Revista Científica del Centro Universitario de la Guardia Civil
Revista Logos Guardia Civil

Research Article

HACKTIVISM: FROM SOCIAL PROTEST TO THE INSTRUMENTALISATION OF THE STATE

English translation with AI assistance (DeepL)

Josué Expósito Guisado
Sergeant of the Guardia Civil
PhD student at the University Pablo de Olavide
Master's degree in Peace, Security and Defence by the Gutiérrez Mellado University Institute (UNED)
jexpgui@gmail.com
ORCID: 0009-0003-4977-3899

Received 18/03/2025
Accepted 05/05/2025
Published 27/06/2025

Recommended citation: Expósito, J. (2025). Hacktivism: from social protest to state instrumentalisation of the State. *Logos Guardia Civil Magazine*, 3(2), pp. 101-122.

Legal Deposit: M-3619-2023
NIPO online: 126-23-019-8
ISSN online: 2952-394X

ORCID: 0009-0003-4977-3899

HACKTIVISM: FROM SOCIAL PROTEST TO THE INSTRUMENTALISATION OF THE STATE

Summary: INTRODUCTION. 2. FROM SOCIAL PROTEST TO CYBER WAR. 3. 3.THE LINK BETWEEN HACKTIVISM AND APT. 4. THE FUTURE OF HACKTIVIST GROUPS. 5.CONCLUSIONS. 6. BIBLIOGRAPHICAL REFERENCES.

Abstract: Hacktivism has evolved from an initial form of digital protest to become a key tool in contemporary geopolitical conflicts. What began as a decentralised movement in defence of freedom of expression and social justice has been progressively instrumentalised by states to execute cyber-attacks, manipulate public opinion and deploy disinformation operations. This phenomenon has been particularly accentuated in the context of the war in Ukraine, where the convergence between Advanced Persistent Threat (APT) groups and patriotic hacktivists has allowed the execution of cyber operations coordinated with state interests. In parallel, the internationalisation of hacktivism has led to the formation of alliances between groups in different regions, broadening its impact beyond the Russian-Ukrainian conflict. Cyberspace has established itself as an ideal arena for confrontation between states in a controlled environment. However, the growing sophistication of attacks and increasingly strategic targeting pose serious challenges to international stability and the security of Western states.

Resumen: El hacktivismo ha evolucionado desde una forma inicial de protesta digital hasta convertirse en una herramienta clave en los conflictos geopolíticos contemporáneos. Lo que comenzó como un movimiento descentralizado en defensa de la libertad de expresión y la justicia social, ha sido progresivamente instrumentalizado por los Estados para ejecutar ciberataques, manipular la opinión pública y desplegar operaciones de desinformación. Un fenómeno que se ha visto especialmente acentuado en el marco de la guerra de Ucrania, donde la convergencia entre grupos de Amenaza Persistente Avanzada (APT) y hacktivistas patrióticos ha permitido la ejecución de operaciones cibernéticas coordinadas con los intereses estatales. Paralelamente, la internacionalización del hacktivismo ha llevado a la formación de alianzas entre grupos de distintas regiones, ampliando su impacto más allá del conflicto ruso-ucraniano. El ciberespacio se ha consolidado como un escenario idóneo para la confrontación entre Estados en un entorno controlado. Sin embargo, la creciente sofisticación de los ataques y la selección de objetivos cada vez más estratégicos plantean serios desafíos a la estabilidad internacional y la seguridad de los Estados occidentales.

Keywords: Hacktivism, APTs, cyberproxies, cyberconflict, cyberattacks.

Palabras clave: Hacktivismo, APT, ciberproxies, ciberconflicto, ciberataques.

ABBREVIATIONS

APT: *Advanced Persistent Threat.*

DDoS: *Distributed Denial of Service* attack.

DOJ: US *Department of Justice*.

FBI: *Federal Bureau of Investigation*.

GRU: Main Intelligence Directorate of Russia (*Glavnoe Razvedyvatel'noe Upravlenie*).

ICS: *Industrial Control Systems.*

IRGC: Iran's *Islamic Revolutionary Guard Corps.*

IT Army of Ukraine: IT Army of Ukraine.

NSA: US National *Security Agency*.

PMC: *Private Military Company*.

PSOA: *Private Sector Offensive Actor*.

SCADA: *Supervisory Control And Data Acquisition.*

Stuxnet: Name of the *malware* used in the "Olympic Games" operation.

1. INTRODUCTION

The war in Ukraine has meant that the Western world has once again been confronted with the impact of political realism. Prior to the invasion of 2022, the vast majority of Western analysts were unable to envision a conventional conflict on the international scene such as the one that continues to occur on Europe's doorstep. Blinded by soft-power doctrines and following the liberal paradigms of capitalist peace or commercial peace theory, European leaders willfully ignored the fact that in some parts of the world, political realism still reigns supreme.

In an increasingly digitised world, where there is virtual interconnection between the intangible plane of information technology and physical space itself, it is not surprising that Europe's current enemy poses a security challenge. As states have become increasingly dependent on information technologies, so have the opportunities for hostile actors (state and non-state) to influence the political and geopolitical environment by deploying actions in cyberspace.

The war in Ukraine has not only marked not only the beginning of an operation of harassment and cyber disruption by cyberthreats linked to the Kremlin, but has also brought about a change in the global hacktivist landscape: what until not so many years ago was the bastion of the defence of freedom of expression, privacy, social justice and human rights, is now a tool with strategic implications and, in many cases, linked directly or indirectly to governments and intelligence services.

The ideological and protest-oriented digital activism that Anonymous once represented is evolving into a phenomenon made up of a myriad of nationalist groups that repeatedly use distributed denial of service (DDoS) attacks to create a climate of tension and persistent harassment of Western enemies.

Hacktivism has become a double-edged tool. On the one hand, it represents a form of expression and struggle for social justice, transparency and human rights. On the other, it has become a weapon used by states to deploy political destabilisation and disinformation campaigns.

The use of cyberattacks for geopolitical purposes has highlighted the fine line between activism and state-sponsored cybercrime. This article seeks to analyse the evolution of hacktivism and its relationship with governments, as well as the role of Advanced Persistent Threat (APT) groups in the use of cyberspace for political and military purposes . [1]

Through a review of concrete cases, it will explore the collaboration (or instrumentalisation) of hacktivists by states, the implications of this practice and its impact on current geopolitics. Finally, a reflection will be offered on the future of

[1] Groups of cyber-attackers often associated with nation states or large criminal organisations, highly sophisticated and persistent, who infiltrate networks for long periods of time for espionage or sabotage and who have abundant resources (technical, economic) to attack high-value targets (governments, large companies) with premeditation and stealth.

hacktivism in an increasingly interconnected world, where artificial intelligence and other emerging technologies could redefine the role of these actors in cyberspace.

Hacktivism is no longer just a marginal phenomenon of digital protest, but a potential security risk for states. Understanding its evolution and implications is fundamental to analysing the future of Spanish cybersecurity.

2. FROM SOCIAL PROTEST TO CYBERWARFARE

Hacktivism has undergone a remarkable transformation since its origins, going from being a form of social protest to a tool used by governments to support a political agenda. If we think about it carefully, this transformation betrays the origins and the very essence of activism, which is why, before analysing the role played by hacktivism as a tool at the service of the state, we believe it is necessary to look at how this phenomenon has evolved since its origins.

In certain contemporary approaches, particularly those oriented towards terminological systematisation, there is a tendency to establish a hierarchical relationship between cyberactivism and hacktivism, understanding the former as a broader phenomenon and necessarily encompassing the latter as a specific manifestation or radicalised variant. This reading, present in both popular literature and some normative analytical frameworks, considers cyberactivism to represent the use of digital technologies for the promotion of social, political or cultural causes through awareness-raising campaigns, online petitions or virtual protests. Hacktivism, on the other hand, would be characterised by the use of hacking tools - such as distributed denial of service (DDoS) attacks, data breaches or website alteration - for similar purposes, albeit by more disruptive or even illicit means.

However, this interpretation, while widespread, is problematically reductionist and does not stand up to closer scrutiny from the historical and critical theory of digital movements. Firstly, the assumption of a linear and progressive evolution - from "moderate" cyberactivism to "radical" hacktivism - ignores the distinct historical trajectories of the two concepts. Hacktivism, far from being a late derivation of cyberactivism, emerges simultaneously and even earlier in certain contexts, rooted in the hacker culture of the 1980s and 1990s, and articulated around principles such as freedom of information, open access to knowledge and civil disobedience in cyberspace (Jordan & Taylor, 2004; Coleman, 2014).

In fact, the term "hacktivism" arises from the etymological combination of "hacker" and "activism", describing the use of computer skills to promote political or social causes; and its roots go back to the mid-1990s, when groups like the "*Cult of the Dead Cow*" (a reference to the Texas slaughterhouse where the group holds its meetings) advocated universal access to online information as a fundamental human right and the fight against oppressive governments .[2]

"Cult of the Dead Cow, considered one of the founders of modern hacktivism, not only disseminated manifestos critical of state and corporate control of the Internet, but

[2] The website of "The Cult of the Dead Cow" can still be consulted at: https://cultdeadcow.com/about.html

also developed tools with a clear disruptive vocation. Among them is *Back Orifice* (1998), a software designed to expose vulnerabilities in the Windows operating system and denounce deficiencies in users' privacy[3] . A year later, in 1999, several of its members promoted the *Hacktivismo* project, a branch explicitly oriented towards the fight against digital censorship that gave rise to the development of tools such as *Six/Four* or *Peekabooty*, designed to circumvent the filters imposed by authoritarian regimes and facilitate free access to information.

In the *Cult of the Dead Cow*'s ideology, access to online information was not only a fundamental right, but also a field of political contestation that demanded innovative forms of technical and symbolic intervention. However, these actions, while non-violent in physical terms, implied a direct confrontation with restrictive legislation on network use and intellectual property; in other words, they revealed the ambiguous character of hacktivism.

On the other hand, conceptualising hacktivism as a simple tactical intensification of cyberactivism makes us lose sight of the ideological and epistemological divergences between the two. While cyberactivism tends to be framed within the logic of citizen participation, institutional advocacy and the strategic use of social media, hacktivism often operates on the basis of direct antagonism, resistance to power structures and the questioning of existing legal frameworks.

While it may be useful to think of hacktivism as a subcategory of cyberactivism from certain descriptive approaches, it is epistemologically insufficient and empirically questionable when addressing the genealogy, normative framework and ethical-political implications of both forms of digital activism. In this article we will focus solely on the evolution of hacktivism, understood as a phenomenon in its own right, leaving aside the formulation of a critical review of this classification.

In the early stages of hacktivism, the main objective was to carry out attacks against government and corporate entities as a form of protest against censorship and social injustices. These messages became more intense as the anti-globalisation movement of the mid-1990s emerged on the social scene (Auty, 2004).

A key milestone in the consolidation of hacktivism as a tool of political confrontation was the Kosovo war in the 1990s (often described as the first war fought online), where the contenders not only shared information and testimonies about the war online, but also spread propaganda and disinformation. *Hackers* even emerged and actively intervened in the conflict by defacing government websites and executing denial-of-service attacks against the opposing side's online infrastructures (Denning, 2001).

Academically and socially, hacktivist movements were perceived as the natural expression of a pre-existing political activism that had found in a new tool (the Internet) the possibility of employing a type of activist with a technical profile to spread its messages in a more mediatic way (Jordan, 2002).

[3] Although initially conceived as a security auditing tool, its creation generated some controversy and was perceived as a threat by the technology industry.

However, the manifest disregard for established norms, the names chosen by the groups (*The Legion of Doom, Bad Ass Mother Fuckers, Toxic Shock*, etc.) and the context of social insecurity opened up by the 9/11 attacks, meant that a phenomenon that was initially perceived positively began to arouse some mistrust. (Torres Soriano, 2018).

The figure of the *hacker* began to be identified with that of the criminal, and by extension, in a geopolitical context marked by the fight against Terror, with that of the cyberterrorist. And hacktivist actions began to be identified basically as a new form of illegitimate political participation, using cyber-attacks to carry out sabotage and cyber-espionage (Vegh, 2005).

At the academic level, the identification of hacktivism with the illegal or criminalisable, frequent in certain discourses, reduces hacktivism to a "radical form of cyberactivism", and thus impoverishes the analysis and explanatory capacity of the social sciences in the face of the complexity of contemporary digital political practices.

The beginnings of this decade reflect a hacktivism marked by the desire of its members to transgress social conventions for the fun of it. In fact, the roots of the best-known hacktivist group (Anonymous) can be traced back to the Japanese forum *2chan*, where the virtual community was dedicated to sharing all kinds of aberrant content related to anime, porn and practical jokes (Bartlett, 2015).

However, around 2003, the first internal tensions arose in a virtual community that had found in the *4chan* forum an ideal place to have fun regardless of the consequences. Precisely in this forum, some users (known as *moralfags*) proposed to focus their activities on more transcendental causes such as the fight against Internet censorship, in order to clean up the image of hacktivism and represent the defence of freedom of expression, transparency and other civil rights.

Under the slogan "*We are Anonymous. We are Legion. We do not forgive. We do not forget. Expect us.*" and the Guy Fawkes mask, a decentralised collective of activists emerged that combined the exfiltration of information and DDoS attacks to vindicate the fight against corruption, censorship and abuses of power.

From the powers-that-be, Anonymous was quickly interpreted as a premonition of the risk posed by a new generation of motivated virtual actors, with a leaderless structure and an operation based on voluntarism and spontaneity (Olson, 2012). However, it was not until the collective began to support the actions of WikiLeaks that the group was perceived as a top-level cyber threat.

In a short time, Anonymous grew from a small group of politically minded *hackers* to a global movement with thousands of followers around the world. However, their appeal did not lie in a structured ideology or a defined programme of action. Beyond their anti-establishment stance, which led them to denounce the manipulation and control exercised by governments and corporations, their philosophy lacked a clear orientation on how politics, society or the economy should be organised. This made Anonymous a difficult phenomenon to pigeonhole, as its identity was based more on action and protest than on a concrete agenda for change (Torres Soriano, 2018).

Under the Guy Fawkes mask gathered individuals who certainly believed they supported positive social change, but also others: those whose inspiration was the nihilistic destruction of the world as we know it and those who sought to hide under the banner of Anonymous for political or economic gain.

From the main legacy of Anonymous - turning hacktivism into a popular practice that transcended the *hacker* sphere - comes a new era of hacktivists operating in a landscape of fragmentation and complexity, where multiple actors with diverse motivations coexist.

Today, while groups such as Anonymous continue to operate in a decentralised way, their impact has diminished compared to the boom they reached in the early 2010s. At the same time, new generations of hacktivists have emerged, who, although they have a lower level of technical expertise, compensate with the use of automation tools and a mastery of media impact and social mobilisation.

Today, hacktivism is used both by independent collectives denouncing injustice and by state-sponsored groups instrumentalising these tactics for geopolitical purposes. The conflict between Russia and Ukraine has highlighted the existence of a cyber war, with pro-Ukrainian and pro-Russian hacktivists carrying out coordinated attacks for the benefit of their respective sides.

The boundary between legitimate digital activism, cybercrime and covert intelligence operations is increasingly blurred. However, we might consider that there are currently three types of hacktivists: cyberterrorists, civic *hackers* and patriotic *hackers* (Dahan, 2013; Denning, 2001; Johnson and Robinson, 2014; Sauter, 2013).

Cyberterrorism would include all hostile actions in cyberspace aimed at perpetrating acts of violence against people or property, with the aim of intimidating or coercing governments or societies to achieve specific political, religious or ideological ends. Their actions mainly involve spreading viruses and *malware*, vandalising *websites* and carrying out denial-of-service (DDoS) or *botnet* attacks (Denning, 2001; Jordan and Taylor, 2004; Goode, 2015).

In the category of civic *hackers* we would find all those organised groups that carry out actions against computer systems with the aim of contributing some good to the community, generally bordering on legality (Hunsinger and Schrock, 2016; Schrock, 2016).

Finally, patriotic *hackers* are those individuals or groups whose efforts are aligned with nationalist ideology and are considered a 'cyber militia' in pursuit of specific interests (Dahan, 2013; Green, 2016). Although from the outside these *hackers* may not appear to be directly sponsored by any state, we can now infer that they are instrumentalised as part of a larger web of state forces.

Patriotic *hacking* originated in China in the 1990s in response to anti-Chinese riots in Indonesia, and has since been used as a tactic by China, Russia, Syria and other states as a means to damage their enemies in the cyber domain. However, none of the operations prior to the Ukrainian war had achieved the scale, impact and governmental ties as robust

and prolonged, nor so blatantly transgressed international norms, as contemporary hacktivism (Healey & Grinberg, 2022).

3. THE NEXUS BETWEEN HACKTIVISM AND APT

Throughout history, states have resorted to proxy actors to carry out their conflict strategies without directly engaging their armed forces. Auxiliary units, mercenary groups, insurgencies, terrorist organisations or private military companies (PMCs) are just some of the forms that third actors have taken to act as substitutes for the strategic action of states.

It is therefore not surprising today that, in the light of an increasingly digitalised society, state action has found in hacktivist groups a new actor to personify the externalisation of authorship, and in cyberspace, the ideal environment to project geopolitical influence.

The concept of *surrogate warfare* has been the subject of extensive debate in the academic and security community, not least because of the difficulty in differentiating it from *proxy* warfare, given the closely intertwined nature of the two concepts.

In both terms, the objectives of the principal actor (the state) and the proxy agent coincide. However, while in *proxy warfare* there are two or more hierarchically related actors (the principal actor works for, with and through the proxy to achieve a common goal), in *surrogate warfare* these actors are aligned only if the principal actor is able to mobilise the adequate support required by the proxy (Fox 2019). In other words, the concepts of *surrogate warfare* and *proxy warfare* differ according to the relationship between the actors and their motivations.

Since hacktivist groups have little independence to resist the control of the state that sponsors (or at least influences or tolerates) them, in our case study we will speak in terms of *proxy* actors.

More specifically, to refer to them we will use Rondeaux and Sterman's (2019) definition of "*proxy* actors", who define them as "*subjects outside the security structure of the states involved in a conflict who act under direct or indirect sponsorship of a conventional actor (a state)*"; and Maurer's (2018) definition of *cyber proxies* as "*intermediaries who carry out offensive actions in cyberspace for the benefit of a principal actor*".

Historically, *cyberproxies* have been personified through various entities linked to the world of cybercrime and cyberespionage. However, the term encompasses a large number of organised entities that, directly or indirectly, pose a risk factor for companies and states. In fact, the list of actors is very long: criminal groups, private *sector offensive actors* (PSOA), terrorist groups, insurgents, insurgents, hacktivists, state actors or APTs are just some of them.

The reasons behind their use are varied: (1) the use of *proxy* actors by governments reduces the risk of escalation in conflicts, since the difficulty of attributing responsibility for a cyber-attack is complex; (2) there is a possibility of plausible deniability that deflects responsibility for an attack to an actor outside government control; (3) it helps states to

prolong the tense situation in conflicts by wearing down their adversary on a social, political and economic level; (4) it allows states to act outside domestic regulations and the criticism of opposing governmental sectors - or even public opinion itself in democracies; (5) it gives states speed and flexibility in responding to their adversaries' offensive actions, as it does not require technical evidence or public legitimation; (6) it offers states an additional tool of deterrence; (7) it allows states to circumvent the application of international law; (8) it facilitates the use of expert personnel without the need to offer legal recruitment; (9) it makes it possible to participate in international conflicts that would otherwise be economically and politically unmanageable (Torres Soriano, 2017; Expósito Guisado, 2024; Marín Gutiérrez, 2023).

However, achieving these benefits is not without its problems. In fact, the main attraction of using a *proxy* (which is none other than obtaining plausible deniability of an aggression) is also its main weakness, as anonymity and clandestinity dilute the coercive and dissuasive capacity of the sponsoring state - after all, we cannot ignore Clausewitz's theories that suggest that for one state to modify its conduct based on the will of another, the latter must know the origin of the coercive act suffered.

Another drawback of the use of *cyber proxies* lies in how the state selects and controls them when they are used. The existence of divergent interests between the two parties can lead to disloyalty on the part of the *proxy*, causing economic or political damage to the actor using them - a fact that is aggravated if we take into account that these *proxies* generally operate in areas where the state neither can nor wants to intervene.

The benefit of *proxies* lies in their ability to act covertly, although it is this very lack of transparency that limits the state sponsor in verifying their background and reliability. The academic literature highlights that control over *proxies* is further complicated if the state does not have effective mechanisms to sanction disloyalty, or if there are decentralised structures that prevent proper enforcement of hierarchical orders (Popovic 2015).

In this paper we will only focus on two actors that represent the two different poles (open activism and silent espionage) of the same phenomenon, but which are not so different in terms of the ends they pursue and the instrumentalisation of them by states.

Broadly speaking, hacktivism and APTs differ in motivation, methods and degree of state support. Thus, while hacktivism is driven by a social-political context (protest, activism, moral causes), APTs focus on strategic espionage and gaining an economic-military advantage.

Operationally, APTs act through stealth and persistence, employing custom *malware*, backdoors and lateral movement; unlike hacktivist actions that usually seek public attention and generally focus on short-term DDosS attacks.

However, it is not uncommon to observe how APTs temporarily act as hacktivists (when they publicly disclose the data they exfiltrate to provoke a political impact) and how hacktivists are instrumentalised by states to achieve their strategic ends.

At the organisational level, hacktivists and APTs also differ: hacktivists generally act decentralised, spontaneously, even anonymously, and without a unified command.

APTs, on the other hand, are usually structured teams, often integrated into a larger organisation (an army, intelligence agency or criminal group), with a defined hierarchy and considerably more powerful funding (CyberZaintza, 2021).

Indeed, the difference in resources and technical training suggests a closer link between APTs and states than hacktivist groups. However, the lines between the two concepts have recently been blurred by the realisation that some pro-Russian hacktivist groups have been receiving covert state support, or act in line with the state agenda, blurring the hitherto clear distinction between "activist *hackers*" and "state operatives" (Muncaster, 2024).

In fact, it cannot be ruled out that certain hacktivist groups are actually formed or backed by APTs or directly by state actors. One example is the "XakNet Team", "Infoccentr" and "CyberArmyofRussia_Reborn", pro-Russian hacktivist groups that, according to Mandiant, are cyberthreat actors sponsored by the Russian Main Intelligence Directorate (GRU) through the APT44 (Mandiant, 2022).

Over the last decade there have been multiple documented cases in which states have used both their own APT groups and hacktivist collectives (or their identities) to carry out cyber-espionage, conflict sabotage and political manipulation.

A paradigmatic example illustrating the interdependence of both concepts can be found in the 2016 US elections, when "*DCLeaks*" and "*Guccifer 2.0*", two identities linked to Russia's Main Intelligence Directorate (*Glavnoe Razvedyvatel'noe Upravlenie*, GRU), stole Democratic Party emails and disseminated them posing as "patriotic American hacktivists" (DOJ, 2018).

In the wake of the war in Ukraine, it is not uncommon to find interdependence between Russian hacktivists and APTs, groups such as *Killnet*, *NoName057(16)*, *Anonymous Sudan* that have attacked government websites and Western companies in support of the Kremlin's narrative show that, while these groups call themselves "spontaneous activists", they suspiciously act in coordination with Russian state action (Van Der Walt, 2025).

However, Russia is not the only state actor that employs APTs and hacktivists to deploy its power. Other states such as China, North Korea or Iran have also been accused for years of conducting their offensive activities in cyberspace in this way.

Specifically, China has been accused for years of sponsoring vast cyber espionage campaigns through military units and paid *hackers*, such as those of the APT1 group, considered by Mandiant in 2013 to be Unit 61398 of the Chinese People's Liberation Army.

Chinese APT operations tend to focus on strategic targets (aerospace, energy, telecommunications, defence, etc.) and are considered part of Chinese state intelligence, but unlike Russia, the use of hacktivism is not as prominent in Chinese strategies.

The government has tolerated and even inspired Chinese "patriotic *hackers*" in some conflicts, one example being the "*Honker Hacker Network*", a *hacker* community

outside government control - according to Chinese sources - that has attacked China's adversarial actors during territorial disputes or diplomatic incidents.

Iran, on the other hand, has shown a tendency to instrumentalise supposedly activist *hacker* groups to carry out retaliatory operations against its adversaries, while developing its own APTs. A significant example of this was the DDoS attacks against US banks in 2012-2013, in retaliation for Western sanctions: an entity claiming to be religious hacktivists and calling itself the '*Cyber Fighters of Izz ad-Din al-Qassam*' claimed credit for the offensive, citing outrage over an anti-Islamic video (CFR, 2012).

US intelligence agencies subsequently concluded that this was an operation orchestrated by Iran (probably its Revolutionary Guard) in response to measures taken against its nuclear programme. In fact, in 2016 the US Department of Justice indicted seven Iranians linked to the Islamic Revolutionary Guard Corps (IRGC) for these attacks.

Another example is the 2012 "*Shamoonj*" attack by the "*Cutting Sword of Justice*", an alleged hacktivist group that wiped data from 30,000 computers at the Saudi oil company Aramco, but which analysts later attributed to an Iranian state operation in response to the *Stuxnet* offensive and regional tensions.

North Korea, despite its isolation, has also managed to build one of the most active cyberthreats, mainly to raise funds and destabilise its geopolitical adversaries. Its most notable APT group, *Lazarus Group* (linked to APT38) has stolen hundreds of millions through attacks on banks.

Another case that illustrates the instrumentalisation of activist campaigns by states can also be found in one of their actions, the *hacking* of Sony Pictures in 2014, when a group called "*Guardians of Peace*" exfiltrated confidential data and destroyed Sony systems in apparent retaliation for the satirical film about the North Korean leader "*The Interview*". (FBI, 2014).

North Korea is the paradigm of direct instrumentalisation, its *hackers* are agents of the state who sometimes assume the names of fictitious groups to disseminate their messages or justify their attacks, but unlike other states, the North Koreans do away with the distinction between APT and state apparatus altogether, keeping the cover only in the public narrative to the outside world.

For their part, Western powers obviously also employ offensive cyber capabilities to attack other states. Perhaps the most relevant case is the 'Olympic Games' operation attributed to the NSA agencies and the (unofficially recognised) 8200 unit, in which the US and Israel developed the *Stuxnet* malware to sabotage Iran's nuclear centrifuges around 2010 (The Guardian, 2017).

However, in the West, although there are APT entities supported by states to act offensively in espionage campaigns, the instrumentalisation of hacktivist groups to hide their actions is practically non-existent. In fact, we can only find one case where a Western hacktivist group links its activity to the cyber offensive capacity of a state: the "*IT Army of Ukraine*".

This case is particularly controversial, as public state support by the Ukrainian government openly violates recently agreed norms on the conduct of states in cyberspace, as well as the foreign policy positions of NATO members (Healey and Grinberg, 2022).

If we use Healey and Grinberg's (2022) "Spectrum of Responsibility" table, where they correlate the activity of groups according to the degree of state responsibility for their *cyber proxy*, we can see how the Ukrainian government's support for the *IT Army of Ukraine* started at least as "state-coordinated (level 6)", (when Ukrainian Minister of Digital Transformation Mikhail Fedorov openly called on hacktivist volunteers from all over the world to support Ukraine on the digital front) and even "encouraged by the state (level 4)".

Table 1: *Spectrum of State responsibility.*

State position	*State-proxy* relationship
1. Banned by the State.	The national government will help stop a third party attack.
2. State ban but inadequate.	The national government cooperates, but it is unable to stop the attack by third parties.
3. Ignored by the state.	The national government is aware of the attacks by third parties, but is unwilling to take no official action.
4. State-sponsored.	Third parties control and direct the attack, but the national government promotes them as a political issue.
5. Shaped by the State.	Third parties control and lead the attack, and the The state provides some support.
6. Coordinated by the State.	The national government coordinates the attack by third parties, e.g. by suggesting details operational.
7. State-mandated.	The national government orders third parties to carry out the attack on their behalf.
8. Managed, but not recognised by the state.	Elements outside the control of the forces cybernetic attacks by the national government lead to orderly attack.
9. State-implemented.	The national government carries out the attack using cybernetic forces under their direct control.
10. State-integrated.	National government attacks using embedded proxies and cyber forces governmental.

(Healey, 2022).

It is especially in geopolitical conflicts that we see the most accelerated convergence between hacktivism and state operations. In the case of the Ukrainian war,

three years after the start of the conflict and despite the fact that the number of hacktivist actors has decreased considerably (from more than 130 groups in 2024 to only about 80 groups in 2025), we can still observe how both sides maintain a crossover of destructive cyberattacks, coordinated with their military campaign and supported in their actions by "patriotic hackers" (Cyberknow, 2025).

On the Ukrainian side, the *IT Army of Ukraine* remains Ukraine's most important hacktivist force, still mobilising volunteers inside and outside the country to attack Russian infrastructure, conduct counter-propaganda and support intelligence missions. In the period 2023-2024, it is credited, for example, with temporarily bringing down internet services in Russian-occupied areas and continuously deploying DDoS campaigns against high-profile Russian entities (Optiv, 2023).

On the pro-Russian side, the most prominent group at present is *NoName057(16)*, a group linked to the GRU, which acts in coordination with the Kremlin's agenda by selecting targets in tune with Russian strategic interests and considers itself a sort of permanent "cyber-spontaneous arm" of the Russian military.

Table 2: *Chronological cases of state instrumentalisation of hacktivism.*

Year	***State***	***Group hacktivist***	***Feature***	***Level of state linkage (Healey & Grinberg).***
1998-1999	Kosovo	Patriotic hackers	First conflict with notable hacktivist intervention.	Ignored / Spontaneous
1999	China	Red Honker	Patriotic hackers active in territorial conflicts. Industrial espionage campaigns and cyber-attacks on critical infrastructure.	Encouraged / Shaped
2012-2013	Iran	Cyber Fighters of Izz ad-Din al-Qassam	DDoS attacks on US banks in retaliation for sanctions. Operation Shamoon against Aramco with mass deletion.	Coordinated / Orderly
2014	North Korea	Lazarus Group	Cyber-attacks for state funding. Attack on Sony Pictures (2014) as symbolic retaliation.	Implemented / Integrated
2022-present	Russia	Killnet/ Cyber Army of Russian Reborn/ NoName057(16)	Hacktivist groups coordinated with the Russian strategy in the Ukrainian war. DDoS attacks.	Coordinated / Encouraged
2022-present	Ukraine	IT Army of Ukraine	Government's public call for hacktivism against Russia. DDoS, sabotage and pro-Ukrainian propaganda.	Coordinated / Encouraged

4. THE FUTURE OF HACKTIVIST GROUPS.

The survival of hacktivist groups indicates that war-integrated hacktivism is here to stay, at least for as long as the underlying conflict lasts and states in conflict find this layer of decentralised action useful. Moreover, the current hacktivist landscape leads us to observe that hacktivism is moving beyond DDoS and into more sophisticated APT attacks, such as attacks on critical infrastructure SCADA and industrial control systems (ICS) . [4]

The fact that groups belonging to the pro-Russian hacktivist ecosystem, such as *Z-Pentest Alliance* or *Sector 16*, have been actively intruding into power plants, drinking water facilities and industries in general, reflects not only a maturation and stateisation of the hacktivist phenomenon, but also the existence of increasingly physical risks of their actions (Antoniuk, 2024).

The reduction in the number of hacktivist groups in the pro-Russian environment suggests that the initial effervescence has given way to a natural selection process in which those groups with better support, organisation and protection survive. A phenomenon that translates into more effective and coordinated operations, but also more predictable as they are aligned with the Russian state agenda.

At the same time, the persistence of daily attacks indicates that low-intensity cyber warfare has become routine. Constant DDoS maintains psychological and propaganda pressure on target populations (daily reminders of the presence of conflict), while the adoption of *ransomware* and attacks on industries raises the potential for real damage to critical infrastructure, blurring the line between hacktivism and cyberterrorism - a fact that may ultimately lead to more forceful responses by victim states and the potential for escalation of conflict.

Another development of relevance is the remarkable development of emerging alliances between hacktivist causes that transcend the theatre of operations beyond Ukraine and involve third countries. One example is the recent alliance between pro-Russian and pro-Palestinian hacktivists, which unites seemingly distinct geopolitical causes under a common narrative of attacking the West.

The global tensions of 2024 (including the Gaza war) created a strange united front of hacktivists. Russian groups (especially *NoName057(16)*) began coordinating operations with Middle Eastern-linked collectives (such as *Mr. Hamza* or *Anonymous Guys*), and synchronised their attacks under the banner of the "*Holy League*" union against countries they perceived as shared adversaries, such as France.

This type of alliance is well known in Spain, and particularly by the Guardia Civil, since in July 2024, the institution was the direct target of a joint cyber-attack campaign, "*#FuckGuardiaCivil*", which responded to an initiative promoted by the group *NoName057(16),* to "take revenge on the Spanish authorities" who had arrested three people in Manacor (Mallorca), Huelva and Seville on suspicion of participating in

[4] Centralised system to monitor, control and collect data from processes and devices in real time.

cyberattacks against public entities and strategic companies in Spain and other NATO countries.

In fact, in April 2025, a new alliance was already registered, including the *Keymous+, Mr Hamza, Alixsec and NoName057(16)* groups, to attack Poland, Germany, France, Italy and Spain under the slogan "*Operation Hack For Humanity V2!*

In the case of Spain alone, on the first day of the "Operation Hack For Humanity V2!" campaign, more than 30 attacks on companies and government websites were registered, with *Mr Hamza*, *NoName057(16)*, *TwoNet* and *Keymous+* being the most active groups in the attack.

The frequency with which this convergence has been occurring in recent months shows that the phenomenon is becoming increasingly international and interconnected. The alliances between hacktivist groups have become mutually supportive, transcending the borders of the Russian-Ukrainian conflict with a single goal: to expand their actions towards the common Western enemy.

The fact that NATO countries such as France, Italy and Spain itself could become targets of Russian patriotic *hackers* could lead to an escalation of the conflict, especially if one of their attacks were to severely damage critical infrastructure, low-intensity cyber warfare could draw a stronger response than usual.

5. CONCLUSIONS

The analysis of hacktivism and its relationship with states shows that this phenomenon has evolved from digital protest to state instrumentalisation with geopolitical and strategic implications. The boundary between activism, cybercrime and state operations is increasingly blurred, especially in conflicts such as the war in Ukraine, where we have observed a growing instrumentalisation of hacktivist groups by government forces in the defence of their national interests.

Indeed, the conflict between Russia and Ukraine has marked a turning point in the use of cyberspace as a battleground, where both state and non-state actors have actively engaged in denial of service (DDoS) attacks, cyber espionage and sabotage of critical infrastructure.

This study, developed through the study of the most prominent cases on the international scene, has allowed us to establish a distinction between civic *hackers* and patriotic *hackers*. While the former embrace nihilistic or socially conflictive causes, the latter are used by states as a covert tool in international conflicts, which entails an externalisation of governmental cyber capabilities and offers a series of strategic advantages: plausible deniability of responsibility, prolongation of situations of tension or the reduction of political and economic costs.

In short, we could say that states have learned to exploit hacktivism as an additional weapon, either by pretending to be hacktivists in order to disinform or exfiltrate data or by encouraging their sympathisers to launch mass cyberattacks against their enemy.

However, this instrumentalisation poses serious challenges at the strategic level. The progressive sophistication of attacks that have moved from digital vandalism to more advanced operations against critical infrastructures only seriously increases the possibilities of retaliation by the affected states and increases the potential risk of escalation in asymmetric conflicts.

Moreover, the convergence between APTs and hacktivists calls into question existing international norms, as attacks perpetrated by *proxy* actors blur state responsibility and make it difficult to implement deterrence or direct retaliation. Especially as hacktivist collectives seem to be evolving towards a new landscape of alliances capable of bringing together hacktivist groups with diverse geopolitical agendas to attack Western countries.

State cyber security must adapt to a new reality in which hacktivist groups play a key role in the projection of state power. Western democracies, traditionally more reluctant to use such tactics, face the dilemma of how to respond effectively without compromising their values.

The current trend not only shows a clear evolution of hacktivism towards an increasing linkage with the state interests of the government that supports them, but also reinforces the idea that cyberspace will continue to become more important in future conflicts. Cases such as Russia, where groups like *Killnet* or *NoName057(16)* have claimed cyber operations coinciding with the Kremlin's geopolitical interests -especially during the war in Ukraine-, or Iran, with groups like *Tapandegan*, whose oppositionist rhetoric does not prevent suspicions of indirect coordination with state agendas, exemplify this drift, and demonstrate a progressive blurring between non-state and state actors in the digital sphere, where hacktivism ceases to be exclusively a form of citizen dissidence and becomes, in certain contexts, an informal tool for the projection of state power.

6. BIBLIOGRAPHICAL REFERENCES

Antoniuk, D. (2024). *Cybervolk: Hacktivists from India and Russia collaborate on ransomware attacks.* The Record. https://therecord.media/cybervolk-india-hacktivists-russia-ransomware

Antoniuk, D. (2024). *Cybervolk: Hacktivists from India and Russia collaborate on ransomware attacks.* The Record. https://therecord.media/cybervolk-india-hacktivists-russia-ransomware

Auty, C. (2004). Political hacktivism: Tool of the underdog or scourge of cyberspace? *Aslib Proceedings, 56*(4), 212-221.

Bartlett, J. (2015). *The dark net: Inside the digital underworld.* Melville House.

CFR (2012). *Denial of service attacks against U.S. banks in 2012-2013.* Council on Foreign Relations (CFR). https://www.cfr.org/cyber-operations/denial-service-attacks-against-us-banks-2012-2013

Coleman, G. (2014). *Hacker, hoaxer, whistleblower, spy: The many faces of Anonymous.* Verso Books.

Cyberknow (2025). *Russia-Ukraine war: Hacktivist update.* https://cyberknow.substack.com/p/russia-ukraine-war-hacktivist-update

CyberZaintza (2021). *APT Group.* https://www.ciberseguridad.eus/ciberglosario/grupo-apt

Dahan, M. (2013). Hacking for the homeland: Patriotic hackers versus hacktivists. *International Conference on Information Warfare and Security*, 51-VII. Academic Conferences International Limited. https://search.proquest.com/docview/1549245919

Denning, D. E. (2001). Activism, hacktivism, and cyberterrorism: The internet as a tool for influencing foreign policy. In J. Arquilla & D. Ronfeldt (Eds.), *Networks and netwars: The future of terror, crime, and militancy* (pp. 239-288). RAND Corporation.

DOJ (2018). *Grand jury indicts 12 Russian intelligence officers for hacking offenses related to the 2016 election.* U.S. Department of Justice. https://www.justice.gov/archives/opa/pr/grand-jury-indicts-12-russian-intelligence-officers-hacking-offenses-related-2016-election

Expósito Guisado, J. (2023). *Cyberproxies: APTs as a future risk factor.* Spanish Institute for Strategic Studies (IEEE). *IEEE Bulletin*, (32), 815-831.

FBI (2014). *Update on Sony Investigation.* Federal Bureau of Investigation (FBI), Washington, D.C. https://www.fbi.gov/news/press-releases/update-on-sony-investigation

Fox, A. C. (2019). *Conflict and the need for a theory of proxy warfare. Journal of Strategic Security, 12*(1), 44-71. JSTOR. www.jstor.org/stable/26623077

Goode, L. (2015). Anonymous and the political ethos of hacktivism. *Popular Communication, 13*(1), 74-86. https://doi.org/10.1080/15405702.2014.978000

Green, K. (2016). People's war in cyberspace: Using China's civilian economy in the information domain. *Military Cyber Affairs, 2*(1). https://doi.org/10.5038/2378-0789.2.1.1022

Healey, J., & Grinberg, A. (2022). *Patriotic hacking: No exception*. Lawfare. https://www.lawfaremedia.org/article/patriotic-hacking-no-exception

Hern, A. (2017). NSA contractor leaked US hacking tools by mistake, Kaspersky says. *The Guardian.* https://www.theguardian.com/technology/2017/oct/26/kaspersky-russia-nsa-contractor-leaked-us-hacking-tools-by-mistake-pirating-microsoft-office

Hunsinger, J., & Schrock, A. (2016). The democratization of hacking and making. *New Media & Society, 18*(4), 535-538. https://doi.org/10.1177/1461444816629466

Johnson, P., & Robinson, P. (2014). Civic hackathons: Innovation, procurement, or civic engagement? *Review of Policy Research, 31*(4), 349-357. https://doi.org/10.1111/ropr.12074

Jordan, T. (2002). *Activism! Direct action, hacktivism and the future of society*. Reaktion Books.

Jordan, T., & Taylor, P. A. (2004). *Hacktivism and cyberwars: Rebels with a cause?* Psychology Press.

Mandiant (2022). *GRU's rise: Telegram minions.* https://cloud.google.com/blog/topics/threat-intelligence/gru-rise-telegram-minions

Marín, F. (2023). *Hacktivism in the service of the state: cyberproxies in Ukraine.* Opinion Paper. Spanish Institute for Strategic Studies (IEEE).

Maurer, T. (2018). *Cyber Mercenaries: The state, hackers, and power*. Cambridge University Press.

Muncaster, P. (2024). *Hacktivism: Evolving threats to organisations*. WeLiveSecurity. https://www.welivesecurity.com/es/cibercrimen/el-hacktivismo-evolucionando-amenazas-organizaciones

Olson, P. (2012). *5 things every organization can learn from Anonymous*. Forbes. http://www.forbes.com/sites/parmyolson/2012/06/05/5-things-every-organization-can-learn-from-anonymous/

OPTIV (2023). *Russia/Ukraine Update - December 2023*. https://www.optiv.com/insights/discover/blog/russiaukraine-update-december-2023

Popovic, M. (2015). Fragile proxies: Explaining rebel defection against their state sponsors. *Terrorism and Political Violence*. https://doi.org/10.1080/09546553.2015.1092437

Rondeaux, C., & Sterman, D. (2019). *Twenty-first century proxy warfare: Confronting strategic innovation in a multipolar world since the 2011 NATO intervention*. New America. https://d1y8sb8igg2f8e.cloudfront.net/documents/Twenty-First_Century_Proxy_Warfare_Final.pdf

Sauter, M. (2013). "LOIC will tear us apart": The impact of tool design and media portrayals in the success of activist DDOS attacks. *American Behavioral Scientist, 57*(7), 983-1007. https://doi.org/10.1177/000276

Schrock, A. R. (2016). Civic hacking as data activism and advocacy: A history from publicity to open government data. *New Media & Society, 18*(4), 581-599. https://doi.org/10.1177/1461444816629469

Torres Soriano, M. (2017). Proxy wars in cyberspace. *Revista del Instituto Español de Estudios Estratégicos, (9)*, 15-36.

(2018). Hacktivism as a communication strategy from Anonymous to the cybercaliphate. *Cuadernos de Estrategia, (197)*, 197-224.

Van Der Walt (2025). *Reflecting on three years of cyber warfare in Ukraine. ComputerWeekly.* https://www.computerweekly.com/opinion/Reflecting-on-three-years-of-cyber-warfare-in-Ukraine

Vegh, S. (2005). *The media's portrayal of hacking, hackers, and hacktivism before and after September 11. First Monday.* http://uncommonculture.org/ojs/index.php/fm/article/view/1206/1126

ORCID: 0009-0003-4977-3899