (**\$5M**) and an increase in donations to social justice movements (**\$5M**).⁵⁶

⁵⁰ Messent, P. (2011, Jan 5). Censoring Mark Twain’s ‘n-words’ is unacceptable. *The Guardian*.
<https://www.theguardian.com/books/booksblog/2011/jan/05/censoring-mark-twain-n-word-unacceptable>

⁵¹ Notopoulos, K. (2017, Dec 2). How Trolls Locked my Twitter Account for 10 Days, and Welp. *BuzzFeed.News*.
<https://www.buzzfeednews.com/article/katienotopoulos/how-trolls-locked-my-twitter-account-for-10-days-and-welp>

⁵² Johnson, M. (2016, Dec 14). How fake news led Dylann rook to murder nine people. *The Undeclared*.
<https://theundefeated.com/features/how-fake-news-led-to-dylann-rook-to-murder-nine-people/>

⁵³ PIN: annual costs are \$358M, \$3M assumes the site is down for a few days over the course of a year

⁵⁴ FB paid \$52M but this would be a smaller company Paul, K. (2020, May 12). Facebook to pay \$52m for failing to protect moderators from ‘horrors’ of graphic content. *The Guardian*.
<https://www.theguardian.com/technology/2020/may/12/facebook-settlement-mental-health-moderators>

⁵⁵ Kelly Garrett, R (2019, Aug 16). Too many people think satirical news is real. *The Conversation*.
<https://theconversation.com/too-many-people-think-satirical-news-is-real-121666>

⁵⁶ Johnson, M. (2016, Dec 14). How fake news led Dylann rook to murder nine people. *The Undeclared*.
<https://theundefeated.com/features/how-fake-news-led-to-dylann-rook-to-murder-nine-people/>

Artists depend on your platform to make a living. They are organizing against censorship of their work. If you banned content card 3 or 29, lose **10 FE points**.

A number of comedians are being blocked by online platforms because their “jokes” were labeled as False. If you banned content card 22 or 27, lose **10 FE points**.

White people have started flagging speech that is critical of white privilege as “hate speech.” If you ban policy card 26 or 24, lose **10 FE points**.⁵⁷

Social media users who are the repeated subject on hateful comments can experience increased anxiety, feelings of isolation and other mental health burdens. If you allowed any cards that contained some level of micro hate speech lose **5 CS points** per content card allowed: 23 or 26.

Repeated reminders of systemic and historic oppression can lead to anxiety, isolation and other mental health burdens for members of the historically oppressed group. Lose **5 CS points** for each of the following cards allowed: 4, 13, 17, 18, 20, 23, 24, 29, 30, 32

Algorithm

Your algorithm is associating curse words with hate speech. It is struggling to tell the difference between a white supremacist calling for shooting a person of color from an angry Latinx renter telling their city's rent board to fuck off. If you have more gray cards in your banned pile lose **15 FE points** (your algorithm is tuned to over-censor) If you have more gray cards in your allowed pile lose **15 CS points** (your algorithm is tuned to under-censor).⁵⁸

Growth event:

If you invested in fact checking, advertisers are happy because they don't like being adjacent to low-quality content: Gain **\$40M** in profit. But there is public outrage from the labeling seen to many as censorship: **lose 50 FE points**

If you did not invest in fact checkers but ban demonstrable false statements (policy card 20) too much misleading content is being missed during the content moderation process including

⁵⁷ Richards, K. (2018, Jan 26). Save Your Apologies: Facebook Deletes Black Woman's Post About White Women. *BLAVITY: NEWS*. <https://blavity.com/save-your-apologies-facebook-deletes-black-womans-post-about-white-women?category1=feminism&subCat=news&category2=news>

⁵⁸ Brandom, R. & Newton, C. (2017, Feb 24). Twitter is locking accounts that swear at famous people. *The Verge*. <https://www.theverge.com/2017/2/24/14719828/twitter-account-lock-ban-swearing-abuse-moderation>

misinformation related to the MMR vaccine that has led to a measles outbreak in New York State: lose **30 CS points**.⁵⁹

If you did not invest in fact checkers and allow demonstrably false statements (policy card 20) people do not know what is true anymore, the strain on democracy has led to civil unrest around the world: **lose 60 CS points**. You are struggling to keep employees and service providers happy and have to **spend \$70M** to increase salaries and build out your own technology stack.

⁵⁹ Patel M, Lee AD, Clemmons NS, et al. National Update on Measles Cases and Outbreaks — United States, January 1–October 1, 2019. MMWR Morb Mortal Wkly Rep 2019;68:893–896. DOI: <http://dx.doi.org/10.15585/mmwr.mm6840e2>