*'truth'*" [**P14**]. It also forces "*people to become rather tribalistic and a priori suspicious of people with different views*" [**P**]. The "political tribalism" on social media [2], in turn, makes it "*easier to demonize people with different opinions and political attitudes and avoid scrutinizing the like-minded ones*" [**P2**], which plays directly in the hands of the "misinformers."

As for the "misinformers", our participants unequivocally identified the state-sponsored "appropriators" that hijacked the original hacktivist playbook to spread *external propaganda* on social media. That nation-states enjoyed a reputation for promulgating disinformation in the past was not a news to the hacktivists (e.g. "*Russia has always been really good at it*" [**P2**]), but instead what caught them aback was the "*audacity and the sophistication*" [**P4**] in utilizing trolling and memes on such a massive scale [134]. Reflecting on this shift in online operations, **P3** believes that "*disinfo ops [sic] and hacking our intellectual property is all these nation-states are left with because they can't beat us militarily or economically*." Not necessarily neoliberal, but nonetheless authoritarian, the elites behind the external propaganda in equal degree conflicts with the *mistrust of authority and promote decentralization* hacker postulate because is a "*blatant effort to control the social media turf and the mass of population spending their time there* [**P15**]. The external propaganda nature of disinformation also conflicts with the *all information should be free* hacker postulate in the view of the hackers in our sample because "*overshadows and complicated an access to other more factual or useful information*" [**P2**].

## 6 Active Countering of Misinformation

Literature on misinformation focuses on helping the social media *users* discern falsehoods with strategies for "prebunking" i.e. forewarning and preemptive refutation of the falsehoods [68] or "debunking" i.e. providing users verifiable corrections of the falsehoods from credible sources to break the illusion of truth [30, 93]. An *algorithmic* is also available for the mainstream social media platforms (the alternative ones do not deem misinformation as a problem [107]) that leverages natural language processing, image analysis, or metadata to detect trolling and memes [50, 51, 122]. Platforms also have the option for algorithmic "soft" moderation by either obscuring trolling and memes with warnings covers or attaching warning labels [108, 125] and "hard" moderation for removing or suspending misinformer accounts [63]. None of these solutions, however fends of troll farms and meme disseminators effectively, so we wanted to know what hacktivists have to propose instead in the second research question.

### 6.1 Leaking, Doxing, and Deplatforming

Suspending user accounts by social media platforms for breach of their code of conduct is referred to as "deplatforming" [1]. In the context of hacktivism, it takes a border meaning as hacktivists do investigative work that entails leaking and doxing but also confrontation with the misinformers that, in their subjective view, breaches the vision of democratic Internet. For example, hacktivists did a massive API scrapping of the alt-platform Parler to leak data that tied users to the Capitol Riots and the QAnon conspiracy [94], which in turn resulted in a massive account deplatforming on Twitter [15]. These activities spur operations to confront and expose the QAnon conspirators on social media (e.g. @QAnonAnonymous [22]), amongst which some of our hacktivists have a direct role in "*dismantling the Qanon infrastructure*" [**P2**].

The deplatforming targets political misinformation campaigns where our hacktivists "*compiled and leaked dossiers on individuals spreading hateful propaganda and those who seek to sow the seeds of violence*" [**P1**] on social media. These operations were targeted both on "*individual spreaders, nation-states, even companies with murky records*" [**P2**]. Several mentioned their direct operations for exposing disinformation relative to the "*Ukrainian conflict*" [**P5**], praising the work of the Ukrainian IT Army outfit for dispelling the myth that Ukraine is committing genocide against Russians in the Donbas region [23]. Hacktivists were dedicated in "*doxing companies and governmental agencies in response to the political meddling in the US internal affairs from places like Russia, Iran, and China*" [**P8**]. Misinformation "*sanctioned by the governments*" was targeted by hackers in attempts to deplatform prominent "*disinformation front agents on social media, like Irina Tsukerman, for example*" [**P3**].

Leaks and doxing were equally utilized for misinformation beyond political counter(argumentation) and external propaganda. One of the hacktivists has dedicated considerable time on exposing cryptocurrency scammers on social media and elsewhere, deeming the feeling of it as "*better than sex*" [**P5**]. Another was focused leaking personal details about predators on social media that spread misinformation to cover their sexual harassment and cyberstalking towards women, "*exposing both their sock puppet accounts and their real name on Twitter*" [**P3**]. Another pushed back against criminal misinformation by doxing "*bullies, liars, and fraudsters*" [**P20**] and one "*anti-cancel culture in case of minors*" hacktivist noted that they "*successfully deplatformed major participants in hate campaigns and stalking of minors*" on social media [**P5**].

### 6.2 Anti-Misinformation "Ops"

The hacktivists in our sample engaged in misinformation saturation ops, true to the their commitment to "fight misinformation with more information." One of the hacktivists stated that it is "*expected from the hacktivist community to combat misinformation in such a way*" and noted that "*it is the sole reason they maintain a Twitter account*" [**P3**]. Another one seconded this posture noting that "*it is frustrating to see misinformation from others and other creators but that is the main reason I continue to post on TikTok*" [**P17**]. In the words of **P2**, "*there is more ideological aspect of it when I am fighting disinformation*," directly invoking the mission of the true hacktivists to become reflexively "*loud and determined*" to speak true information in response to the "*general assholery of misinformation on internet*" [**P9**].

Partaking in operation #NAFO (North Atlantic Fellas Organization) dedicated to countering Russian propaganda and disinformation in Ukraine by weaponizing memes [103], our participants materialized a combination of saturation and doxing to "*curtail misinformers' ability to gain followers*" [**P1**]. They extended their work to counter "*extremists and fascists and their toxic conspiracy theories*" [**P1**] by disrupting their funding and deplatforming prominent followers, true to the spirit of the "Antifa" hacktivist counterculture [131]. In a similar vein, one of the hacktivists proclaimed that they "*greatly contributed in the #OpJane operation*" [**P10**]. #OpJane is the latest operation launched by Anonymous against Texas for enacting the anti-abortion Bill 8 that allows "abortion bounty" for anyone who will investigate and report abortion in the state of Texas [38]. Interestingly, in the announcement of the operation, Anonymous calls for "fighting misinformation with enough plausible and difficult to disprove misinformation" to make any data these bounty hunters gather as useless [5].

## 7 Misinformation Evolution

As there is virtually no cost of disseminating misinformation [85], it is unlikely that the online discourse will shed off the alternative narratives soon. If this gloomy prediction will eventually materialize [78] or the Internet will improve because the new technologies will upgrade public's ability to judge the quality and veracity of content [4], remains an open issue. Because the hacktivists are nonetheless stakeholders in resolving this issue, our third research question aimed to bring their prediction of how online spaces will fare with trolling, memes, and falsehoods in the near future.

### 7.1 Counter-Misinformation Tactics

The hacktivists in our sample unanimously posit that "*it is hard for social media platforms to keep up with removing it, so people stepping in to help is going to be of critical importance*" [**P13**] for preserving a healthy discourse. The mobilization for "*justice and truth as a cause*" [**P15**] is important not just for curbing misinformation but "*reclaiming information back from the political hold*" [**P1**]. To help "*expose misinformation charlatans*" [**P4**], hacktivists call for maintaining a code of conduct where "*no leak, doxing, or exposure action should cause anyone else harm (physical, reputation, mental)*" [**P3**].

To begin with, **P3** recommends that we should "*stop treating disinformation as a freedom of speech*." As misinformers usually use this cloak to act very aggressively on social media, the next step is to "*identify what their weakness are and what triggers them - deplatforming or provocation?*" [**P14**]. If the misinformers are unresponsive spreaders, then "*exposing, doxing, and putting their real faces through OSINT*" [**P15**] is in due place not just on mainstream social media but also alt-platforms, forums and everywhere on Internet. If they itch for a provocation, then "*orchestrated saturation*" [**P5**] might work better with "*shitposts, absurd trolling, and ridiculing memes*" [**P18**]. Here, the hacktivists note, it is vitally important to *a priori* distance from a "*political whataboutery*" [**P14**] and avoid "*coming across as censorship, disagreement, canceling that only could cause argument or dismissal*" [**P5**].

Some of the hacktivists were on the opinion that "*doxing is not hacking anymore per se because you can get stuff with a credit card and documents could be easily faked nowadays*" [**P1**]. One possible tactic, proposed by **P1**, was to "*find exploits, vulnerabilities in their platforms and step-by-step expose misinformers' amateurish way of doing trolling, using bots, and feeding think tanks to get a credibility behind their propaganda*." Another tactic, proposed by **P2**, was "*doxing for the purpose of having advertisers pull from supporting known misinformer influencers, like for example in the case of Andy Ngo*." Proposing more of a hybrid hacktivist tactics, **P4** suggested "*a latent, yet coordinated psychological warfare where psychologists rip apart these people, conduct serious OSINT to find incriminating leaks on them, and even pay for billboards and radio ads to publicly shame them*." Along these lines, **P11** even suggested throwing the book at them, *targeting them with a social engineering attack and attempting to compromise a piece of their core infrastructure, be that their servers, Internet access, or bot credentials*."

### 7.2 Misinformation Literacy

Hacktivists in our sample echo the sentiment regarding the social media users' susceptibility to false information found in scientific literature: laziness to check facts [**P2**] [89], resistance to authoritative suggestions [**P7**] [57], allegiance [**P13**] [120], and simple ignorance [**P16**] [17]. As people that resort to action, hacktivists do feel the obligation to propose ways for addressing this susceptibility. In the view of **P5**, "*misinformation needs to be seen as something everyone is being watched for, and not just one group of people on the left or the right,*" A "misinformation social contract" [136] necessitates interventions such as "*a critical thinking curricula in schools*" [**P18**], "*teaching hacking OpSec skills as social responsibility and rise to action*" [**P5**], and "*forcing professional communication norms on platforms*" [**P16**].

As hacktivists have little control over these interventions, they were happy to help with a development of "*truth-spreading bots for a 'standoff' with misinformation-spreading bots*" as something that could append the practice of leaks, doxing, and exposure [**P13**]. They recognized that these "truth-spreading bots" must help ordinary users to better find and locate facts, as information literacy is the single most effective one in dispelling falsehoods [55]. Hacktivists reiterate that platforms do have to let "*misinformation to float on social media and make bots visible, so they gets overwhelmed with factual information*" to demonstrate to ordinary users how to do help themselves [**P14**].

Regardless if it these stances are realistic or not, the hacktivists in our sample believe that the current approach to raising misinformation literacy is ineffective because it does not signal an "*unbiased attitude*" [**P7**] to the social media users in the wrong. Instead of an educational and respectable tone, "*rather a 'cancel culture' infused or a 'your opinion is wrong' tone*" [**P3**] plagues any attempt to help people to navigate and locate factual information. Rejection of misinformation, as a result of misinformation literacy, must come as an agreement that "*scientific facts do not have political properties, even if the social media platforms inherently do*" [**P5**].

### 7.3 Misinformation hacktivism

The participants in our sample acknowledge that orchestrated *misinformation hacktivism*, bar individual instances of ops against misinformers, is largely absent from social media. For the hacktivists to assume misinformation as a worthy cause for action, the conflict between the past "hacking for political causes" and [58] future "hacking against using falsehoods in furthering political causes" [22] must be resolved. Though this conflict is complex and evolving, several of the hacktivists worried that it could nevertheless create a "*division between the hacktivists on political lines*" [**P2**].

As a relative threat to the misinformation activism, one participant mentioned the hijacking of the hacktivists image for self-promotion, e.g. "*some like to portrait themselves as woke gods of the web with zero fuck-ups*" [**P12**]. Another threat is the temptation of using misinformation against misinformation, as in the #OpJane campaign [**P10**]. While this strategy is true to the "fight-fire-with-fire" approach, it might backfire in circumstances where abiding to the hacktivist ethic comes secondary to expressing social and political angst on social media [79]. On top of this, one could argue that this conflict *per se* might be hard to resolve in the misinformation instance as external propaganda, because even if the hacktivists are "hacking for the homeland," they nonetheless are doing it on political terms [26].

## 8 Discussion

### 8.1 Implications

The new brand of misinformation, our findings show, draws the ire of the hacktivists, reprehending the hijacked discourse for political and propagandistic proposes. The "fight-fire-with-fire" response – leaks, doxing, and deplatforming – though individually employed by some of the participants in our sample, is yet to be orchestrated and tested against serious disinformation outfits that, unfortunately, are still out there on social media [48]. The early evidence outside of the US shows that this orchestration works as the IT Army leaked data from Russian organizations in response to the troll farms' disinformation narrative that Ukraine is committing genocide against Russians in the Donbas region [23].

The hacktivists' resoluteness to go after the misinformers would certainly have implications for the content/user moderation on social media, user participation, and future of Internet activism overall. Moderating users and content on social media was, and still is, the response by the mainstream platforms to the political and public health misinformation [108]. Alternative platforms like Gab, Gettr, and Parler, seen as the seeding grounds for this misinformation [133], on the other hand, never did, nor currently do, employ any content/user moderation [107]. While the content/user moderation incites a migration from the mainstream to the alt-platforms [133], it remains to be seen whether the deplatforming will have the same effect. Mainstream social media had a mixed response to leaks and doxing in the past (e.g. allowing WikiLeaks [114] and barring the Hunter Biden's laptop leaks [28]), so this also adds uncertainty if and how the hacktivists' "fight-fire-with-fire" approach will be allowed, moderated, or perhaps even forced to migrate entirely outside of the social media space.

Trolling and memes might still maintain the popularity amongst the misinformers, however, the latest modes of social media participation like short videos on TikTok open new "fronts" for both the misinformers and the hacktivists. TikTok has increasingly been tested as the next "battlefield" of alternative narratives with evidence of health and abortion misinformation [8, 112] and an individual engagement by at least one of participants. Recalling that the hacktivists' #OpJane was waged in response to the abortion ban laws in Texas and called for "misinformation-against-(mis)information" [38], it is yet to see how the leaks, doxing, and deplatforming will materialize with meme-ified videos and trolling. TikTok claims it does health and abortion misinformation moderation [123], but evidence shows that this is lax and largely ineffective [14], adding an additional incentive for shifting the disinformation campaigns on this platform.

TikTok is also the next platform for Internet activism where the hashtag activism is appended with videos expanding the developing news narratives, such as the coverage of the Black Lives Matter movement and the Capitol riot [72]. TikTok presents content not just from viral hashtags but also their variations (e.g. #abotion but also #abôrtion [112]) so the threat of hashtag hijacking, co-opting, and counter hash tagging will inevitably materialize here too. This particular affordance likely will allow for weaponizing deepfakes in appending the hashtag war in near future, as they already appeared in misinformation videos about the COVID-19 pandemic on TikTok [106]. All of these developments would certainly necessitate a dynamic adaptation in the way doxing, leaking, and deplatforming is performed in order not just to avoid disintegration of the Internet activism and hacktivism, but prevent