Re: Thanks!

abrockwehl@usaid.gov Alexander Brockwehl
To: [redacted]@fb.com Pedro Less
Cc: mescobari@usaid.gov Marcela Escobari, [redacted]@fb.com Brian Rice, [redacted]@fb.com Sandra Luff

Pedro and Sandy,

Thank you so much for the Wednesday meeting and for these follow-ups. I was just talking with Marcela and we are really looking forward to continuing the conversation with you and seeing h
expanding your network of trusted partners, connecting our missions in the LAC region with your content teams, or potentially engaging your "She Means Business" initiative with a focus on V

Let's keep the discussion going.

Have a nice weekend,

Alex

On Fri, Apr 29, 2022 at 3:02 PM Pedro Less <[redacted]@fb.com> wrote:

> Dear Assistant Administrator Escobari,
>
> It has been a great pleasure meeting you and Alex. As Sandy mentioned we will put together a comprehensive report on ongoing work in the region on the topics of common interest, future
> discuss with you and see if we can leverage the upcoming Summit of the Americas.
>
> Warm greetings from Buenos Aires, thank you for your time and wishing you a great weekend!
>
> Pedro
>
> **Pedro Less (he/him)**
> Vice President | Public Policy Latin America
> [redacted]

De: Sandra Luff <[redacted]@fb.com>
Enviado: jueves, 28 de abril de 2022 13:23
Para: Marcela Escobari <mescobari@usaid.gov>
Cc: Pedro Less <[redacted]@fb.com>; abrockwehl@usaid.gov <abrockwehl@usaid.gov>; Brian Rice <[redacted]@fb.com>
Asunto: Thanks!

Dear Assistant Administrator Escobari,

Thanks again for taking the time to meet this week. We hope you and Alex found the meeting as instructive as we did.

As discussed, I wanted share our latest Transparency Report on Community Standards Enforcement. Pedro mentioned it during our discussion and I thought it might pr
quarter report dated March 1st from Guy Rosen, our VP of integrity. Hope they help provide some greater context on our enforcement efforts & actions.

Finally, we will huddle internally on the topics we discussed and look forward to further discussions.

Many thanks again for being so generous with your time & best for a great weekend!

Sandy

Community Standards Enforcement Report, Fo

March 1, 2022
By Guy Rosen, VP of Integrity



- **Prevalence of harmful content on Facebook and Instagram remained relatively consistent — and decreased in some of our policy**

Today we're publishing our Community Standards Enforcement Report for the fourth quarter of 2021 and provides metrics on how we enforced our polici

We're also sharing:

- The Widely Viewed Content Report for the fourth quarter of 2021
- The Oversight Board Quarterly Report covering the fourth quarter of 2021
- A post on our Design Blog about our integrity design efforts to prevent misuse of our technologies and promote effective and fair enforceme

All of these reports are available in the Transparency Center.

Highlights From the Report

Prevalence of harmful content on Facebook and Instagram remained relatively consistent — and decreased in some of our policy areas — from Q3 to Q

We continued to see steady numbers on content we took action on across many areas.

On Facebook in Q4 we took action on:

- 4 million pieces of drug content, an increase from 2.7 million in Q3 2021, due to improvements made to our proactive detection technologies
- 1.5 million pieces of firearm-related content, an increase from 1.1 million in Q3, due to improved and expanded proactive detection technolo
- 1.2 billion pieces of spam content, an increase from 777 million in Q3, due to a large amount of content removed in December.

On Instagram in Q4 we took action on:

- 905K pieces of content related to terrorism, an increase from 685K pieces of content in Q3 due to improvements made to our proactive dete
- 195K pieces of firearm-related content, an increase from 154K in Q3, due to improved and expanded proactive detection technologies.

In the Q3 report, we began reporting prevalence metrics on bullying and harassment for the first time. In Q4, bullying and harassment prevalence on Inst harassment was at 0.11-0.12% in Q4, down from 0.14-0.16% in Q3. The reduction in prevalence was in part due to policy updates accounting for langua between friends.

Our Steady Improvement in Reducing Harmful Content

Our steady improvement can be attributed to a holistic approach including our policies, the technology that helps us enforce them, the operational piece

With advancements in AI technologies, like our Few Shot Learner and shift to generalized AI, we've been able to correctly take action on harmful content



Our product design teams play an important role in our effort to reduce harmful content. By carefully designing our social media products, we can promot applied, informative overlays and labels provide more context and limit exposure to misinformation or graphic content. Read more about our integrity des



Last fall, we launched Account Status on Instagram, which helps people understand why content has been removed and what policies it violates. We kno
systems by giving us indications when we've gotten things wrong. Following this launch, we saw an increase of restored content on Instagram across se
sexuality.



We're always trying to identify policies we may be under- or over-enforcing so that we can improve. For example, in 2020 during Breast Cancer Awarene
that were medical in nature. While these don't violate our policies, many were enforced on incorrectly by our systems, which led to content being taken d
needed to improve on how we enforce this type of content. The Oversight Board also heard a case related to breast cancer symptoms and nudity and as

Since then, we've been working to improve the accuracy of enforcement on health content, including content related to breast cancer and surgeries. We
with mastectomy scars and escalating content flagged as medical nudity for human review to further improve accuracy. We also send content with freque
images and avoid overenforcement in the future. As a result, we saw significantly less overenforcement during last year's Breast Cancer Awareness Mon

Integrity and Transparency Looking Forward

We're constantly looking for new ways to improve our transparency and the integrity of our platforms. For many years, we've published biannual transpar
but doesn't go against our Community Standards. We've now agreed to contribute to Lumen, an independent research project hosted by Harvard's Berkr
concerning online content. We're joining Project Lumen because it allows us to hold governments, courts, regulators and ourselves accountable for the c

We'll continue building on our progress to enforce our policies in 2022, as we're always working to improve, be transparent and share meaningful data.

Attachments:

image001.jpg 1.3M
image002.jpg 160k
image003.jpg 202k
image004.jpg 87k
Outlook-z1dwt0bc.gif 334k