The Effectiveness of Deplatforming an Extremist Forum to Suppress Online Hate and Harassment

Anh V. Vu, Alice Hutchings, Ross Anderson anh.vu@cst.cam.ac.uk





Forums Chat What's new Members

New posts Find threads Watched Search forums Mark forums read



Lolcows



Lolcows

Lolcows are people with eccentric behavior who can be "milked" for entertainment. Threads: 1.8K Messages: 1.4M



Community Watch

Lolcows tend to thrive in communities with like-minded people who do not criticize their behavior. Threads: 746 Messages: 793K



Internationale Clique

Do you speak multiple languages and know about people who'd only appeal to speakers of your language? Read this, and make your contribution to the Internationale Clique today !!

Threads: 106 Messages: 45.1K



Animal Control

The people, communities, and drama of the furry and pony fandoms.

Threads: 338 Messages: 342.7K



Christian Weston Chandler

Chris-Chan is the author of *Sonichu* and is considered one of the most heavily documented people in human history.

Threads: 3.7K Messages: 685.2K



Ethan Ralph / Da Sektur

The #Killstream and the smoldering ruins of a once great Sektur.

Threads: 427 Messages: 428.2K

What is Kiwi Farms?

- The largest forum known for online harassment.
- Tied to various suicides.
- 10+ year lifespan.
- 15M+ posts in total.
- Used Cloudflare DDoS protection for years.

until...



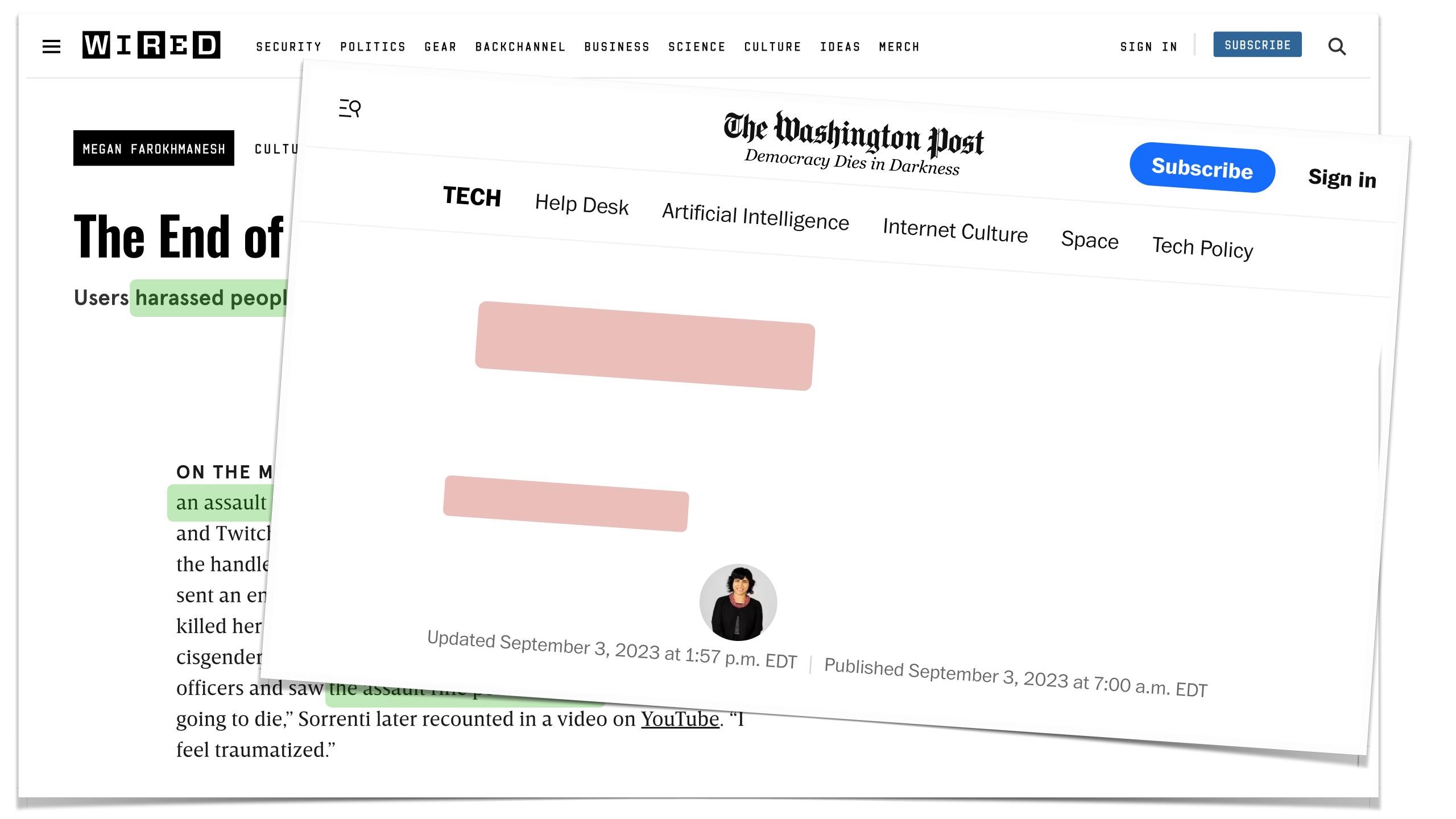
WIRED

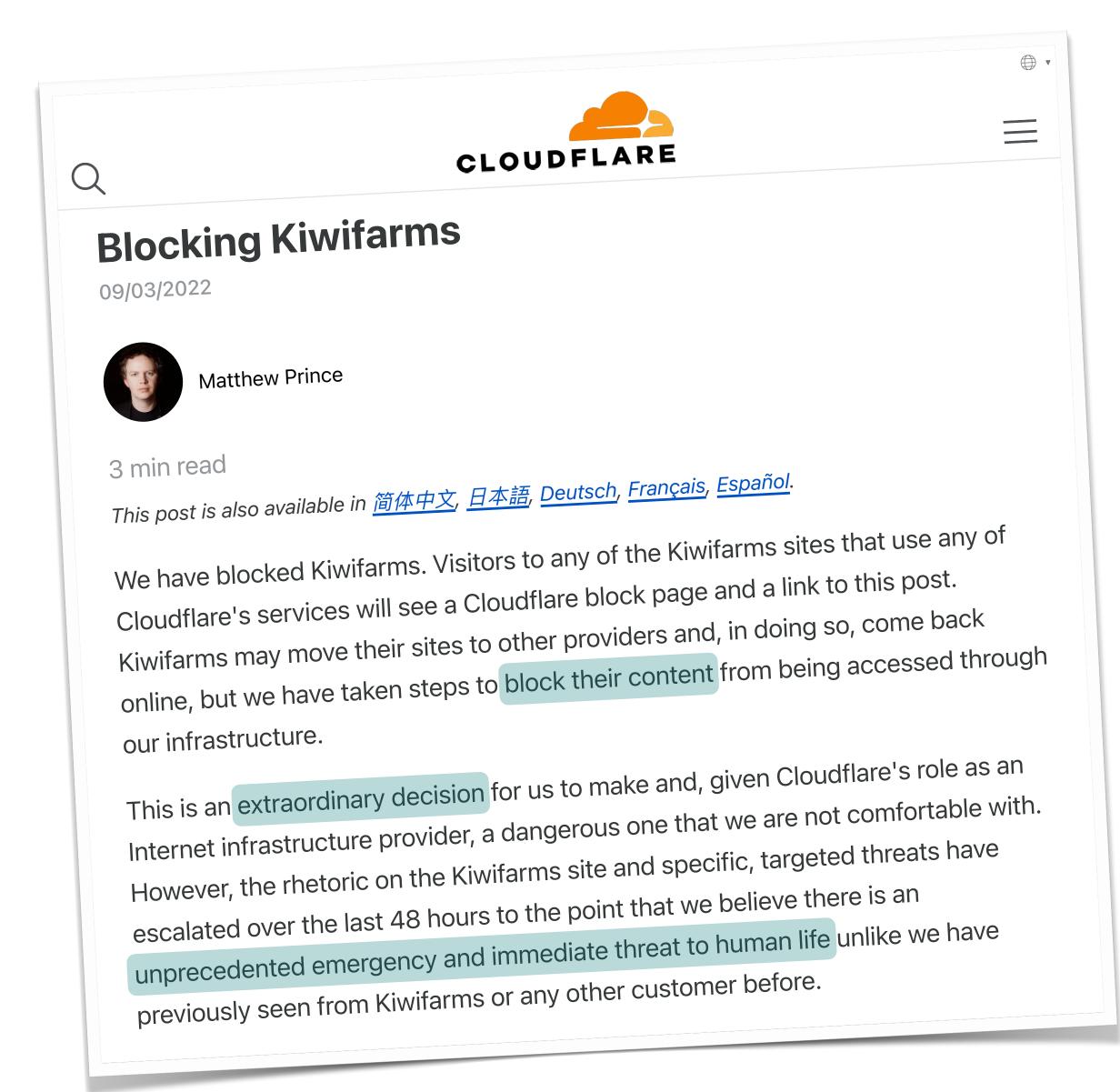
CULTURE SEP 8, 2022 12:47 PM

The End of Kiwi Farms, the Web's Most Notorious Stalker Site

Users harassed people for a decade. Then they messed with the wrong woman.

ON THE MORNING of August 5, in London, Ontario, police put an assault rifle in Clara Sorrenti's face. Sorrenti is a trans activist and Twitch streamer who provides political commentary under the handle Keffals. Earlier that morning, an impersonator had sent an email to city councillors claiming that Sorrenti had killed her mother and would soon go to City Hall to shoot every cisgender person she saw. "When I was woken up by police officers and saw the assault rifle pointed at me, I thought I was going to die," Sorrenti later recounted in a video on <u>YouTube</u>. "I feel traumatized."











Blocking Kiwifarms

09/03/2022



3 min read

This post is also available in <u>简体中文</u>, <u>日本語</u>, <u>Deutsch</u>, <u>Français</u>

We have blocked Kiwifarms. Visitors to any of the Kiv Cloudflare's services will see a Cloudflare block page Kiwifarms may move their sites to other providers ar online, but we have taken steps to block their conte our infrastructure.

This is an extraordinary decision for us to make an Internet infrastructure provider, a dangerous one t However, the rhetoric on the Kiwifarms site and sr escalated over the last 48 hours to the point that unprecedented emergency and immediate threa previously seen from Kiwifarms or any other cus



Home / Blog / DDoS-Guard terminating services for Kiwi Farms

DDoS-Guard terminating services for Kiwi Farms

On September 4, Kiwi Farms forum administrators activated DDoS-Guard protection

As a DDoS protection provider, we have customers around the world and it is not our duty to moderate content on their sites. We don't have to decide whether a website violates the law. We only can restrict access to the customer's website if it is reasonable. For example, if there's the official notice from the court.

However, today we did not wait for the official notification and stopped providing services for the Kiwi Farms forum. We have received multiple complaints from users saying that it violates the DDoS-Guard Acceptable Use Policy.

Having analyzed the content of the site, we decided on the termination of DDoS protection services for kiwifarms.ru. To all those who brought this incident to our



Blocking Kiwifarms

09/03/2022



3 min read

This post is also available in <u>简体中文</u>, <u>日本語</u>, <u>Deutsch</u>, <u>Français</u>

We have blocked Kiwifarms. Visitors to any of the Kiv Cloudflare's services will see a Cloudflare block page Kiwifarms may move their sites to other providers ar online, but we have taken steps to block their conte our infrastructure.

This is an extraordinary decision for us to make an Internet infrastructure provider, a dangerous one t However, the rhetoric on the Kiwifarms site and sr escalated over the last 48 hours to the point that unprecedented emergency and immediate threa previously seen from Kiwifarms or any other cus



Home / Blog / DDoS-Guard terminating

DDoS-Guard terminal

September 5, 2022 • 1303

On September 4, Kiwi Farms forum a for their domain.

As a DDoS protection provider, we ha duty to moderate content on their site violates the law. We only can restrict a reasonable. For example, if there's the

However, today we did not wait for the services for the Kiwi Farms forum. We h saying that it violates the DDoS-Guard A

Service Continuation Of Kiwi Farms...

My name is Hugo Carvalho and I'm the CEO of DiamWall, today I'm bringing an important statement that I never thought I would need to do anytime soon.

About our CDN...

Before even starting to talk about Kiwi Farms I think it is important to let people know what a CDN is and does. We are only a proxy that filters website traffic and blocks malicious requests, we **DO NOT** host any website, we **ARE NOT**

responsible for any website's content and we **CANNOT** control every service's content.

Everyone is free to register an account at our website, pay for the product and start using it. We **DO NOT** and **CANNOT** analyze every single website that starts using our services.

DiamWall is completely new in this market (we opened to the public one month ago) and this is not a good way to start for sure, nor is it the way that we wanted to start.

DiamWall was formed by a small group of people that worked for many years in this area with private customers and we created the best Anti-DDoS technologies that ever existed.

Now that you understand how it works, let's talk about Kiwi Farms...

The owner of Kiwi Farms came in need of DDoS Protection and because their website was offline due to DDoS, we didn't really know about their website's content. They had a **PROBLEM** and we had the **SOLUTION**.

Soon enough the reports started to arrive and we started digging more and more about this website, soon enough we found that Kiwi Farms hosts a lot of revolting content.

We do not think that is fair to terminate any service because of public pressure but in this case, we think there is some foundation behind all those requests and we really do not want to have anything to do with it.

So, after careful consideration, DiamWall does not feel like Kiwi Farms is a good fit for our company, which means that today we are SUSPENDING their service inside DiamWall's CDN.

Please be **AWARE** that this will not fix the issue, only delay it, they will find another provider that can protect them and will be online once again.

Best Regards, Hugo Carvalho, DiamWall CEO

Having analyzed the content of the site, we decided on the termination of DDoS protection services for kiwifarms.ru. To all those who brought this incident to our

also: Internet Archive, hCaptcha, Voxility, Lumen, Arelion, GTT, Zayo...



Blocking Kiwifarms

09/03/2022



3 min read

This post is also available in <u>简体中文</u>, <u>日本語</u>, <u>Deutsch</u>, <u>Français</u>

We have blocked Kiwifarms. Visitors to any of the Kiv Cloudflare's services will see a Cloudflare block page Kiwifarms may move their sites to other providers ar online, but we have taken steps to block their conte our infrastructure.

This is an extraordinary decision for us to make an Internet infrastructure provider, a dangerous one t However, the rhetoric on the Kiwifarms site and sr escalated over the last 48 hours to the point that unprecedented emergency and immediate threa previously seen from Kiwifarms or any other cus



Home / Blog / DDoS-Guard terminating

DDoS-Guard terminal

September 5, 2022 • 1303

On September 4, Kiwi Farms forum a for their domain.

As a DDoS protection provider, we ha duty to moderate content on their site violates the law. We only can restrict a reasonable. For example, if there's the

However, today we did not wait for the services for the Kiwi Farms forum. We h saying that it violates the DDoS-Guard A

Service Continuation Of Kiwi Farms...

My name is Hugo Carvalho and I'm the CEO of DiamWall, today I'm bringing an important statement that I never thought I would need to do anytime soon.

About our CDN...

Before even starting to talk about Kiwi Farms I think it is important to let people know what a CDN is and does.

We are only a proxy that filters website traffic and blocks malicious requests, we **DO NOT** host any website, we **ARE NOT** responsible for any website's content and we **CANNOT** control every service's content.

Everyone is free to register an account at our website, pay for the product and start using it. We **DO NOT** and **CANNOT** analyze every single website that starts using our services.

DiamWall is completely new in this market (we opened to the public one month ago) and this is not a good way to start for sure, nor is it the way that we wanted to start.

DiamWall was formed by a small group of people that worked for many years in this area with private customers and we created the best Anti-DDoS technologies that ever existed.

Now that you understand how it works, let's talk about Kiwi Farms...

The owner of Kiwi Farms came in need of DDoS Protection and because their website was offline due to DDoS, we didn't really know about their website's content. They had a **PROBLEM** and we had the **SOLUTION**.

Soon enough the reports started to arrive and we started digging more and more about this website, soon enough we found that Kiwi Farms hosts a lot of revolting content.

We do not think that is fair to terminate any service because of public pressure but in this case, we think there is some foundation behind all those requests and we really do not want to have anything to do with it.

So, after careful consideration, DiamWall does not feel like Kiwi Farms is a good fit for our company, which means that today we are SUSPENDING their service inside DiamWall's CDN.

Please be **AWARE** that this will not fix the issue, only delay it, they will find another provider that can protect them and will be online once again.

Best Regards, Hugo Carvalho, DiamWall CEO

Having analyzed the content of the site, we decided on the termination of DDoS protection services for kiwifarms.ru. To all those who brought this incident to our

how effective is this coordinated effort?



Blocking Kiwifarms

09/03/2022



3 min read

This post is also available in <u>简体中文</u>, <u>日本語</u>, <u>Deutsch</u>, <u>Français</u>

We have blocked Kiwifarms. Visitors to any of the Kiv Cloudflare's services will see a Cloudflare block page Kiwifarms may move their sites to other providers ar online, but we have taken steps to block their conte our infrastructure.

This is an extraordinary decision for us to make an



Home / Blog / DDoS-Guard terminating

DDoS-Guard terminal

September 5, 2022 • 1303

On September 4, Kiwi Farms forum a for their domain.

As a DDoS protection provider, we ha duty to moderate content on their site violates the law. We only can restrict a reasonable. For example, if there's the

However, today we did not wait for the services for the Kiwi Farms forum Wob

Service Continuation Of Kiwi Farms...

My name is Hugo Carvalho and I'm the CEO of DiamWall, today I'm bringing an important statement that I never thought I would need to do anytime soon.

About our CDN...

Before even starting to talk about Kiwi Farms I think it is important to let people know what a CDN is and does.

We are only a proxy that filters website traffic and blocks malicious requests, we **DO NOT** host any website, we **ARE NOT** responsible for any website's content and we **CANNOT** control every service's content.

Everyone is free to register an account at our website, pay for the product and start using it. We **DO NOT** and **CANNOT** analyze every single website that starts using our services.

DiamWall is completely new in this market (we opened to the public one month ago) and this is not a good way to start for sure, nor is it the way that we wanted to start.

DiamWall was formed by a small group of people that worked for many years in this area with private customers and we created the best Anti-DDoS technologies that ever existed.

Now that you understand how it works, let's talk about Kiwi Farms...

The owner of Kiwi Farms came in need of DDoS Protection and because their website was offline due to DDoS, we didn't really know about their website's content. They had a **PROBLEM** and we had the **SOLUTION**.

Soon enough the reports started to arrive and we started digging more and more about this website, soon enough we found that Kiwi Farms hosts a lot of revolting content.

We do not think that is fair to terminate any service because of public pressure but in this case, we think there is some foundation behind all those requests and we really do not want to have anything to do with it.

So, after careful consideration, DiamWall does not feel like Kiwi Farms is a good fit for our company, which means that today we are SUSPENDING their service inside DiamWall's CDN.

Please be **AWARE** that this will not fix the issue, only delay it, they will find another provider that can protect them and will be online once again.

Best Regards,

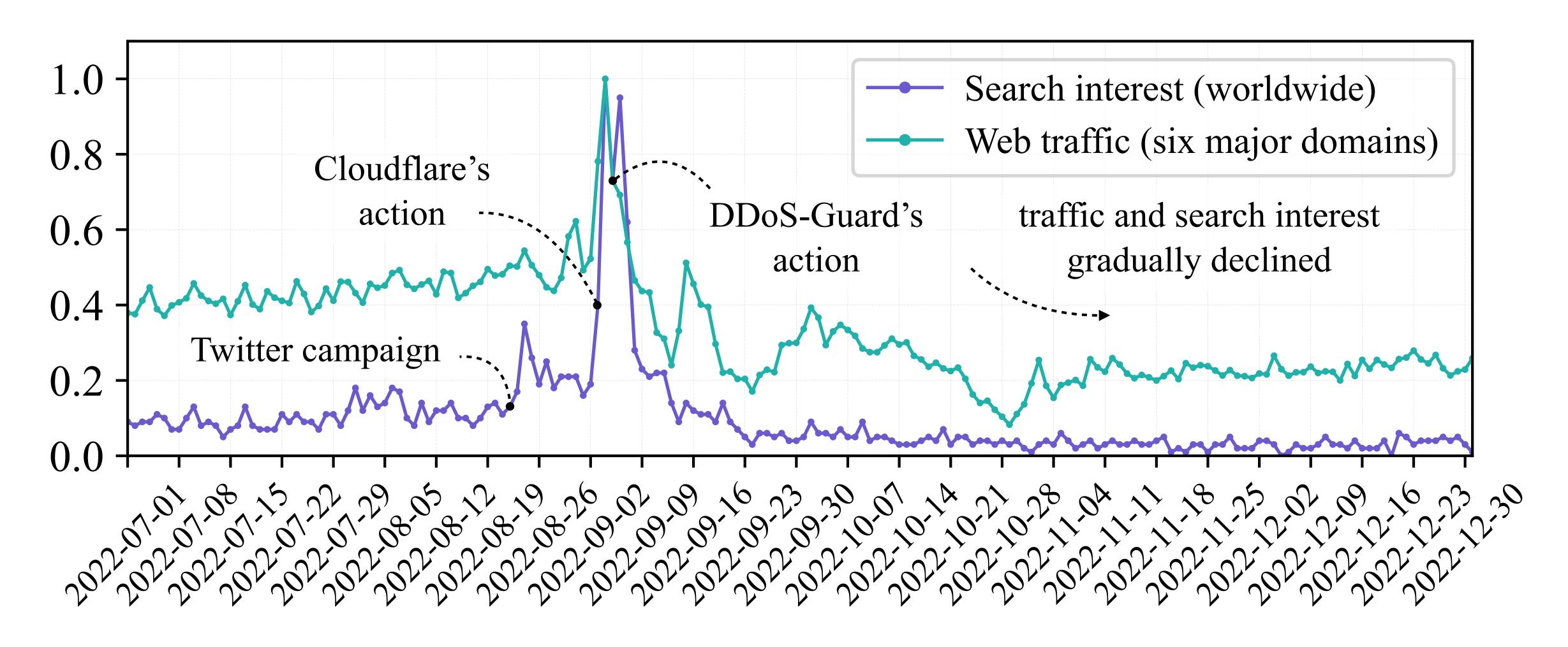
Methods and **Datasets**

- I5M+ forum posts, 500k+ Telegram messages, and I Ik+ Tweets.
- Web traffic to 6 major domains of Kiwi Farms, by Similarweb.
- Search interest in Kiwi Farms, as shown by Google Trends.

how effective is this coordinated effort?

Unintended consequences: the Streisand effect

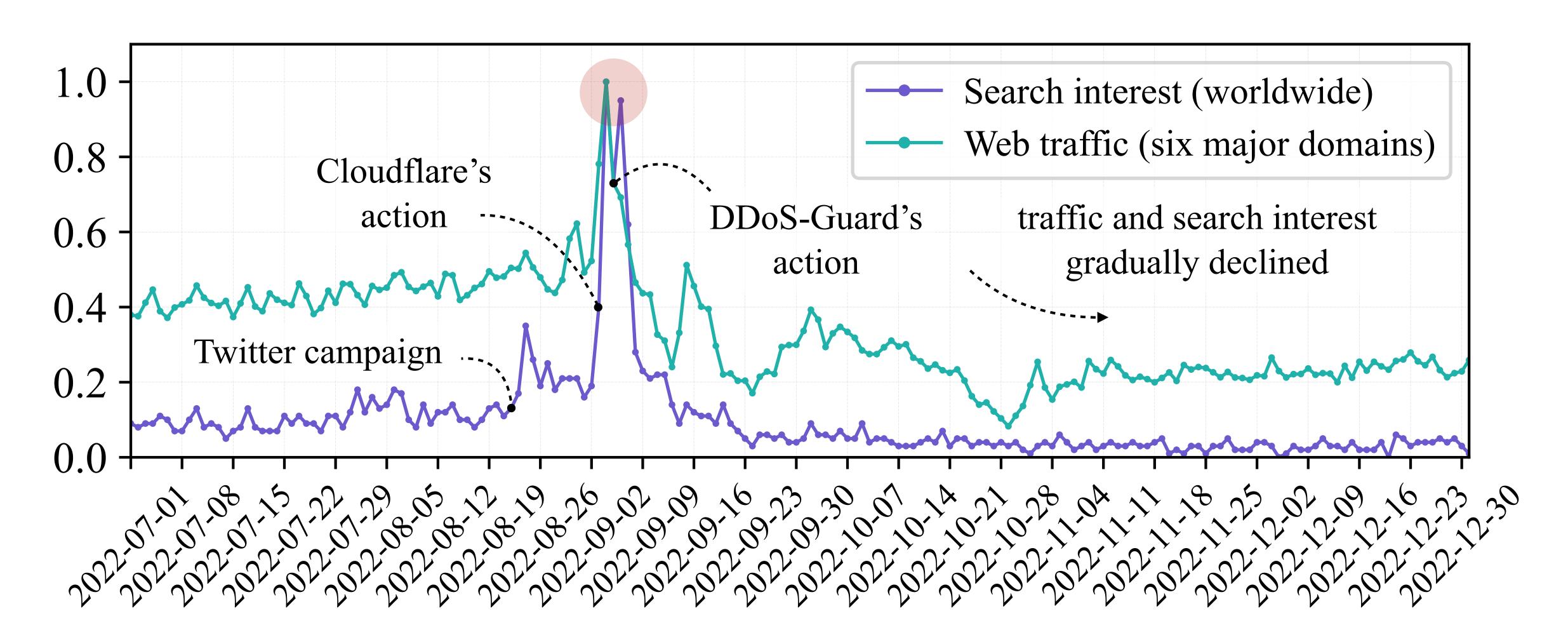




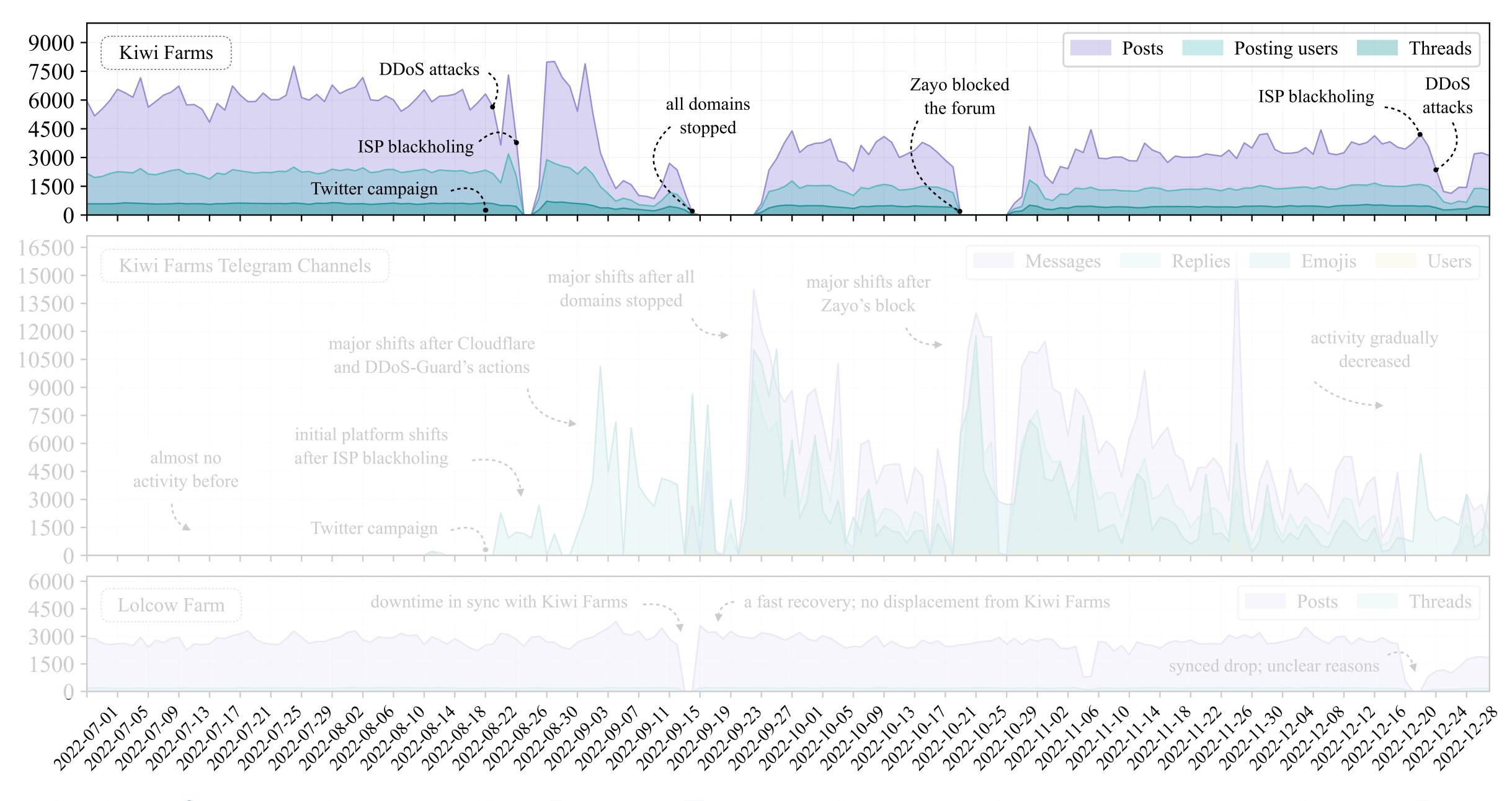
The global search trends for Kiwi Farms and visit traffic to its six major domains

Unintended consequences: the Streisand effect

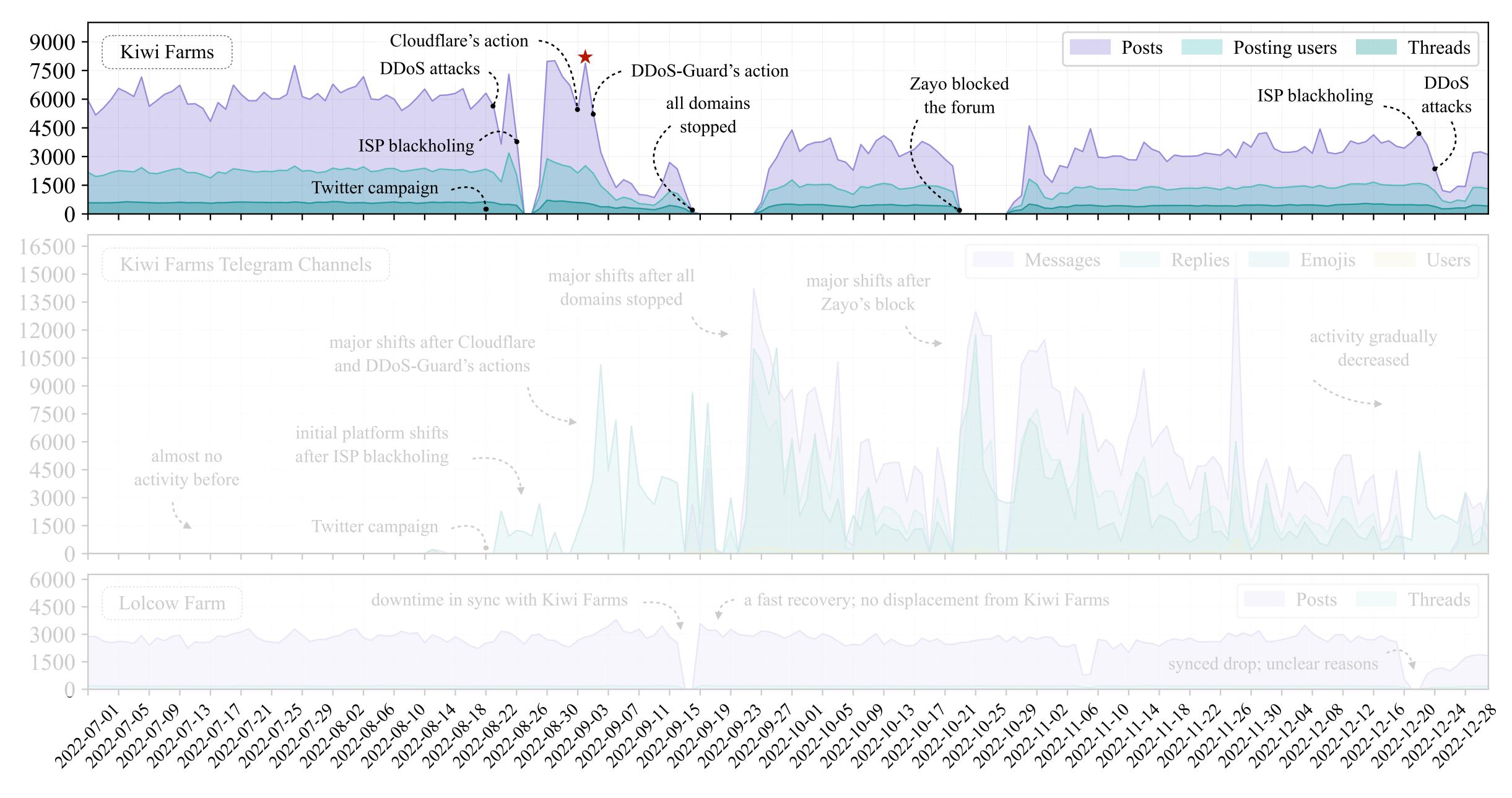




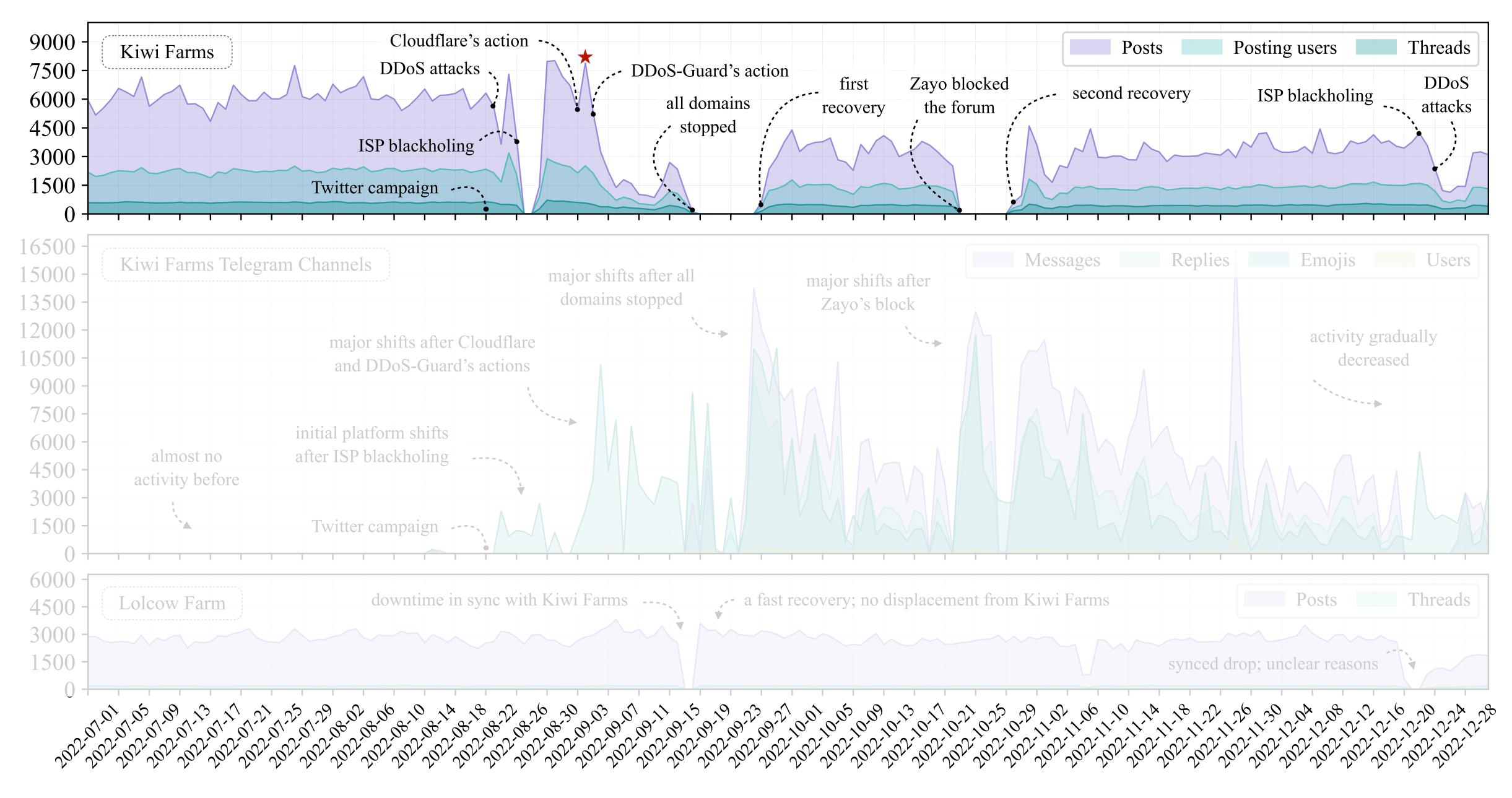
The global search trends for Kiwi Farms and visit traffic to its six major domains



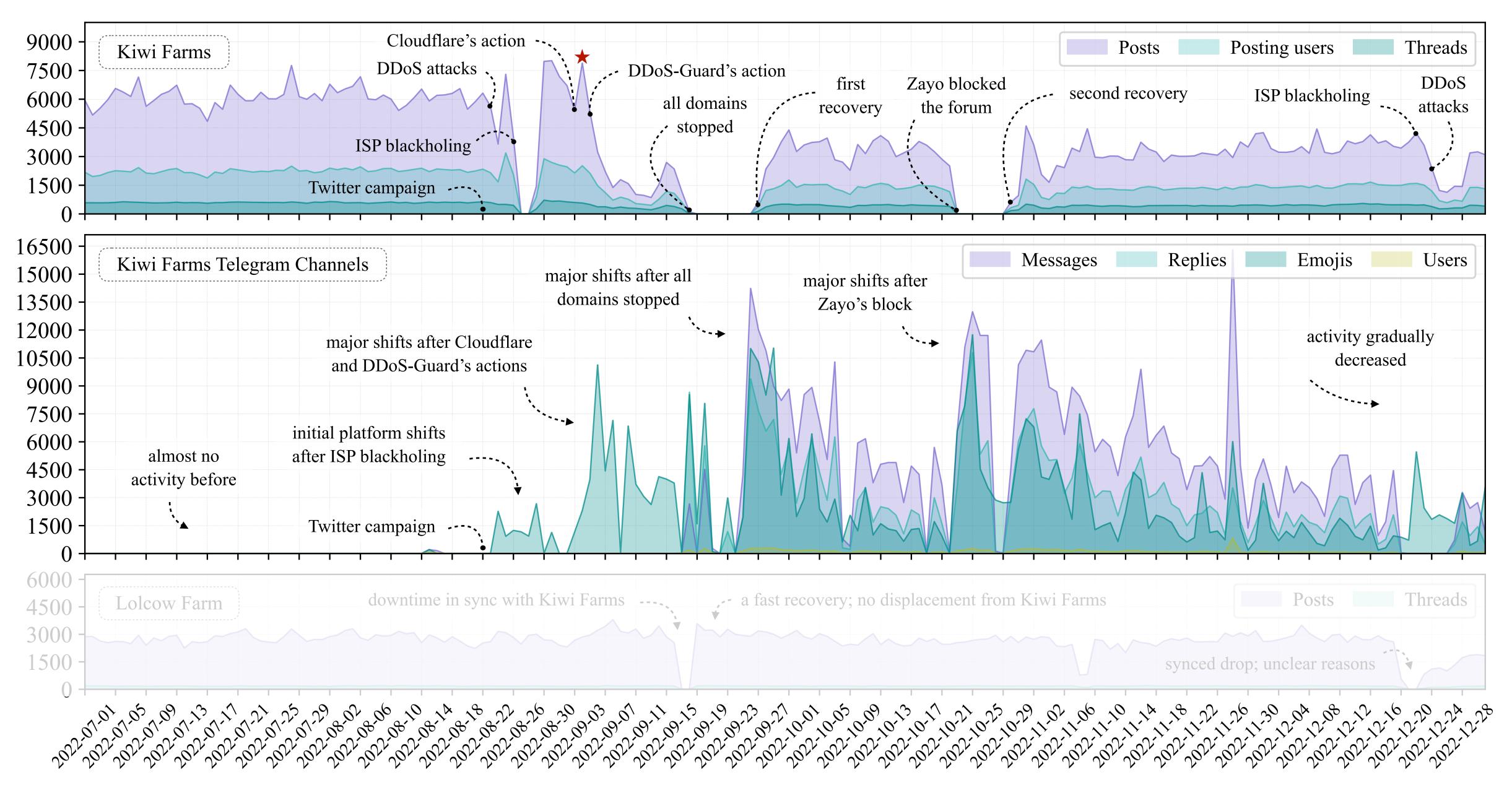
Number of posts and users on Kiwi Farms, its Telegram channels, and primary competitor Lolcow Farm



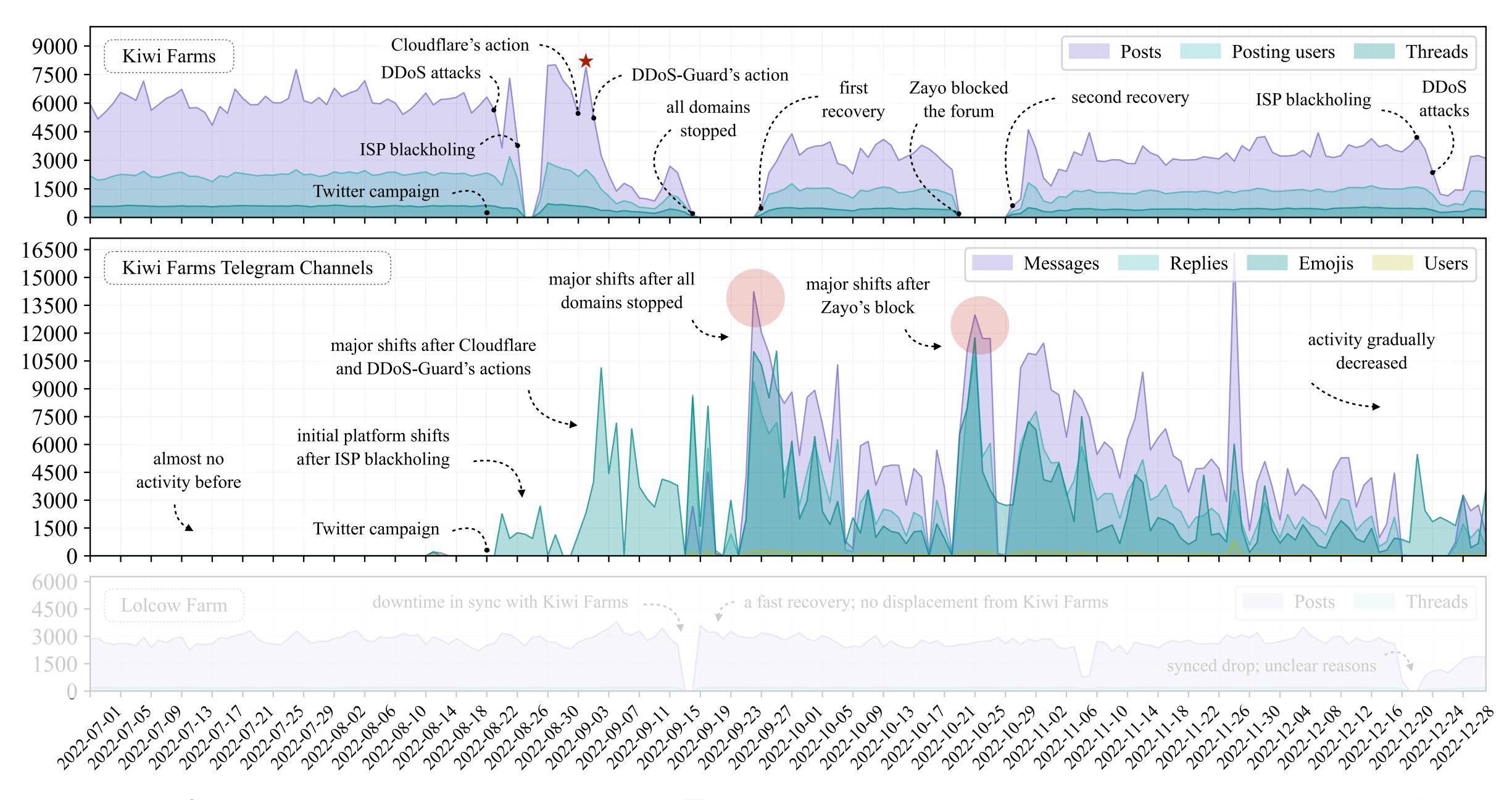
Number of posts and users on Kiwi Farms, its Telegram channels, and primary competitor Lolcow Farm



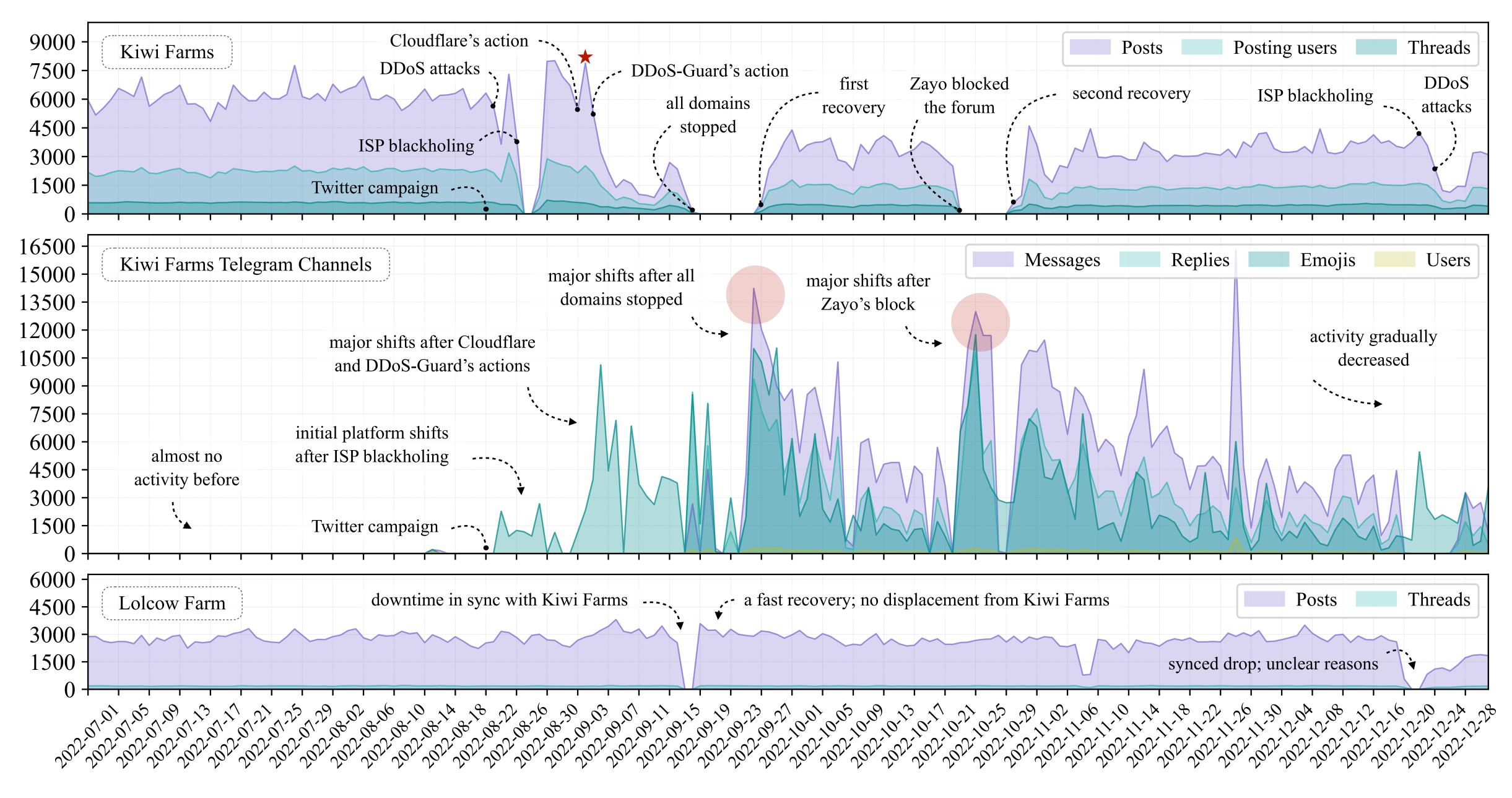
Number of posts and users on Kiwi Farms, its Telegram channels, and primary competitor Lolcow Farm



Number of posts and users on Kiwi Farms, its Telegram channels, and primary competitor Lolcow Farm

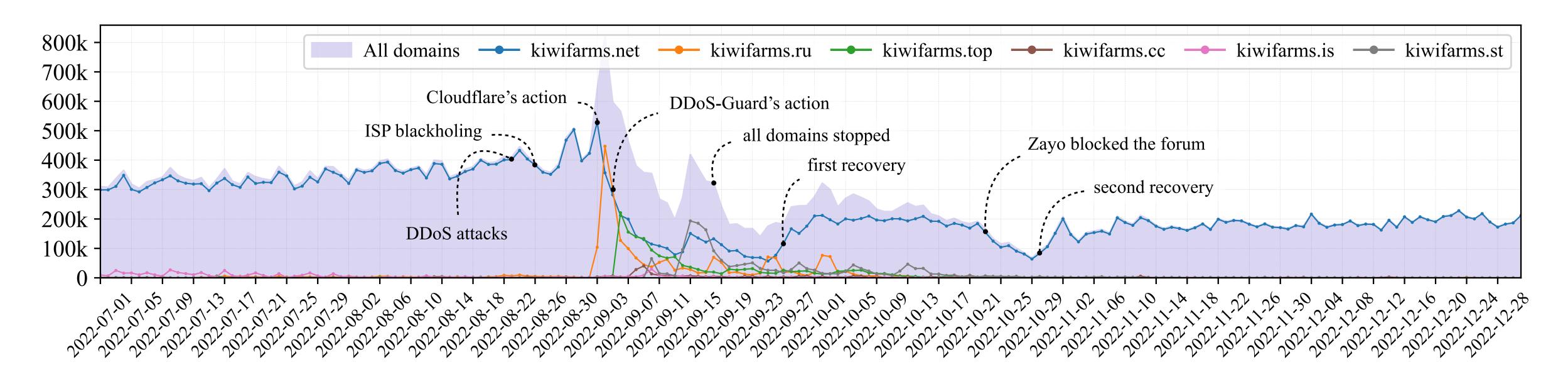


Number of posts and users on Kiwi Farms, its Telegram channels, and primary competitor Lolcow Farm



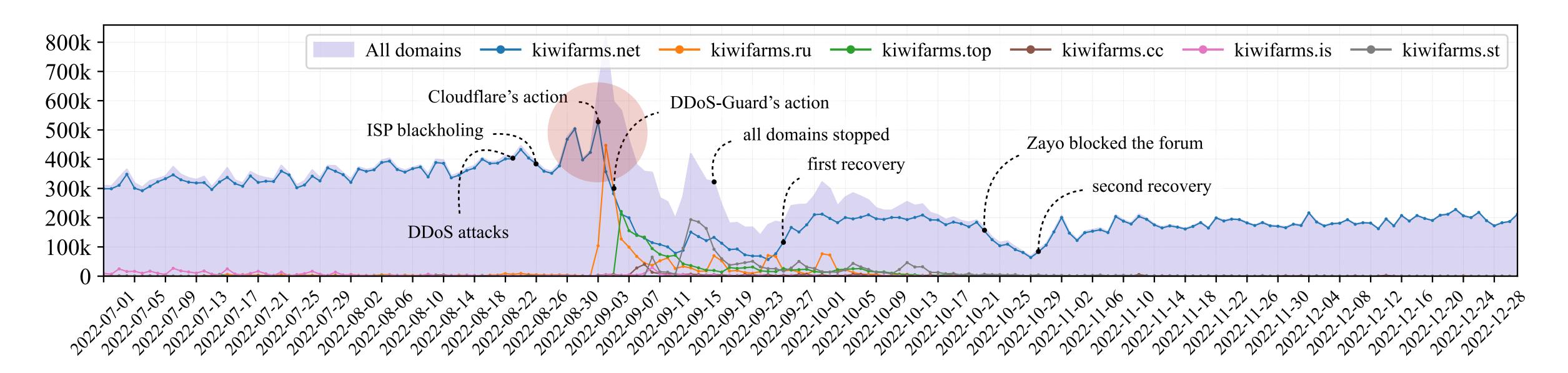
Number of posts and users on Kiwi Farms, its Telegram channels, and primary competitor Lolcow Farm





The number of daily estimated Similar web visits and the traffic fragmentation into multiple old domains

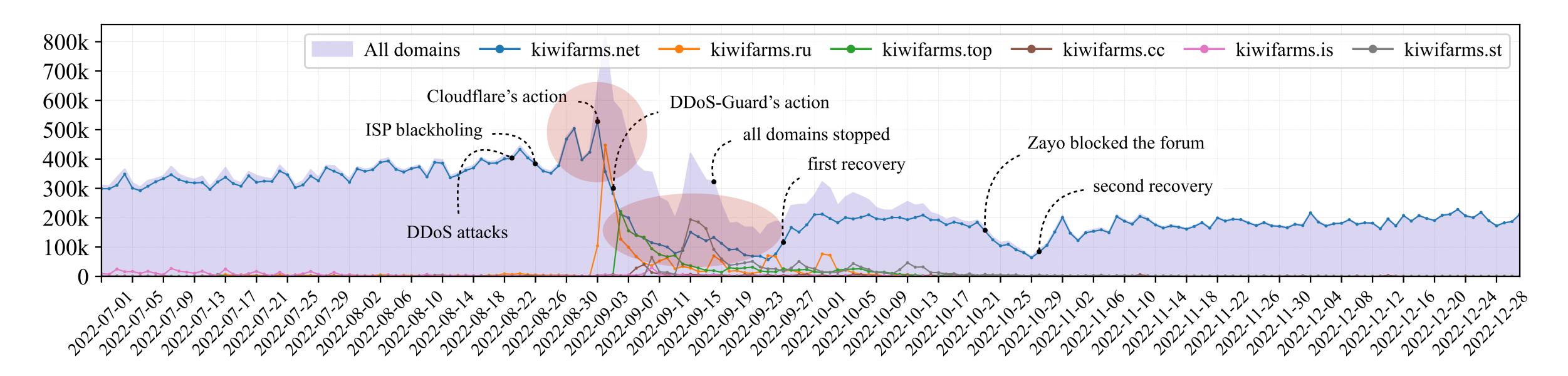




The number of daily estimated Similar web visits and the traffic fragmentation into multiple old domains

• The publicity given by the takedown increased awareness and attracted more visitors.

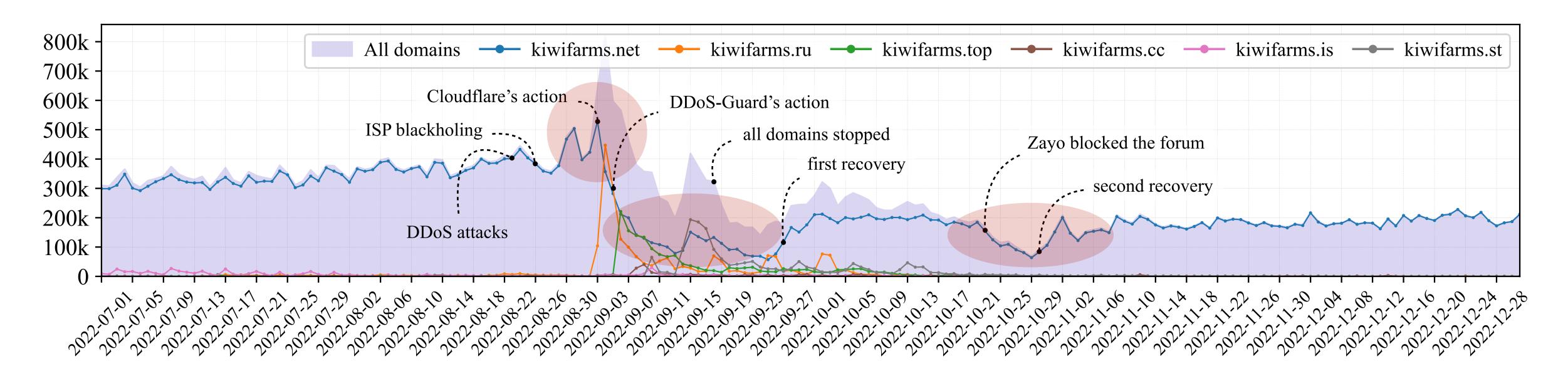




The number of daily estimated Similar web visits and the traffic fragmentation into multiple old domains

- The publicity given by the takedown increased awareness and attracted more visitors.
- Most traffic was fragmented into other old domains shortly after the disruption.





The number of daily estimated Similar web visits and the traffic fragmentation into multiple old domains

- The publicity given by the takedown increased awareness and attracted more visitors.
- Most traffic was fragmented into other old domains shortly after the disruption.
- The forum recovery got all the traffic back to the primary domain.
- The forum traffic has now been fully recovered, and even more active than before!

Concluding remarks

See paper for



- Forum users were more connected.
- Details of the industry responses.
- The reactions of the leading community.
- User discussions about the campaign.





No Easy Way Out: the Effectiveness of Deplatforming an Extremist Forum to Suppress Hate and Harassment

Anh V. Vu
University of Cambridge
Cambridge Cybercrime Centre
anh.vu@cl.cam.ac.uk

Alice Hutchings
University of Cambridge
Cambridge Cybercrime Centre
alice.hutchings@cl.cam.ac.uk

Ross Anderson
University of Cambridge
and University of Edinburgh
ross.anderson@cl.cam.ac.uk

Abstract-Legislators and policymakers worldwide are debating options for suppressing illegal, harmful and undesirable material online. Drawing on several quantitative data sources, we show that deplatforming an active community to suppress online hate and harassment, even with a substantial concerted effort involving several tech firms, can be hard. Our case study is the disruption of the largest and longest-running harassment forum KIWI FARMS in late 2022, which is probably the most extensive industry effort to date. Despite the active participation of a number of tech companies over several consecutive months, this campaign failed to shut down the forum and remove its objectionable content. While briefly raising public awareness, it led to rapid platform displacement and traffic fragmentation. Part of the activity decamped to Telegram, while traffic shifted from the primary domain to previously abandoned alternatives. The forum experienced intermittent outages for several weeks, after which the community leading the campaign lost interest, traffic was directed back to the main domain, users quickly returned, and the forum was back online and became even more connected. The forum members themselves stopped discussing the incident shortly thereafter, and the net effect was that forum activity, active users, threads, posts and traffic were all cut by about half. The disruption largely affected casual users (of whom roughly 87% left), while half the core members remained engaged. It also drew many newcomers, who exhibited increasing levels of toxicity during the first few weeks of participation. Deplatforming a community without a court order raises philosophical issues about censorship versus free speech; ethical and legal issues about the role of industry in online content moderation; and practical issues on the efficacy of private-sector versus government action. Deplatforming a dispersed community using a series of court orders against individual service providers appears unlikely to be very effective if the censor cannot incapacitate the key maintainers, whether by arresting them, enjoining them or otherwise deterring them.

1. Introduction

Online content is now prevalent, widely accessible, and influential in shaping public discourse. Yet while online places facilitate free speech, they do the same for hate speech [1],

and the line between the two is often contested. Some cases of stalking, bullying, and doxxing such as Gamergate have had real-world consequences, including violent crime and political mobilisation [2]. Content moderation has become a critical function of tech companies, but also a political tussle space, since abusive accounts may affect online communities in significantly different ways [3]. Online social platforms employ various mechanisms, for example, artificial intelligence [4], to detect, moderate, and suppress objectionable content [5], including "hard" and "soft" techniques [6]. These range from reporting users of illegal content to the police, through deplatforming users breaking terms of service [7], to moderating legal but obnoxious content [8]. which may involve actions such as flagging it with warnings, downranking it in recommendation algorithms [9], or preventing its being monetised through ads [10].

Deplatforming may mean blocking individual users, but sometimes the target is not a single bad actor, but a whole community, such as one involved in crime [11]. It can be undertaken by industry, as when Cloudflare, GoDaddy, Google and some other firms terminated service for the DAILY STORMER after the Unite the Right rally in Virginia in 2017 [12] and for 8CHAN in August 2019 [13]; or by law enforcement, as with the FBI taking down DDoS-forhire services in 2018 [14], [15] and 2022 [16], [17], and seizing RAID FORUMS in 2022 [18]. Industry disruption has often been short-lived; both 8CHAN and DAILY STORMER reemerged or relocated shortly after being disrupted. Police intervention is often slow and less effective, and its impact may also be temporary [11]. After the FBI terminated SILK ROAD [19], the online drug market fragmented among multiple smaller ones [20]. The seizure of RAID FORUMS [18] led to the emergence of its successors Breach Forums, EXPOSED FORUMS, and ONNI FORUMS. Furthermore, the FBI takedowns of DDoS-for-hire services cut the attack volume significantly, yet the market recovered rapidly [14], [15].

KIWI FARMS is the largest and longest-running online harassment forum [21]. It is often associated with real-life trolling and doxxing campaigns against feminists, gay rights campaigners and minorities such as disabled, transgender, and autistic individuals; some have killed themselves after being harassed [22]. Despite being unpleasant and widely controversial, the forum has been online for a decade and

All our datasets are available!
Anh V. Vu — anh.vu@cst.cam.ac.uk



Concluding remarks

- Deplatforming an active community to suppress online hate and harassment without incapacitating the owners can be challenging!
- A series of competent tech firms, with a swift effort, have not succeeded. Government and police actions may be less effective.
- The UK Online Safety Bill empowering a regulator (Ofcom) to stop 'harmful but not illegal' services may not be very effective.

See paper for <</p>



- Forum users were more connected.
- Details of the industry responses.
- The reactions of the leading community.
 - User discussions about the campaign.





No Easy Way Out: the Effectiveness of Deplatforming an Extremist Forum to Suppress Hate and Harassment

Anh V. Vu
University of Cambridge
Cambridge Cybercrime Centre
anh.vu@cl.cam.ac.uk

Alice Hutchings
University of Cambridge
Cambridge Cybercrime Centre
alice.hutchings@cl.cam.ac.uk

Ross Anderson
University of Cambridge
and University of Edinburgh
ross.anderson@cl.cam.ac.uk

Abstract-Legislators and policymakers worldwide are debating options for suppressing illegal, harmful and undesirable material online. Drawing on several quantitative data sources, we show that deplatforming an active community to suppress online hate and harassment, even with a substantial concerted effort involving several tech firms, can be hard. Our case study is the disruption of the largest and longest-running harassment forum KIWI FARMS in late 2022, which is probably the most extensive industry effort to date. Despite the active participation of a number of tech companies over several consecutive months, this campaign failed to shut down the forum and remove its objectionable content. While briefly raising public awareness, it led to rapid platform displacement and traffic fragmentation. Part of the activity decamped to Telegram, while traffic shifted from the primary domain to previously abandoned alternatives. The forum experienced intermittent outages for several weeks, after which the community leading the campaign lost interest, traffic was directed back to the main domain, users quickly returned, and the forum was back online and became even more connected. The forum members themselves stopped discussing the incident shortly thereafter, and the net effect was that forum activity, active users, threads, posts and traffic were all cut by about half. The disruption largely affected casual users (of whom roughly 87% left), while half the core members remained engaged. It also drew many newcomers, who exhibited increasing levels of toxicity during the first few weeks of participation. Deplatforming a community without a court order raises philosophical issues about censorship versus free speech; ethical and legal issues about the role of industry in online content moderation; and practical issues on the efficacy of private-sector versus government action. Deplatforming a dispersed community using a series of court orders against individual service providers appears unlikely to be very effective if the censor cannot incapacitate the key maintainers, whether by arresting them, enjoining them or otherwise deterring them.

1. Introduction

Online content is now prevalent, widely accessible, and influential in shaping public discourse. Yet while online places facilitate free speech, they do the same for hate speech [1],

and the line between the two is often contested. Some cases of stalking, bullying, and doxxing such as Gamergate have had real-world consequences, including violent crime and political mobilisation [2]. Content moderation has become a critical function of tech companies, but also a political tussle space, since abusive accounts may affect online communities in significantly different ways [3]. Online social platforms employ various mechanisms, for example, artificial intelligence [4], to detect, moderate, and suppress objectionable content [5], including "hard" and "soft" techniques [6]. These range from reporting users of illegal content to the police, through deplatforming users breaking terms of service [7], to moderating legal but obnoxious content [8], which may involve actions such as flagging it with warnings, downranking it in recommendation algorithms [9], or preventing its being monetised through ads [10].

Deplatforming may mean blocking individual users, but sometimes the target is not a single bad actor, but a whole community, such as one involved in crime [11]. It can be undertaken by industry, as when Cloudflare, GoDaddy, Google and some other firms terminated service for the DAILY STORMER after the Unite the Right rally in Virginia in 2017 [12] and for 8CHAN in August 2019 [13]; or by law enforcement, as with the FBI taking down DDoS-forhire services in 2018 [14], [15] and 2022 [16], [17], and seizing RAID FORUMS in 2022 [18]. Industry disruption has often been short-lived; both 8CHAN and DAILY STORMER reemerged or relocated shortly after being disrupted. Police intervention is often slow and less effective, and its impact may also be temporary [11]. After the FBI terminated SILK ROAD [19], the online drug market fragmented among multiple smaller ones [20]. The seizure of RAID FORUMS [18] led to the emergence of its successors Breach Forums, EXPOSED FORUMS, and ONNI FORUMS. Furthermore, the FBI takedowns of DDoS-for-hire services cut the attack volume significantly, yet the market recovered rapidly [14], [15].

KIWI FARMS is the largest and longest-running online harassment forum [21]. It is often associated with real-life trolling and doxxing campaigns against feminists, gay rights campaigners and minorities such as disabled, transgender, and autistic individuals; some have killed themselves after being harassed [22]. Despite being unpleasant and widely controversial, the forum has been online for a decade and

All our datasets are available!
Anh V. Vu — anh.vu@cst.cam.ac.uk



Concluding remarks

- Deplatforming an active community to suppress online hate and harassment without incapacitating the owners can be challenging!
- A series of competent tech firms, with a swift effort, have not succeeded. Government and police actions may be less effective.
- The UK Online Safety Bill empowering a regulator (Ofcom) to stop 'harmful but not illegal' services may not be very effective.

See paper for



- Forum users were more connected.
- Details of the industry responses.
- The reactions of the leading community.
 - User discussions about the campaign.



Professor Ross Anderson (1956 - 2024)





All our datasets are available! Anh V. Vu — anh.vu@cst.cam.ac.uk

