Taylor & Francis
Taylor & Francis Group

Critical Studies in Media Communication

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rcsm20

"Hacking and information disorder: the weaponization of leaking"

Philip Di Salvo

To cite this article: Philip Di Salvo (07 Mar 2025): "Hacking and information disorder: the weaponization of leaking", Critical Studies in Media Communication, DOI: 10.1080/15295036.2025.2465693

To link to this article: https://doi.org/10.1080/15295036.2025.2465693

Published online: 07 Mar 2025.

CRITICAL STUDIES IN MEDIA COMMUNICATION
https://doi.org/10.1080/15295036.2025.2465693

DISCUSSION

OPEN ACCESS Check for updates

"Hacking and information disorder: the weaponization of leaking"

Philip Di Salvo

School of Umanities and Social Sciences, Universität St. Gallen, St. Gallen, Switzerland

ABSTRACT

This essay examines the dual role of leaks as both a journalistic tool and a vehicle for propaganda, with a focus on state-sponsored hack-and-leak operations. While leaks have traditionally served investigative journalism, the rise of cyberattacks and disinformation has redefined their role within the broader context of information disorder. Hack-and-leak operations leverage authentic documents obtained through cyberattacks, releasing them tactically to shape media narratives, foster doubt, and undermine targets. Unlike conventional disinformation campaigns, these operations weaponize real information, aligning with the concept of "malinformation."[1] State-sponsored hackers play a critical role by introducing "tainted leaks" (Hulcoop, A., Scott-Railton, J., Tanchak, P., Brooks, M., & Deibert, R. (2017). *Tainted leaks: Disinformation and phishing with a Russian nexus*. The Citizen Lab, Munk School of Global Affairs & Public Policy & University of Toronto. https://utoronto.scholaris.ca/server/api/core/bitstreams/d45f78d4-e53e-409a-b81f-5bba4689b442/content) that blend genuine and falsified content to deceive the public. Journalists face ethical challenges when covering such leaks, as they must balance public interest with the risk of amplifying malicious agendas. In essence, hack-and-leak operations threaten journalistic credibility, information integrity, and democratic processes. By exploiting the perceived "authenticity" of leaked materials, these operations erode trust in media and public institutions. This essay underscores the importance of understanding the nuanced relationship between "real" and "fake" in leak-driven disinformation, highlighting the need for strategies to mitigate their broader impact on the digital public sphere.

ARTICLE HISTORY
Received 13 December 2024
Accepted 6 February 2025

KEYWORDS
Hacking; leaks; information disorder

Leaks have become increasingly central to the dynamics of the digital public sphere, emerging from the actions of various actors motivated by different agendas and aims (Di Salvo, 2024). On one hand, leaks have sparked crucial journalistic investigations, inspired new journalistic genres, and become normalized sources for journalists (Lee,

CONTACT Philip Di Salvo philip.disalvo@unisg.ch MCM - HSG Blumenbergplatz 9, CH-9000, St.Gallen

2023). On the other hand, leaks have also become a common strategy associated with influence campaigns, state-sponsored hacking, and the cyber tactics of contemporary warfare (Shires, 2019). Leaks have also been employed as a strategy within information disorder, particularly in "malinformation" actions (Wardle & Derakhshan, 2017), often with explicit propagandistic aims. The aim of this essay is to discuss how leaks can become effective propaganda tools when they result from state-sponsored cyber actions targeting individuals and organizations during tense political events, such as elections. Viewed in this light, leaks reveal that the interplay between "fake" and "real" within information disorder and propaganda is not necessarily straightforward. This poses new challenges to journalists and other actors involved in the public sphere, affecting various areas such as journalistic authority, cyber influence, hacking, and informational chaos.

Leaks have long been part of the information ecosystem, predating digitalization. The impact of leaks driven by whistleblowers is well-documented in the history of journalism and political communication, where they have been some of the most powerful and consequential actions in this domain (McCurdy, 2013). With the influx of digitalization, new and more numerous actors have become involved in leaking, both as perpetrators and as disseminators, thereby expanding the reach of digital leaks far beyond whistleblowing (Di Salvo, 2024).

While the rise of leak-centered organizations such as WikiLeaks—which focuses on the publication of leaked materials—has been widely analyzed, the role of hackers in the procurement and dissemination of leaks remains relatively underexplored. Hackers from diverse backgrounds, often wearing different "hats" (Jaquet-Chiffelle & Loi, 2020) have incorporated leaking into their actions by cyberattacking organizations to obtain information and make it public. These hack-and-leak operations involve two distinct stages: the acquisition of data and content through cyberattacks ("hack") and its dissemination through various strategies ("leak").

At the heart of these practices are the "political motivations" (Shires, 2019) and an intrinsic "public" dimension (Coleman, 2017). Hackers affiliated with state actors or intelligence agencies are among the most active groups in this field, often targeting governments, companies, political parties, and individuals (Coleman, 2017).

The very existence or announcement of the existence of a leak of inner information can have a significant impact on an organization's public image and reputation, causing severe informational chaos. Not surprisingly, emails have become the preferred "loot" of hack-and-leak operations due to their private nature and their ability to expose the internal workings of affected organizations or individuals, as well as to reveal potentially controversial personal details and communications (Hosaka, 2025).

Overall, hack-and-leak operations are increasingly becoming one of the preferred "active measures"[2] of contemporary influence campaigns, particularly because of their capacity to cause damaging exposure in already tense situations, such as elections, scandals, or controversies (Rid, 2020, pp. 3–16). This raises a critical question: what makes the exposure of "real" information—such as hacked emails—a propaganda or disinformation operation? The answer lies not in the content of the leaks themselves, but in the potential informational chaos they generate, as well as the "weaponization" that typically follow these unauthorized releases. Moreover, the impact of hack-and-leak operations is often amplified by ambiguity surrounding the source and credibility of the information (Shires, 2019). At the same time, hack-and-leak campaigns by state actors can be

accompanied by other propaganda strategies, such as social media campaigns or attention-seeking tactics aimed at journalists who might amplify specific messages or cover the leaked materials in sensationalistic ways. Moreover, the difficulty of attribution associated with this type of leak is another potential source of confusion, which can even spill over into the realm of conspiracy (Briant, 2023).

Additionally, hack-and-leak operations also give rise to "competing interests" (Briant & Wanless, 2018) shared by the leak perpetrators and the journalists tasked with reporting on them. The friction between these interests further contributes to the informational chaos directly sought by hackers. This dynamic is linked to the nature of leaks themselves, as they involve "real" materials and content that may (or may not) contain newsworthy information with public interest value, despite their malicious origins and the agendas involved.

In this sense, state-sponsored hack-and-leak diverge from other forms of propaganda, as they do not necessarily rely on forgery, the creation of fake content, or the use of deceptive distribution channels. Instead, when state actors engage in information disorder through leaks, they weaponize a communication and media practice more commonly associated with whistleblowing, hacktivism, and investigative reporting. According to Wardle and Derakhshan's framework (2017), the distinctive nature of hack-and-leak operations indeed aligns with the concept of "malinformation," which involves "information that is based on reality but is used to inflict harm on a person, organization, or country" (2017).

The rationale behind these operations is to employ them as "mechanisms for delegitimization," regardless of the veracity of the content being exposed (Shires, 2019). As research has shown, U.S. media outlets themselves have provided extensive coverage and hype to certain leaks, thereby enabling influence campaigns to shape the media agenda (Watts & Rothschild, 2017). In this context, the disinformation associated with leaks arises also from the excessive media coverage and hype that unauthorized disclosures can generate, potentially leading to the creation of conspiracy theories and amplifying doubt and mistrust. This is not the space for an in-depth discussion of the ethical and editorial questions that pertain to the journalistic field. However, journalists and journalism scholars have debated how the press should report on the newsworthy elements of state-sponsored leaks without becoming unwitting amplifiers of the hackers' agendas (Tofel, 2024) or even "puppets", as cyber journalist Joseph Cox put it (2017). There is no consensus on how to achieve this goal, as the issue is deeply tied to differences in journalistic cultures and newsroom policies, which vary significantly. A fundamental issue is certainly the need to differentiate between the coverage of whistleblowing cases and this type of leak (Briant, 2023). As a matter of principle, even when journalists choose to cover certain information emerging from hack-and-leak operations, these cases should be treated for what they are. While some of the disclosed information may have public value, the motivations of the hackers who made it available do not.

Moreover, it is important to note that some hack-and-leak operations have also incorporated fake elements. This was particularly evident during the Clinton campaign and the Democratic Party leaks that occurred in the context of the 2016 U.S. Presidential election, which were carried out by hackers linked to the Russian intelligence agencies.

In this case the Russian hackers behind the operation created fake news websites to release the information they obtained through cyberattacks. They also used the fake

persona "Guccifer 2.0," who presented themselves as an independent Romanian hacker claiming responsibility for the hack-and-leak operation. This persona played a crucial role in transferring the stolen information to WikiLeaks and U.S. journalists, for instance. The creation of fake personas like Guccifer 2.0 was a deliberate strategy to conceal the real identities and intentions of the perpetrators. Guccifer 2.0 also served as the public face of the operation, facilitating outreach to journalists and news organizations to promote the leaks while maintaining the appearance of legitimacy.

Other hack-and-leak operations have demonstrated even more explicit layers of forgery. In 2017, individuals connected with Emmanuel Macron's political party, En Marche!, were hacked, and their emails were subsequently leaked in an attempt to undermine Macron's campaign just days before the French general election. The influence campaign—whose attribution remains inconclusive, although most available evidence points once again to Russia—ultimately failed for various reasons, as it did not achieve significant coverage or impact in France (Soesanto, 2017).

This hack-and-leak operation occurred within the broader context of a disinformation campaign, which also spread rumors about the existence of offshore bank accounts linked to Macron himself (BBC News, 2017). The leaked documents did not contain any evidence to support these claims—or any other significant newsworthy revelations. However, fake and photoshopped documents were reportedly included within the leaked files, along with others that appeared to be clearly forged or manipulated (Soesanto, 2017). The combination of attempts to fuel controversy around the election, the release of a leak with limited substance, and the presence of fake documents highlights the strategic goals of hack-and-leak perpetrators. Instances where altered documents and forged information are inserted into leaked document sets have been termed "tainted leaks" by a research report from The Citizen Lab. The report investigated a Russian hack-and-leak operation targeting journalist and prominent Kremlin critic David Satter (Hulcoop et al., 2017).

In this case, hacked emails were released by CyberBerkut, a group linked to the hacking collective Fancy Bear or APT28, which is widely believed to be part of Russian intelligence. One of the leaked emails appeared to have been deliberately modified by the perpetrators to falsely link Satter to "U.S. efforts to meddle in Russian politics and even to inspire a popular revolution" (Hulcoop et al., 2017). The clear aim of this forgery was to construct a weaponized narrative, using the leak as a delivery mechanism to instill disinformation into the public sphere.

By definition, leaks involve the unauthorized dissemination of information not intended for public access, a dynamic that can either serve investigative journalism and activism in the public interest or, as we have discussed here, support more controversial interests. The purpose of certain hack-and-leak operation has been describe as "simulation of a scandal" to reflect those instance where the goal are attempts to "direct public moral judgment against their target" (Shires, 2020) through doxing, which can interest individuals or also be "political" or organizational (Schneier, 2015 as quoted in Shires, 2020). Besides the obvious negative implications for the affected organizations, which are not targeted by actors aiming to serve transparency or the public interest, the weaponization of leaking also has broader systemic effects on the media ecosystem and ultimately on democratic processes. Toxic operations conducted by state actors could cast a negative light on whistleblowing or other types of leaks

that occur in the public interest or provide genuine democratic benefits, pushing for the silencing of the voices (Briant, 2023). Additionally, hack-and-leak operations can also target individuals—particularly journalists (Waisbord, 2020)—with the aim of exposing their personal information. Although this article has primarily focused on cases related to the U.S. and Europe, state-sponsored hack-and-leak operations are not limited to this geographical area. As highlighted in various analyses (Coleman, 2017; Shires, 2020), conflict zones such as the Middle East and Ukraine are also affected. Cyber activities from China, North Korea, and Iran, among others, should also be taken into account, always keeping in mind that leaking operations are often not clearly separate from more conventional cyberattacks. The implications of this phenomenon should indeed be regarded as global.

In more general terms, the goal of influence campaigns is to disseminate mistrust. In doing so, these leaks also serve as delegitimizing actions against journalism and other democratic processes. By giving leaks a different meaning, hack-and-leak operations aim to sow doubt not only about the affected organizations but also about the media organizations and journalists who are supposed to report on them, thereby dishonestly shaping public opinion.

Notes

1. "Malinformation" has been defined as "information that is based on reality, used to inflict harm on a person, organization or country" (Wardle & Derakhshan, 2017).
2. "Active measures" are covert operations carried out by intelligence agencies to shape events and sway public perception in foreign nations. These efforts involve various strategies, such as spreading false information, propaganda, fabrications, and leaks, among others. See Rid (2020) for an overview of their historical evolution.

Disclosure statement

No potential conflict of interest was reported by the author.

References

BBC News. (2017, May 4). French election: Macron takes action over offshore claims. *BBC News*. https://www.bbc.com/news/world-europe-39802776

Briant, E. L. (2023). Hack attacks. How cyber intimidation and conspiracy theories drive the spiral of "secrecy hacking". In J. Steel & J. Petley (Eds.), *The Routledge companion to freedom of expression and censorship* (pp. 285–296).

Briant, E. L., & Wanless, A. (2018). A digital ménage à trois: Strategic leaks, propaganda and journalism. In C. Bjola & J. Pamment (Eds.), *Countering online propaganda and extremism* (pp. 44–65). Routledge.

Coleman, E. G. (2017). The public interest hack. *limn*, *1*(8), 18–23.

Cox, J. (2017, May 5). How to report on a hack without becoming a puppet. *Motherboard*. https://www.vice.com/en/article/how-to-report-on-a-hack-without-becoming-a-puppet/

Di Salvo, P. (2024). A typology of digital leaks as journalistic source materials. In H. S. Dunn, M. Ragnedda, M. L. Ruiu, & Laura Robinson (Eds.), *The palgrave handbook of everyday digital life* (pp. 469–488). Springer International Publishing.

Hosaka, S. (2025). Leaked email data: A new source for the study of authoritarian regimes. *Digital War*, *6*(1), 1. https://doi.org/10.1057/s42984-024-00097-w

Hulcoop, A., Scott-Railton, J., Tanchak, P., Brooks, M., & Deibert, R. (2017). *Tainted leaks: Disinformation and phishing with a Russian nexus*. The Citizen Lab, Munk School of Global Affairs & Public Policy & University of Toronto. https://utoronto.scholaris.ca/server/api/core/bitstreams/d45f78d4-e53e-409a-b81f-5bba4689b442/content

Jaquet-Chiffelle, D. O., & Loi, M. (2020). Ethical and unethical hacking. In M. Christen, B. Gordijn, & M. Loi (Eds.), *The ethics of cybersecurity* (pp. 179–204). Springer Internationl Publishing. https://link.springer.com/chapter/10.1007978-3-030-29053-5_9

Lee, M. (2023). *Hacks, leaks, and revelations. The art of analyzing hacked and leaked data*. https://hacksandleaks.com/

McCurdy, P. (2013). From the Pentagon papers to cablegate: How the network society has changed leaking. In B. Brevini, A. Hintz, & P. McCurdy, *Beyond WikiLeaks implications for the future of communications, journalism and society* (pp. 123–145). Springer International Publishing.

Rid, T. (2020). *Active measures. The secret history of disinformation and political warfare*. Profile Books.

Schneier, B. (2015, July 10). Organizational doxing. *Schneier on Security*. https://www.schneier.com/blog/archives/2015/07/organizational_.html

Shires, J. (2019). Hack-and-leak operations: Intrusion and influence in the Gulf. *Journal of Cyber Policy*, *4*(2), 235–256. https://doi.org/10.1080/23738871.2019.1636108

Shires, J. (2020). The simulation of scandal: Hack-and-leak operations, the Gulf States, and U.S. Politics. *Texas National Security Review*, *3*(4), 10–29.

Soesanto, S. (2017, May 9). The Macron leak that wasn't. *European Council on Foreign Relations*. https://ecfr.eu/article/commentary_the_macron_leak_that_wasnt_7285/

Tofel, R. J. (2024, August 29). On hacked documents, journalism and the motives of sources. *Ethics & Journalism*. https://ethicsandjournalism.org/2024/08/29/on-hacked-documents-journalism-and-the-motives-of-sources/

Waisbord, S. (2020). Mob censorship: Online harassment of US. *Journalists in Times of Digital Hate and Populism. Digital Journalism*, *8*(8), 1030–1046.

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (Vol. 27, pp. 1–107). Council of Europe.

Watts, D. J., & Rothschild, D. M. (2017, December 5). Don't blame the election on fake news. Blame it on the media. *Columbia Journalism Review*. https://www.cjr.org/analysis/fake-news-media-election-trump.php.