

# A Case Study in Censorship

Anh V. Vu, Alice Hutchings, Ross Anderson

Legislators worldwide are mulling over options for censoring undesirable and illegal material online. We report an interesting case from late 2022 when members of a radical online forum, KIWI FARMS, harassed a Canadian trans activist so seriously that a series of tech firms tried to take the forum down. The trigger was a swatting incident on 5 August 2022. The forum was inaccessible intermittently for several weeks, and many of its members decamped to Telegram for a while. This extraordinary disruption may have been more effective than many governments could organise, yet it only cut the number of users, posts, threads and traffic towards the forum roughly by half [1].

## Datasets

Our measurements were made on several large-scale longitudinal data sources, including a complete snapshot of the KIWI FARMS forum from the EXTREMEBB dataset [2] with over 10.1M posts spanning a decade. They are supplemented with a full scrape of a Telegram channel used by forum members to circulate updates during the disruption, consisting of 525k messages, 298k replies, and 356k emojis made by 2 502 users. We also collected search interest about KIWI FARMS from Google Trends, and web traffic to its six major domains from Similarweb.

## The impact

A DDoS attack on 23 August 2022 led to a drop of roughly 35% in posting volume, then an ISP industry blackholing on 26 August 2022 silenced the forum for two consecutive days. On 3 September 2022, Cloudflare discontinued its DDoS prevention service, which attracted major publicity. This led to a seven-fold spike in global search interest in the forum, and the web traffic to its six major domains doubled on 4 September 2022 (see Figure 1).

Figure 2 shows the timeline in more detail. Forum activity dropped by around 20% after Cloudflare’s action on 3 September 2022, then rebounded as the forum was still up on *kiwifarms.ru* until DDoS-Guard took action on 5 September 2022. By 18 September, the industry had taken down all the domains (including the .onion domain by finding and killing its hosting). Activity dropped to zero for over a week. The forum operator managed to get online on 27 September 2022, and lasted until Zayo (a Tier-1 ISP) blocked it on 23 October 2022. The forum recovered a second time on 30 October 2022 and has been accessible on both the clearnet and darknet since then.

The second graph in Figure 2 shows the shift of forum activity to Telegram, which started on 27 August after the first ISP action. Telegram activity increased quickly; it was occasionally even more prominent than the forum itself, especially after the publicity given to the Cloudflare and DDoS-Guard actions. The largest displacement happened once all the domains were taken down, and again when Zayo blocked the forum’s second incarnation. The Telegram displacement appears to have been rather temporary; users returned rapidly to the forum when it became available. While there was a significant web traffic fragmentation from the primary domain to previously abandoned alternatives, the forum recovery brought back most of the traffic to the primary domain, despite being halved (Figure 3). We found no evidence of ‘external’ displacement to the major competing forum (with over 4.6M posts), instead it was ‘internal’ shifts within the KIWI FARMS ecosystem. The net effect of the concerted disruption effort over several consecutive months was that the web traffic, posting volume and number of active users were roughly halved.

## Policy implication

Deplatforming raises philosophical issues about censorship versus free speech, and also practical issues about the efficacy of private-sector versus government action. A vigorous online community may have capable defenders,

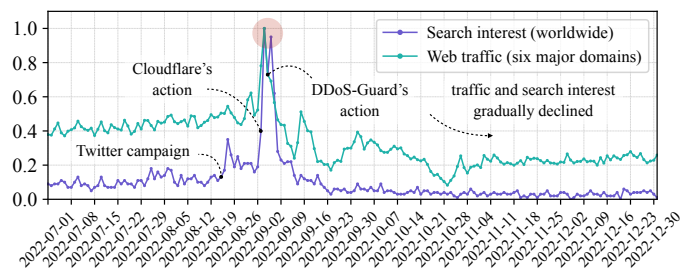


Figure 1: Normalised levels of the global search and traffic to KIWI FARMS. The red bubble marks the Streisand effect.

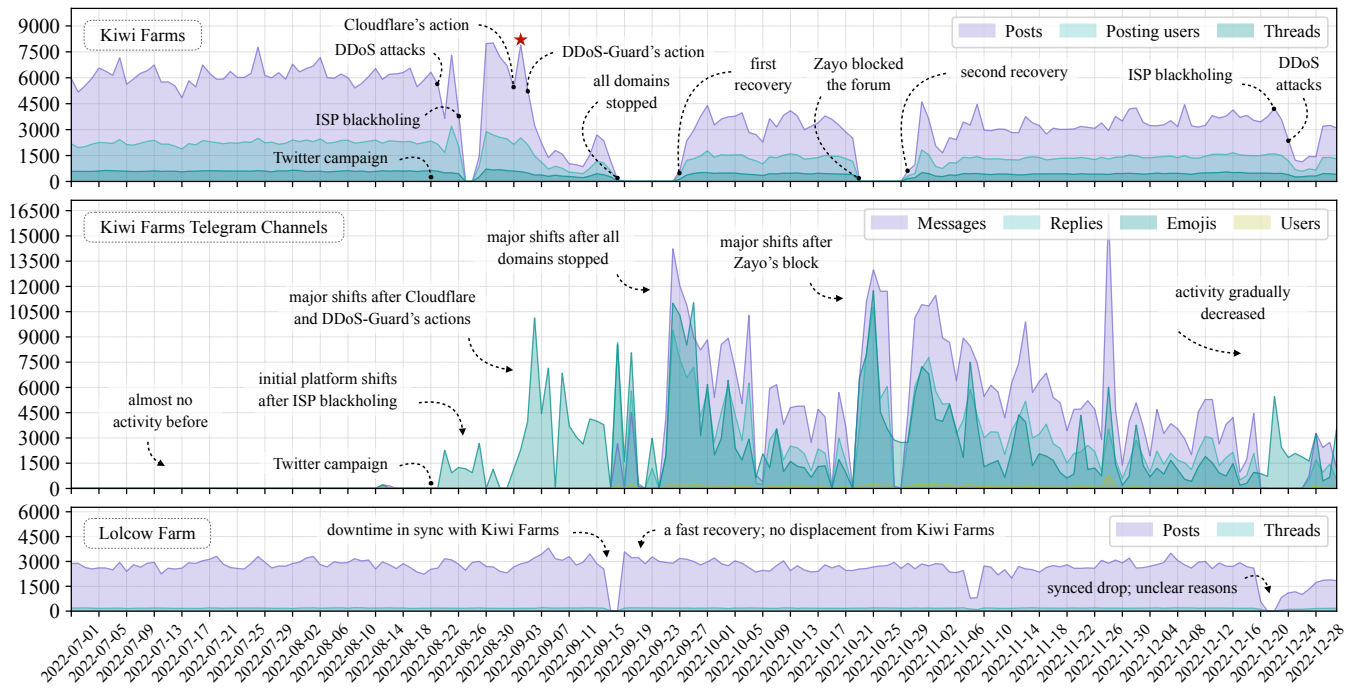


Figure 2: Daily posts, threads, and active users on KIWI FARMS, its Telegram channels, and its primary competitor LOLCOW FARM, as well as major disruptions and displacement across platforms. Red star marks the Streisand effect.

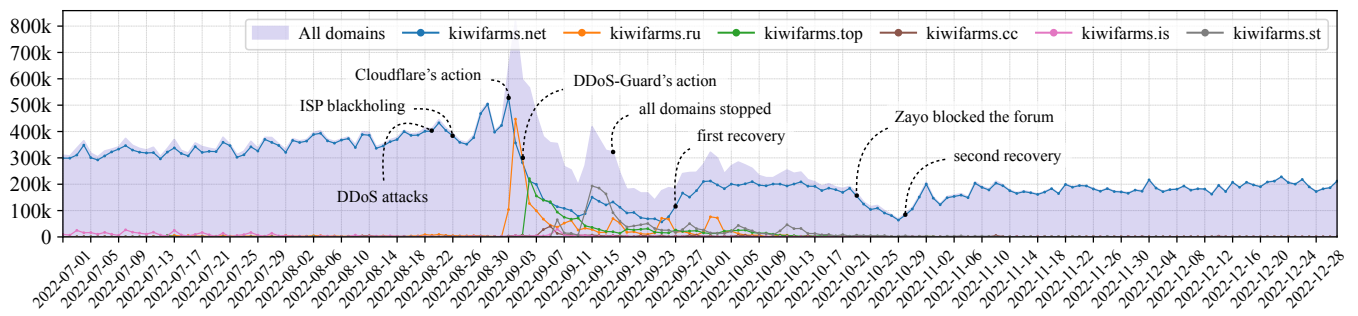


Figure 3: Number of daily estimated visits to KIWI FARMS and the fragmentation to previously abandoned domains.

whether forum operators or distributed volunteers. If some of its members break the law, are they a dissident organisation with a few bad actors, or a terrorist group that should be hunted down? Many troublesome organisations do attract hot-headed young members, including animal-rights activists and trade union organisers; whether they are labelled as terrorists or extremists is often a political matter.

Elected politicians might prefer to take these decisions than leave them to industry. On the face of it, the approach in the UK's Online Safety Bill of empowering a politically-appointed regulator to seek court orders against individual tech companies [3] might make political sense. Yet it is likely to be much less effective than a concerted industry action. Our previous work has explored why governments are less able to take down bad websites than the private sector [4]. That study related to single bad websites, such as those hosting malware, phishing lures or sex-abuse images. This study demonstrates why merely taking down an active standalone community is likely to be even harder and is not enough to deal effectively with hate and harassment, especially when the censor cannot incapacitate the key maintainers, whether by exhausting them, arresting them, enjoining them or otherwise deterring them. Even when the industry rolls its sleeves up and tries to suppress a community some of whose members have indulged in crime and against whom there is an industry consensus, the results can be modest at best. In turbulent debates, policymakers should first ask which tools are likely to (not) work, and it is what we offer in this case study.

[1] Anh V. Vu, Alice Hutchings, and Ross Anderson. "No Easy Way Out: the Effectiveness of Deplatforming an Extremist Forum to Suppress Hate and Harassment." *IEEE Symposium on Security and Privacy (S&P)*, 2024.

[2] Anh V. Vu, Lydia Wilson, Yi Ting Chua, Iliia Shumailov, and Ross Anderson. "ExtremeBB: A Database for Large-Scale Research into Online Hate, Harassment, the Manosphere and Extremism", *ACL Workshop on Online Abuse and Harms (WOAH)*, 2023.

[3] Ross Anderson and Sam Gilbert. "The Online Safety Bill", *Policy Brief, Bennett Institute for Public Policy*, 2022.

[4] Alice Hutchings, Richard Clayton, and Ross Anderson. "Taking Down Websites to Prevent Crime", *APWG Symposium on Electronic Crime Research (eCrime)*, 2016.