Pharos Series: Report No. 2

April 2025

The Ransomware Trust Paradox

Max Smeets

virtual routes

PHAROS SERIES

Virtual Routes | www.virtual-routes.org

Design & Layout by Frank Wo | Cover by Vahram Muradyan | Edited by Katharine Khamhaengwong

Pharos Series: Report No. 2 April 2025

The Ransomware Trust Paradox

Max Smeets

virtual routes

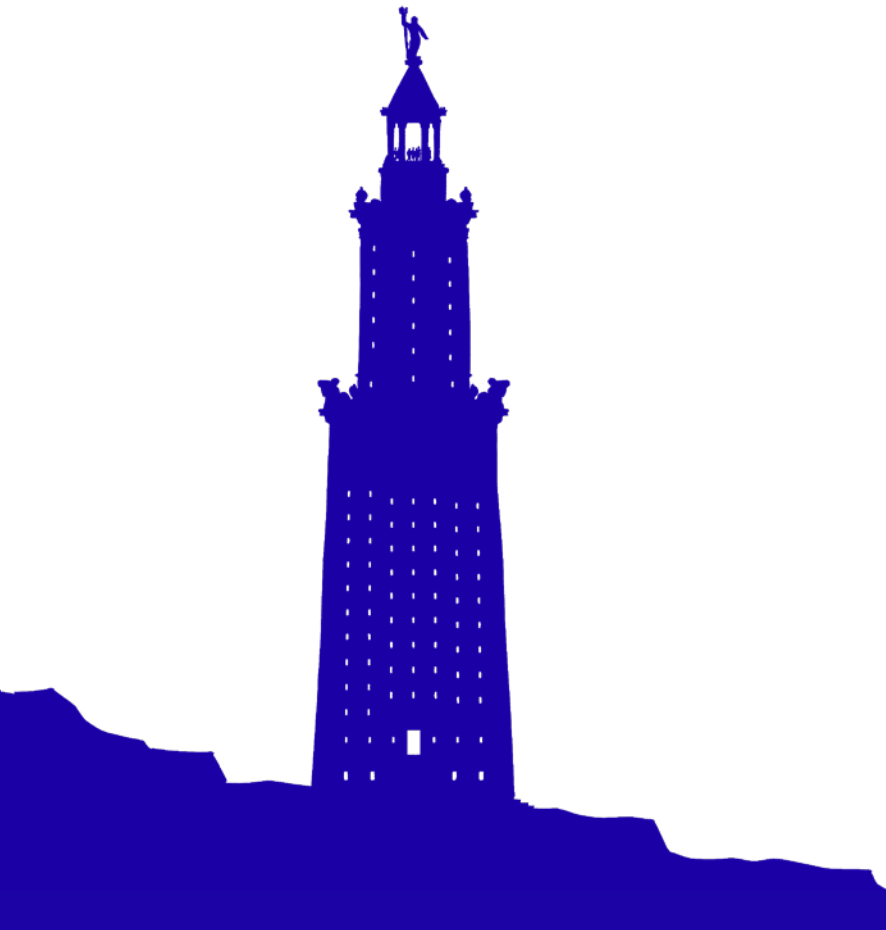
About the Author

Max Smeets

Max Smeets is the Co-Director of Virtual Routes and serves as Managing Editor of Binding Hook. He also holds research positions at ETH Zurich, the Royal United Services Institute (RUSI), and Stanford University's Center for International Security and Cooperation. Max is the author of *Ransom War: How Cyber Crime Became a Threat to National Security* and *No Shortcuts: Why States Struggle to Develop a Military Cyber Force*.

Max received a BA in Economics, Politics, and Statistics from University College Roosevelt, Utrecht University, and an MPhil (Brasenose College) and DPhil (St. John's College) in International Relations from the University of Oxford.

Table of Contents



Introduction

On May 8, 2022, Costa Rican President Rodrigo Chaves Robles issued an executive order declaring a national emergency and pronouncing that the country was in a 'state of war'. In the preceding weeks, a group known as Conti had launched a series of devastating ransomware attacks, encrypting data from several ministries and threatening to release the information online if the government did not comply with their demands.

> This executive order marked the first time in history that a country formally declared a state of emergency in direct response to ransomware, or, in fact, any cyberattack.

The gravity of the response is illustrative of the rising threat of ransomware to national security. After Conti's disruptive activities in Costa Rica, we saw the Cuba ransomware group attack Montenegro's Department for Public Relations in August 2022, Quantum attack the Dominican Agrarian Institute the same month, and RansomHouse attack Colombian government ministries in September 2023. We have also seen major supply-chain attacks, such as REvil's ransomware attack on 1,500 companies through Kaseya, a company that provides IT management software to managed service providers, and critical infrastructure attacks, like Darkside's ransomware attack on the Colonial Pipeline, the largest pipeline system for refined oil products in the United States, which disrupted fuel supplies across the East Coast.

Unfortunately, it is often the most vulnerable who are victims of ransomware attacks. One chilling example took place in March 2023, when Lehigh Valley Health Network, a healthcare network in Pennsylvania, was targeted by a criminal ransomware group. When the healthcare organization refused to pay to have their data decrypted, the hackers resorted to a despicable act, leaking personal data including photos of topless female breast cancer patients.

Despite its increasingly high profile, the current public understanding of ransomware, as well as the government's approach to countering it, is misguided. At the heart of this issue is a fundamental aspect of ransomware groups that has been ill-understood: their need to overcome what I term the 'Ransomware Trust Paradox'.

> Despite their inherently deceptive activities – breaking into systems, stealing data, and encrypting vital information – ransomware groups must convince their victims of their trustworthiness. This trust encompasses not just the promise not to release the stolen data but also the assurance that payment will result in the decryption of the affected systems.

This trust is crucial; no organisation wants to pay a lot of money and then get nothing in return or receive a broken tool to unlock their files. Similarly, no organisation wants to make a payment only to discover that their data has still been released.

This paradox sets ransomware groups apart from state hacking groups. State actors – especially intelligence agencies – aim for secrecy and ambiguity, avoiding detection and typically denying attribution if uncovered. Conversely, ransomware groups start covertly but later embrace self-attribution to enhance their brand. Unlike nation-state hacking groups, who shun publicity, ransomware groups often benefit from media exposure, using it to strengthen their reputation within the cybercriminal ecosystem and to the wider public.

Recognising this difference necessitates a fundamental shift in counter-ransomware policy. Ransomware is often viewed as a subset of broader cybersecurity policy challenges, requiring general responses like disruption, information sharing, and public-private resilience. While these approaches are essential, ransomware also shares significant parallels with disinformation and terrorism, in that the public perception and media narratives surrounding it are critical to how the threat evolves. This suggests the need for more targeted actions, including establishing a code of ethics for ransomware reporting and providing journalist training, to ensure responsible coverage and reduce the reputation and influence of these criminal groups.

Establishing Trust

There are three foundational types of trust underpinning relationships: identification-based, deterrence-based, and knowledge-based.[1] The most successful ransomware groups primarily use the latter two types to navigate the 'Ransomware Trust Paradox'.

Identification-based trust reflects the highest level of trust, where one party deeply understands and aligns with the other's values and preferences. A common example of this is a tribe, where members often show greater trust towards fellow tribesmen than outsiders, attributed to common experiences and values.[2] Shared objectives, closeness, and similar values foster trust in such environments.

For any ransomware group, establishing such profound trust with their victims is nearly impossible. Some groups have tried; for example, the LostTrust ransomware group claims to be 'specialists in the field of network security with at least 15 years of experience,' who have turned to ransomware due to poor pay for their legitimate services.[3] They aim to cultivate trust by projecting an image of professionalism and shared experience.

Deterrence-based trust arises in scenarios where the potential costs of ending a relationship or the threat of retaliation outweigh the (short-term) benefits of acting deceitfully.[4] Nobel laureate Thomas Schelling demonstrated that repeated interactions between two parties can establish a pattern of expected behaviour. Parties are less likely to deceive each other in a single transaction if future, beneficial transactions are anticipated.[5]

Another mechanism facilitating deterrence-based trust is reputational 'hostage taking'.[6] During the Middle Ages, a lord might take another lord's only child as a means of ensuring trustworthiness. The potential loss of a valued child often sufficed to ensure that the lord adhered to agreements.[7] In today's context, consider a well-known online retailer recognized for their distinctive and high-quality products. Such a retailer would steer clear of deceiving customers, concerned that negative feedback could damage their online reputation and dissuade potential buyers.[8]

---

[1] These forms of trust are not mutually exclusive. Debra L. Shapiro, Blair H. Sheppard, Lisa Cheraskin, 'Business on a Handshake,' *Negotiation Journal*, 8:4 (1992):365-377

[2] Ibid.

[3] Lawrence Abrams, 'Meet LostTrust ransomware — A likely rebrand of the MetaEncryptor gang,' *Bleeping Computer*, October 1, 2023, https://www.bleepingcomputer.com/news/security/meet-losttrust-ransomware-a-likely-rebrand-of-the-metaencryptor-gang/

[4] Shapiro, Sheppard, Cheraskin, 'Business on a Handshake'

[5] Thomas C. Schelling, *Strategy of Conflict*, (Harvard University Press 1960).

[6] For a general discussion on the key differences between ransomware attacks and hostage-taking see; Max Smeets, 'Why hostage negotiation tactics don't work on ransomware,' Binding Hook, November 28, 2024, Https://bindinghook.com/articles-binding-edge/why-hostage-negotiation-tactics-dont-work-on-ransomware/

[7] Shapiro, Sheppard, Cheraskin, 'Business on a Handshake'

[8] Ibid.

Ransomware group's interactions with victims are typically singular events rather than evolving relationships with repeated interactions. Unlike businesses that build relationships over time and create a foundation of trust through repeated positive interactions, ransomware groups typically have a one-time interaction with their victims. This sets aside the rare scenario where victims are 'ransomwared' continuously by the same group, building a pattern over the years.

> However, ransomware recovery companies, negotiators, and insurance companies do have such repeated interactions. These entities often find themselves dealing with the same ransomware groups repeatedly on behalf of different victims.

This repeated interaction can influence the dynamics between these intermediaries and ransomware operators, potentially impacting negotiation processes and outcomes.

As such, trust through reputational hostage taking is important for ransomware groups. Each time they interact with a victim, they are negotiating not just for that particular ransom but also for their reputation. Their reputation acts as a 'hostage'. If they fail to uphold their end of the bargain, they risk damaging their reputation, which deters future potential victims from trusting and engaging with them.

Some criminal groups are very explicit about the importance of their reputation. For example, Darkside states in their ransom note: 'We value our reputation. If we do not do our work and liabilities, nobody will pay us. This is not in our interests.'[9] Similarly, Karma writes in their message to victims: 'Decryption is only possible with a private key that only we possess. Our group's only aim is to financially benefit from our brief acquaintance, this is a guarantee that we will do what we promise. Scamming is just bad for business in this line of work.'[10]

[9] See Darkside's ransom note: ransomware_notes/darkside/darkside.txt at main · ThreatLabz/ransomware_notes · GitHub

[10] See Karma's ransom note: ransomware_notes/karma/KARMA-ENCRYPTED.txt at main · ThreatLabz/ransomware_notes · GitHub

Spotlight Case Study: **DarkSide Ransom Note**

```
---------- [ Welcome to DarkSide ] ------------->

What happend?
---------------------------------------------
Your computers and servers are encrypted, backups are deleted. We use strong encryption
algorithms, so you cannot decrypt your data.
But you can restore everything by purchasing a special program from us - universal
decryptor. This program will restore all your network.
Follow our instructions below and you will recover all your data.

What guarantees?
---------------------------------------------
We value our reputation. If we do not do our work and liabilities, nobody will pay
us. This is not in our interests.
All our decryption software is perfectly tested and will decrypt your data. We will
also provide support in case of problems.
We guarantee to decrypt one file for free. Go to the site and contact us.

How to get access on website?
 ---------------------------------------------
Using a TOR browser:
1) Download and install TOR browser from this site: https://torproject.org/
2) Open our website: http://dark24zz36xm4y2phwe7yvnkkkkhxionhfrwp67awpb3r3bdcneivoqd.
onion/ZWQHXVE7MW9JXE5N1EGIP6IMEFAGC7LNN6WJCBVKJFKB5QXP6LUZV654ASG7977V

When you open our website, put the following data in the input form:
Key:

[snip]

!!! DANGER !!!
DO NOT MODIFY or try to RECOVER any files yourself. We WILL NOT be able to RESTORE
them.
!!! DANGER !!!
```

Some groups even issue 'press releases' to correct mistakes made by journalists that might affect their reputation. In one Snatch group release, they state:

> 'First of all we have nothing to do with the Snatch ransomware project that appeared in 2019 and existed for about 2 years. We are the Security Notification Attachment (SNAtch for short) Team, a group specializing exclusively in leaked sensitive data. We don't deal with locking companies or critical infrastructure, we don't aim to stop a company from operating by attacking it with software that blocks the control servers. If journalists analyze our work carefully, they will see that not a single client of ours has been attacked by a malware that can be called Snatch. [...] So the main thing that we want to say and convey to you is that the Security Notification Attachment Team (SNAtch for short) has nothing to do with the Snatch ransomware project.'[11]

Another example is BlackCat, also known as ALPHV, which emerged in late 2021. They released a 1,300-word article on their leak site titled 'Statement on MGM Resorts International: Setting the Record Straight.' In this, they rebuked several publications, including *VXUnderground* and *TechCrunch*, for failing to verify sources and disseminating inaccurate information.[12]

[11] Dissent, 'At some point, SNAtch Team stopped being the Snatch ransomware gang. Were journalists the last to know?' *Databreaches*, September 1, 2023.

[12] Sophox X-Ops, 'Press and pressure: Ransomware gangs and the media,' *Sophos News*, December 13, 2023, https://news.sophos.com/en-us/2023/12/13/press-and-pressure-ransomware-gangs-and-the-media/

Spotlight Case Study: **Excerpt Statement BlackCat/ ALPHV on MGM Resorts International**

Statement on MGM Resorts International: Setting the record straight

9/14/2023, 7:46:49 PM

We have made multiple attempts to reach out to MGM Resorts International, “MGM”. As reported, MGM shutdown computers inside their network as a response to us. We intend to set the record straight.

No ransomware was deployed prior to the initial take down of their infrastructure by their internal teams.

MGM made the hasty decision to shut down each and every one of their Okta Sync servers after learning that we had been lurking on their Okta Agent servers sniffing passwords of people whose passwords couldn’t be cracked from their domain controller hash dumps. Resulting in their Okta being completely locked out. Meanwhile we continued having super administrator privileges to their Okta, along with Global Administrator privileges to their Azure tenant. They made an attempt to evict us after discovering that we had access to their Okta environment, but things did not go according to plan.

On Sunday night, MGM implemented conditional restrictions that barred all access to their Okta (MGMResorts.okta.com) environment due to inadequate administrative capabilities and weak incident response playbooks. Their network has been infiltrated since Friday. Due to their network engineers’ lack of understanding of how the network functions, network access was problematic on Saturday. They then made the decision to “take offline” seemingly important components of their infrastructure on Sunday.

After waiting a day, we successfully launched ransomware attacks against more than 100 ESXi hypervisors in their environment on September 11th after trying to get in touch but failing. This was after they brought in external firms for assistance in containing the incident.

In our MGM victim chat, a user suddenly surfaced a few hours after the ransomware was deployed. As they were not responding to our emails with the special link provided (In order to

prevent other IT Personnel from reading the chats) we could not actively identify if the user in the victim chat was authorized by MGM Leadership to be present.

[...]

At this point, we have no choice but to criticize VX Underground for falsely reporting events that never happened. We typically consider their information to be highly reliable and timely, but we did not attempt to tamper with MGM's slot machines to spit out money because doing so would not be to our benefit and would decrease the chances of any sort of deal.

The rumors about teenagers from the US and UK breaking into this organization are still just that—rumors. We are waiting for these ostensibly respected cybersecurity firms who continue to make this claim to start providing solid evidence to support it. Starting to the actors' identities as they are so well-versed in them.

The truth is that these specialists find it difficult to delineate between the actions of various threat groupings, therefore they have grouped them together. Two wrongs do not make a right, thus they chose to make false attribution claims and then leak them to the press when they are still unable to confirm attribution with high degrees of certainty after doing this. The tactics, procedures, and indicators of compromise (TTPs) used by the people they blame for the attacks are known to the public and are relatively easy for anyone to imitate.

The ALPHV ransomware group has not before privately or publicly claimed responsibility for an attack before this point. Rumors were leaked from MGM Resorts International by unhappy employees or outside cybersecurity experts prior to this disclosure. Based on unverified disclosures, news outlets made the decision to falsely claim that we had claimed responsibility for the attack before we had.

We still continue to have access to some of MGM's infrastructure. If a deal is not reached, we shall carry out additional attacks. We continue to wait for MGM to grow a pair and reach out as they have clearly demonstrated that they know where to contact us.

Tech Crunch: neither you nor anybody else was contacted by the hacker who took control of MGM. Next time, verify your sources more thoroughly, or at the very least, give some hint that you do.

Lastly, knowledge-based trust is built on behavioural predictability. This type of trust develops as parties come to understand each other's behaviour patterns, values, and intentions over time. Several methods can improve understanding and predictability between parties.[13] Ongoing communication is one such method: for example, trust between teammates can be enhanced by routine meetings and status updates. Another way to bolster knowledge-based trust is through what is called 'courtship', the process of carefully and deliberately gathering information, assessing compatibility, and building a relationship before making a formal commitment.[14] This kind of courtship is about ensuring that the relationship – whether in business, purchasing, or any other partnership – is built on a solid foundation of trust, compatibility, and mutual understanding.

More established groups have a communication plan to project a sense of behavioural predictability to their victims. Many ransomware actors, such as Cl0p, which emerged in early 2019, establish dedicated channels for victims to negotiate ransom payments and receive instructions. They adopt a service-oriented tone, referring to victims as 'customers' and themselves as 'support,' creating the illusion of a legitimate business transaction. Groups like 8Base, which appeared in 2022, even include FAQs and guidelines on their darknet sites to seem more transparent in their operations.[15]

After receiving the ransom payment, some groups offer comprehensive technical support to assist victims in decrypting their data. This support often includes detailed, step-by-step instructions, troubleshooting services for issues with the decryption key, and, occasionally, help desks staffed with operators prepared to assist with the decryption process. To demonstrate their capability and build trust, some groups even decrypt a few files for free, reassuring victims that they have the means and will indeed unlock the files once payment is made (also see Box 1: DarkSide Ransom Note).

An example of this approach is illustrated in the initial message from the ransomware group HelloKitty to their victims: 'Unfortunately, your files were encrypted, and more than 200 GB of your critical date was leaked from your File, DEV and SQL servers (Administration and Finance, Direzione, Legal, HR, Risorse umane). For a more detailed list of documents, please contact us and we will send you the samples we have. We are also ready to help you recover your files, prevent the spread of leaks, as well as help solve problems in your IT infrastructure that were the cause of the current situation, so that this does not happen again in the future. Just contact support using the following methods and we will decrypt one non-important file for free to convince you of our honesty.'[16]

---

[13] Shapiro, Sheppard, Cheraskin, "Business on a Handshake"

[14] Ibid.

[15] On the FAQs see: Brian Krebs, 'Who's Behind the 8Base Ransomware Website?,' *KrebsonSecurity*, September 18, 2023, https://krebsonsecurity.com/2023/09/whos-behind-the-8base-ransomware-website/; on their background as pentesters see: Intel Cocktail, '8BASE Ransomware Group Interview:
'We Are Honest and Simple Pentesters', 2024, https://intelcocktail.com/8base-interview/

[16] See HelloKitty's ransom note: ransomware_notes/hellokitty/[File_Name].README_TO_RESTORE at main · ThreatLabz/ransomware_notes · GitHub

Brand Strategy

Ransomware groups often come up with meaningful names that help their branding. REvil, short for 'Ransomware Evil,' draws inspiration from the Resident Evil film franchise. Meanwhile, the Cl0p ransomware, which made its debut in February 2019, takes its name from the Russian word for bug, 'klop'.

Beyond names, these criminal groups create distinctive brand identities through symbols and logos. The Vice Society, a group targeting educational institutions since 2021, uses imagery reminiscent of the video game *Grand Theft Auto: Vice City*. Lockbit, another well-known group, features a retro red, white, and black logo prominently across their communication channels. In a striking example of how far these groups go to build notoriety, Lockbit offered individuals up to $1,000 in Bitcoin to tattoo the Lockbit logo on their bodies.

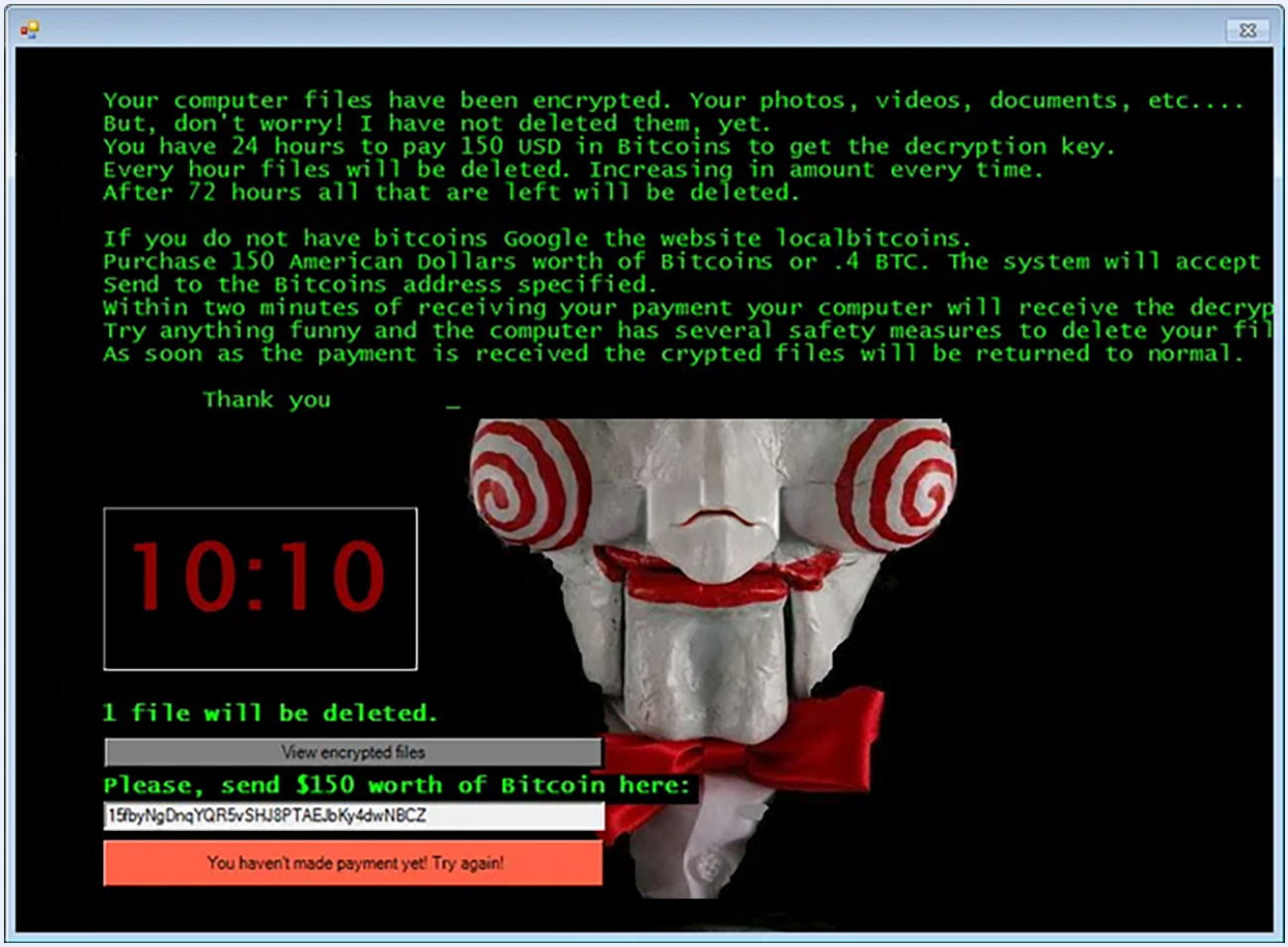
This approach to branding marks a departure from the behaviour of earlier ransomware groups like Archiveus, which would disguise its malicious nature under the guise of authoritative entities, misleading victims with warnings of illegal online behaviour and demanding fines for release. The practice of naming these groups and variants often fell to external observers, as seen with Jigsaw ransomware, initially dubbed 'BitcoinBlackmailer' but later renamed due to its association with imagery from the Saw movie series.

> Early forms of ransomware targeted victims indiscriminately, relying less on a public image or reputation to attract collaborators, given their smaller scale and less sophisticated organisation.

The emphasis on branding was less pronounced among these predecessors, who relied more on what is called a 'spray and pray' approach – sending out large volumes of indiscriminate attacks aimed at smaller victims.

[17] Lucian Constantin, 'REvil Ransomware Explained: A Widespread Extortion Operation,' *CSO Online*, November 12, 2021, https://www.csoonline.com/article/570101/revil-ransomware-explained-a-widespread-extortion-operation.html.

Spotlight Case Study: **Ransom Note Jigsaw**



An important aspect of the emergence of more developed ransomware group branding is the potential to establish multiple brands within the same organisational structure. An analogy can be drawn from The Coca-Cola Company, which offers over 200 brands ranging from well-known soft drinks like Coca-Cola, Sprite, and Fanta, to products like Dasani water, Vitaminwater, and Costa coffee and tea.

There are various benefits to creating multiple brands. First, given that law enforcement and regulatory bodies are constantly on the lookout to disrupt the actions of ransomware groups, the ability to operate under various brands can safeguard against disruptions. Authorities often alert the public to the activities of specific cybercriminal groups, urging non-compliance and actively working to dismantle their networks or expose their encryption keys. By maintaining multiple brands, each with distinct technology and methodologies, a ransomware organisation can ensure the continuity of its operations, even if one brand comes under scrutiny.

Spotlight Case Study: **Diavol: Conti's Failed Brand Diversion**

In early June 2021, the Endpoint Detection and Response team of Fortinet, a cybersecurity company, stopped a ransomware attack on a client. Following a common pattern, the ransom note left in each affected folder contained a URL leading to a site with a red banner reading 'Diavol,' the Romanian word for 'devil.'[18]

The Fortinet team noticed that the deployment of the Diavol locker coincided with the appearance of locker64.dll, an encryption module used by a later version of the Conti ransomware variant, on the victim's system. At the time, Conti was the most dominant ransomware brand. Researchers initially noted similarities but lacked proof to link them.[19] A month later, IBM researchers found stronger links between Diavol and Trickbot, Conti's close associate.[20] The FBI later confirmed these links in an official advisory.[21]

Conti chat messages leaked in February 2022 revealed more about what was happening behind the scenes.

On August 19, 2021, a senior Conti member called 'Professor' was outraged upon discovering a critical oversight: Diavol had mistakenly included a TrickBot module blocking attacks on Commonwealth of Independent States. This mistake provided security experts with the evidence needed to tie TrickBot and Conti to the same cybercriminal entity. In his communication with his boss, Professor vented: 'Stern. Did you see how they fucked it up? Regarding the affiliation? I fucking almost exploded. They put a part of trickbot's code which is responsible for the reply on CIS into the build of Diavol. Despite me explicitly saying not to touch anything related to geo-targeting. And immediately the entire project is on the news as fully affiliated.'[22]

Diavol was meant to be distinct from Conti's flagship brand. The exact motives remain uncertain, but based on Professor's messages, Conti's leadership likely wanted to evade law enforcement, threat intelligence companies, and media scrutiny. They likely assumed Diavol would draw less attention, allowing them to operate with lower risk.

---

[18] Lawrence Abrams, 'FBI Links Diavol Ransomware to the TrickBot Cybercrime Group.' BleepingComputer, January 20, 2022, https://www.bleepingcomputer.com/news/security/fbi-links-diavol-ransomware-to-the-trickbot-cybercrime-group/; Dor Neemani, and Asaf Rubinfeld, 'Diavol - A New Ransomware Used By Wizard Spider?' Fortinet, July 1, 2021, https://www.fortinet.com/blog/threat-research/diavol-new-ransomware-used-by-wizard-spider.

[19] Dor Neemani, and Asaf Rubinfeld, 'Diavol - A New Ransomware Used By Wizard Spider?' Fortinet, July 1, 2021, https://www.fortinet.com/blog/threat-research/diavol-new-ransomware-used-by-wizard-spider.

[20] Ionut Ilascu, 'Diavol Ransomware Sample Shows Stronger Connection to TrickBot Gang.' BleepingComputer, August 18, 2021, https://www.bleepingcomputer.com/news/security/diavol-ransomware-sample-shows-stronger-connection-to-trickbot-gang/; Charlotte Hammond and Chris Caridi, 'Analysis of Diavol Ransomware Reveals Possible Link to TrickBot Gang.' *Security Intelligence* (blog), August 17, 2021, https://securityintelligence.com/posts/analysis-of-diavol-ransomware-link-trickbot-gang/. Also see Binary Defense, 'New Ransomware 'Diavol' Being Dropped by Trickbot.' Binary Defense, April 18, 2023, https://www.binarydefense.com/resources/threat-watch/new-ransomware-diavol-being-dropped-by-trickbot/.

[21] Abrams, 'FBI Links Diavol Ransomware to the TrickBot Cybercrime Group.'.

[22] Ibid.

Another significant motive for cultivating diverse brands is tactical re-engagement with previous victims while preserving the credibility of the organisation's other brands. When a victim complies with a ransom demands, ransomware actors typically pledge not only to decrypt the victim's data and refrain from leaking stolen information, but also to avoid any future targeting of the same entity. However, with multiple brands at their disposal, these groups can revisit previous targets under a new guise. Armed with valuable data including detailed knowledge of the target's systems, they can initiate a second round of extortion. Presenting the victim with a new 'shame site' adorned with a different name and logo, possibly using alternative encryption tools, subtly increases the odds of a successful extortion, all while maintaining the facade of a trustworthy entity in the broader public sphere.

Misguided Government Measures

Whilst branding and reputation building are central, not peripheral, to the success of ransomware groups in overcoming the 'Ransomware Trust Paradox', governments and other organisations rarely attempt to undermine this aspect of ransomware groups.

When the International Counter Ransomware Initiative (CRI) was launched in 2021 by the White House and international partners, it prioritised enhancing network resilience, disrupting ransomware operations through law enforcement collaboration, and countering the financial mechanisms that sustain ransomware profitability.[23] At the second CRI Summit, members reaffirmed their commitment to these goals and established the International Counter Ransomware Task Force (ICRTF) to coordinate and disrupt ransomware activities on an operational level.[24] By the third meeting in 2023, the CRI had expanded its scope, incorporating capacity-building efforts and private sector collaboration into all aspects of its strategy.[25] Despite these advances, the CRI has not yet addressed the 'Ransomware Trust Paradox' and the crucial role that branding and reputation play in the success of ransomware groups.

Similarly, the Ransomware Task Force (RTF), a commendable multi-stakeholder effort by the Institute for Security and Technology (IST), includes participants from government, industry, and civil society.[26] In May 2021, the RTF released a thorough report on countering ransomware with 48 key recommendations aimed primarily at governments and the private sector. This report has significantly improved efforts to disrupt ransomware operations, enhance information sharing, and strengthen strategies for mitigating and recovering from attacks.[27] Yet, like the CRI, the RTF's report offers no policy recommendations to undermine the reputation or credibility of ransomware groups.[28]

---

[23] The White House, 'Joint Statement of the Ministers and Representatives from the Counter Ransomware Initiative Meeting October 2021,' October 2021, https://uk.usembassy.gov/joint-statement-of-the-ministers-and-representatives-from-the-counter-ransomware-initiative-meeting-october-2021/

[24] The White House, 'FACT SHEET: The Second International Counter Ransomware Initiative Summit,' November 1, 2022, https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2022/11/01/fact-sheet-the-second-international-counter-ransomware-initiative-summit/

[25] The White House, 'International Counter Ransomware Initiative 2023 Joint Statement,' November 2023, https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/10/02/international-counter-ransomware-initiative-2024-joint-statement/

[26] Institute for Security and Technology, 'Ransomware Task Force (RTF),' https://securityandtechnology.org/ransomwaretaskforce/

[27] Ransomware Task Force, 'Combating Ransomware A Comprehensive Framework for Action: Key Recommendations from the Ransomware Task Force.', https://securityandtechnology.org/wp-content/uploads/2021/04/IST-Ransomware-Task-Force-Report.pdf; Ransomware Task Force, 'Ransomware Task Force: Doubling Down,' *IST: Institute for Security and Policy*, 2024, https://securityandtechnology.org/virtual-library/reports/ransomware-task-force-doubling-down/

[28] The cornerstone report from the RTF outlines four key goals; deterring ransomware attacks, disrupting the ransomware business model, helping organizations prepare for ransomware attacks, and enhancing responses to such attacks. These goals lead to detailed objectives, which in turn guide specific recommended actions.

> We have seen only sporadic attempts to undercut ransomware groups' credibility and reliability.

For instance, government disclosures have highlighted cases where ransomware groups failed to decrypt data after receiving payments or targeted victims again, despite promises to the contrary. Similarly, after the UK National Crime Agency (NCA) and its international partners disrupted the Lockbit ransomware group, officials stressed on numerous occasions that the ransomware group does not fully delete data even after receiving a payment, contradicting their previous assurances.[29] This revelation has made it harder for Lockbit as a brand to make a comeback; even if it manages to revive some infrastructure, it carries a stigma that will undermine its operations. Negotiators have also begun using this knowledge in subsequent cases, leveraging it to secure lower ransom demands from other ransomware groups by challenging the credibility of attackers' promises based on this precedent.[30]

These efforts should go much further. Over the past few years, the life cycle of ransomware brands has been shorter. New ransomware brands seem to come and go more quickly. This volatility can be attributed to intensified government efforts to combat ransomware. Initiatives like taking down command and control infrastructure of certain ransomware groups have influenced this dynamic.

However, it has also become notably easier for a ransomware newcomer to quickly build brand equity, thanks to the burgeoning industry of ransomware reporting. Before cybersecurity coverage was the multi-billion dollar industry it is today, ransomware groups had to make substantial efforts to garner public attention, relying on a small cadre of security experts for exposure. Now, the scenario has transformed: any release by a ransomware group on a leak site gains immediate traction across social media platforms, news websites, and expert company analyses.

The race to cover new entrants in the cybercriminal ecosystem introduces a unique and complex challenge for the public dissemination of information on ransomware. As previously noted, this is unlike the situation with state hacking groups. The difference arises from the contrasting preferences of these groups; state hacking groups – especially intelligence agencies – aim to operate under the radar, whereas ransomware groups often seek out and benefit from the visibility provided by cybersecurity reporting.

---

[29] See the press conference; also see the BCC story and Alexander Martin, "LockBit held victims' data even after receiving ransom payments to delete it," The Record, February 21, 2024, https://therecord.media/lockbit-lied-about-deleting-exfiltrated-data-after-ransom-payments; for more information on the NCA led operation see: NCA: National Crime Agency, "International investigation disrupts the world's most harmful cyber crime group," February 20, 2024, https://www.nationalcrimeagency.gov.uk/news/nca-leads-international-investigation-targeting-worlds-most-harmful-ransomware-group; Matt Burgess, "A Global Police Operation Just Took Down the Notorious LockBit Ransomware Gang," Wired, February 20, 2024, https://www.wired.com/story/lockbit-ransomware-takedown-website-nca-fbi/

Consequently, the cybersecurity community's eagerness to report plays a critical role in shaping the ransomware ecosystem. Public reporting speeds up the rebranding efforts of these groups, aids in the development of their reputations, and significantly affects their positioning within the constantly shifting cybersecurity environment and relationships with victims.

> The cybersecurity community's eagerness to report plays a critical role in shaping the ransomware ecosystem.

Journalists and cybersecurity researchers therefore need to be cautious when reporting on ransomware attacks to avoid inadvertently promoting the groups behind these attacks. Conti, the group behind the attack against the government of Costa Rica, established its efficient and ruthless reputation largely through media narratives and research analyses. Between 2020 and 2021, over a thousand articles on various aspects of Conti's operations were published. This extensive coverage enhanced Conti's brand and even featured in their ransom note: 'If you don't who we are – just 'google it." Similarly, in a blog post, Vice Society thanked a specific journalist for an article in which it was part of a 'Top 5' of ransomware and malware groups in 2022.[31]

[30] Remarks at Oxford Cyber Forum. Virtual Routes. "ECCRI Holds the Oxford Cyber Forum." Virtual Routes (blog), June 27, 2024. https://virtual-routes.org/eccri-holds-the-oxford-cyber-forum/.

[31] Sophos X-ops, 'Press and Pressure'

A New Code of Ethics

> Governments should advocate for a code of ethics to help journalists and researchers responsibly report on ransomware, to prevent the glorification or legitimisation of these groups.

The development of such a code can benefit from lessons learned in other fields where reporting carries similar risks of unintended consequences.

One pertinent example is the realm of disinformation, where media coverage can inadvertently reinforce the very dynamics it seeks to dismantle.[32] To address this challenge, UNESCO developed a handbook for journalism training on fake news, prompted by growing international concerns over a 'disinformation war' targeting journalists and the broader media landscape. This handbook highlights seven core principles – fairness, independence, accuracy, contextuality, transparency, protection of confidential sources, and perspicacity – that collectively build trust, credibility, and public confidence.[33]

Additionally, a report by Whitney Phillips for Data & Society encourages journalists to assess whether a story has reached a tipping point, offers a public health takeaway, or provides a political or social action point, and whether the risk of entrenching or rewarding a falsehood outweighs the benefits of debunking it.[34] If these criteria are not met, the story might best be left unreported at that time.

Perhaps even more instructive is the field of counterterrorism. Terrorists seek to instill widespread fear, and while accurate reporting on terrorist incidents is essential, it can paradoxically help achieve terrorists' aims. Consequently, a range of guidelines and trainings have been developed to help journalists navigate this sensitive reporting terrain.[35]

---

[32] Olga Belogolova, Lee Foster, Thomas Rid, and Gavin Wilde, 'Don't Hype the Disinformation Threat,' Foreign Affairs, May 3, 2024, https://www.foreignaffairs.com/russian-federation/dont-hype-disinformation-threat

[33] Cherilyn Ireton and Julie Posetti (eds.), Journalism, fake news & disinformation: handbook for journalism education and training, UNESCO, 2018, https://unesdoc.unesco.org/ark:/48223/pf0000265552

[34] Whitney Phillips, 'The Oxygen of Amplification: Better Practices for Reporting,' Data & Society, 2018, https://datasociety.net/library/oxygen-of-amplification/

[35] Eg. Jean Paul Marthoz and Khalid Aoutail, 'Media and the coverage of terrorism: manual for trainers and journalism educators,' UNESCO, 2022, https://unesdoc.unesco.org/ark:/48223/pf0000380356

Four critical elements should form the foundation of any code of ethics for ransomware reporting. First, accuracy must be prioritised over speed. Some ransomware observers rely on automated bots that scrape and publish information from leak sites without verification. These bots can quickly spread misinformation released by ransomware groups aware of the attention these sites garner. In contrast, Andy Greenberg at *Wired* has exemplified a desire for accuracy, remembering that, 'we at *Wired* held off on reporting on the second ransomware gang, RansomHub, threatening to leak a trove of stolen data from Change Healthcare until the hackers provided evidence of their claims.' [36]

Second, journalists must use language cautiously when describing ransomware groups. It is important to avoid terms like 'cyber gangsters' that might inadvertently glorify these groups or enhance their fearsome image. Such portrayals can play into the hands of ransomware operators by bolstering their intended reputation as highly skilled and ruthless.

Third, reporting should avoid overemphasising successful attacks by ransomware groups. It is beneficial to highlight instances where security measures or backup strategies have successfully mitigated the impact of an attack or where organisations have managed to recover without complying with ransom demands. This helps to provide a more balanced view of ransomware's actual consequences and the effectiveness of protective measures.

Fourth, like any cyber incident, careful consideration should be given to the amount of technical detail shared about security vulnerabilities. Information about unpatched vulnerabilities should only be published if it is essential for public safety and awareness and does not provide cybercriminals with exploitable information.

[36] Andy Greenberg, Linkedin, April 2024, https://www.linkedin.com/posts/andygreenbergjournalist_change-healthcare-faces-another-ransomware-activity-7184620359134912512-kVEK/. The article: Andy Greenberg and Matt Burgess, 'Change Healthcare Faces Another Ransomware Threat—and It Looks Credible,' *Wired*, April 12, 2024, https://www.wired.com/story/change-healthcare-ransomhub-threat/

A Unique Opportunity for Intervention

The rising threat of ransomware requires a shift in both public understanding and governmental response. The 'Ransomware Trust Paradox' offers a unique opportunity for intervention. Governments and organisations must move beyond traditional tactics and consider the nuanced role of trust in the ransomware ecosystem. By adopting an approach that not only targets the operational capabilities of these groups, but also diminishes their ability to project trustworthiness and credibility, we can undermine the foundations of ransomware groups' profitability. As we refine our counter-ransomware efforts, it is essential not to inadvertently strengthen these groups through media coverage. The first step in this direction is to carefully craft and implement a code of ethics on ransomware reporting.

This report is adapted from Max Smeets' book* Ransom War: How Cyber Crime Became a Threat to National Security, *published by Oxford University Press and Hurst Publishers (2025). He also gave a keynote at LABSCON 2024 on the Ransomware Trust Paradox.

References

Abrams, Lawrence. 'FBI Links Diavol Ransomware to the TrickBot Cybercrime Group.' BleepingComputer, January 20, 2022. https://www.bleepingcomputer.com/news/security/fbi-links-diavol-ransomware-to-the-trickbot-cybercrime-group/.

———. 'Meet LostTrust Ransomware — A Likely Rebrand of the MetaEncryptor Gang.' BleepingComputer, October 1, 2023. https://www.bleepingcomputer.com/news/security/meet-losttrust-ransomware-a-likely-rebrand-of-the-metaencryptor-gang/.

Belogolova, Olga, Lee Foster, Thomas Rid, and Gavin Wilde. 'Don't Hype the Disinformation Threat | Foreign Affairs,' May 3, 2024. https://www.foreignaffairs.com/russian-federation/dont-hype-disinformation-threat.

Binary Defense. 'New Ransomware 'Diavol' Being Dropped by Trickbot.' Binary Defense, April 18, 2023. https://www.binarydefense.com/resources/threat-watch/new-ransomware-diavol-being-dropped-by-trickbot/.

Burgess, Matt. 'A Global Police Operation Just Took Down the Notorious LockBit Ransomware Gang.' *Wired*, February 28, 2024. https://www.wired.com/story/lockbit-ransomware-takedown-website-nca-fbi/.

Constantin, Lucian. 'REvil Ransomware Explained: A Widespread Extortion Operation.' CSO Online, November 12, 2021. https://www.csoonline.com/article/570101/revil-ransomware-explained-a-widespread-extortion-operation.html.

Dissent. 'At Some Point, SNAtch Team Stopped Being the Snatch Ransomware Gang. Were Journalists the Last to Know? – DataBreaches.Net.' DataBreaches.Net, September 1, 2023. https://databreaches.net/2023/09/01/at-some-point-snatch-team-stopped-being-the-snatch-ransomware-gang-were-journalists-the-last-to-know/.

Greenberg, Andy. 'Andy Greenberg on LinkedIn.' LinkedIn, April 2024. https://www.linkedin.com/posts/andygreenbergjournalist_change-healthcare-faces-another-ransomware-activity-7184620359134912512-kVEK.

Greenberg, Andy, and Matt Burgess. 'Change Healthcare Faces Another Ransomware Threat—and It Looks Credible.' *Wired*, April 12, 2024. https://www.wired.com/story/change-healthcare-ransomhub-threat/.

Hammond, Charlotte, and Chris Caridi. 'Analysis of Diavol Ransomware Reveals Possible Link to TrickBot Gang.' *Security Intelligence* (blog), August 17, 2021. https://securityintelligence.com/posts/analysis-of-diavol-ransomware-link-trickbot-gang/.

Ilascu, Ionut. 'Diavol Ransomware Sample Shows Stronger Connection to TrickBot Gang.' BleepingComputer, August 18, 2021. https://www.bleepingcomputer.com/news/security/diavol-ransomware-sample-shows-stronger-connection-to-trickbot-gang/.

Institute for Security + Technology (IST). 'Combating Ransomware: A Comprehensive Framework for Action: Key Recommendations from the Ransomware Task Force.' Institute for Security + Technology (IST), n.d. https://securityandtechnology.org/wp-content/uploads/2021/04/IST-Ransomware-Task-Force-Report.pdf.

———. 'Ransomware Task Force (RTF): Combating the Ransomware Threat with a Cross-Sector Approach.' Ransomware Task Force, n.d.

Intel Cocktail. '8BASE Ransomware Group Interview: 'We Are Honest and Simple Pentesters."' *Intel Cocktail* (blog), 2024. https://intelcocktail.com/8base-interview/.

Ireton, Cherilyn, Julie Posetti, and UNESCO, eds. *Journalism, 'Fake News' & Disinformation: Handbook for Journalism Education and Training*. UNESCO Series on Journalism Education. Paris: United Nations Educational, Scientific and Cultural Organisation, 2018. https://unesdoc.unesco.org/ark:/48223/pf0000265552.

Krebs, Brian. 'Who's Behind the 8Base Ransomware Website?' KrebsOnSecurity (blog), September 18, 2023. https://krebsonsecurity.com/2023/09/whos-behind-the-8base-ransomware-website/.

Marthos, Jean Paul, Khalid Aoutail, and UNESCO. *Media and the Coverage of Terrorism: Manual for Trainers and Journalism Educators.* UNESCO Series on Journalism Education. United Nations Educational, Scientific and Cultural Organisation, 2022. https://unesdoc.unesco.org/ark:/48223/pf0000380356.

Martin, Alexander. 'LockBit Held Victims' Data Even after Receiving Ransom Payments to Delete It.' The Record, February 21, 2024. https://therecord.media/lockbit-lied-about-deleting-exfiltrated-data-after-ransom-payments.

National Crime Agency (NCA). 'International Investigation Disrupts the World's Most Harmful Cyber Crime Group,' February 20, 2024. https://www.nationalcrimeagency.gov.uk/news/nca-leads-international-investigation-targeting-worlds-most-harmful-ransomware-group.

Neemani, Dor, and Asaf Rubinfeld. 'Diavol - A New Ransomware Used By Wizard Spider?' Fortinet, July 1, 2021. https://www.fortinet.com/blog/threat-research/diavol-new-ransomware-used-by-wizard-spider.

Phillips, Whitney. 'The Oxygen of Amplification: Better Practices for Reporting on Extremists, Antagonists, and Manipulators.' Data & Society, May 22, 2018. https://datasociety.net/library/oxygen-of-amplification.

Ransomware Task Force. 'Ransomware Task Force: Doubling Down.' Institute for Security + Technology (IST), 2024. https://securityandtechnology.org/virtual-library/reports/ransomware-task-force-doubling-down/.

Schelling, Thomas C. *The Strategy of Conflict.* 1st Ed. Cambridge, Mass: Harvard Univ. Pr, 1960.

Shapiro, Debra L., Blair H. Sheppard, and Lisa Cheraskin. 'Business on a Handshake.' *Negotiation Journal* 8, no. 4 (1992): 365–77. https://doi.org/10.1111/j.1571-9979.1992.tb00679.x.

Smeets, Max. 'Why Hostage Negotiation Tactics Don't Work on Ransomware.' *Https://Bindinghook.Com/* (blog), November 28, 2024. https://bindinghook.com/articles-binding-edge/why-hostage-negotiation-tactics-dont-work-on-ransomware/.

Sophos X-Ops. 'Press and Pressure: Ransomware Gangs and the Media.' *Sophos News* (blog), December 13, 2023. https://news.sophos.com/en-us/2023/12/13/press-and-pressure-ransomware-gangs-and-the-media/.

———. 'International Counter Ransomware Initiative 2024 Joint Statement.' The White House, October 2, 2024. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/10/02/international-counter-ransomware-initiative-2024-joint-statement/.

———. 'Joint Statement of the Ministers and Representatives from the Counter Ransomware Initiative Meeting October 2021.' U.S. Embassy & Consulates in the United Kingdom, October 14, 2021. https://uk.usembassy.gov/joint-statement-of-the-ministers-and-representatives-from-the-counter-ransomware-initiative-meeting-october-2021/.

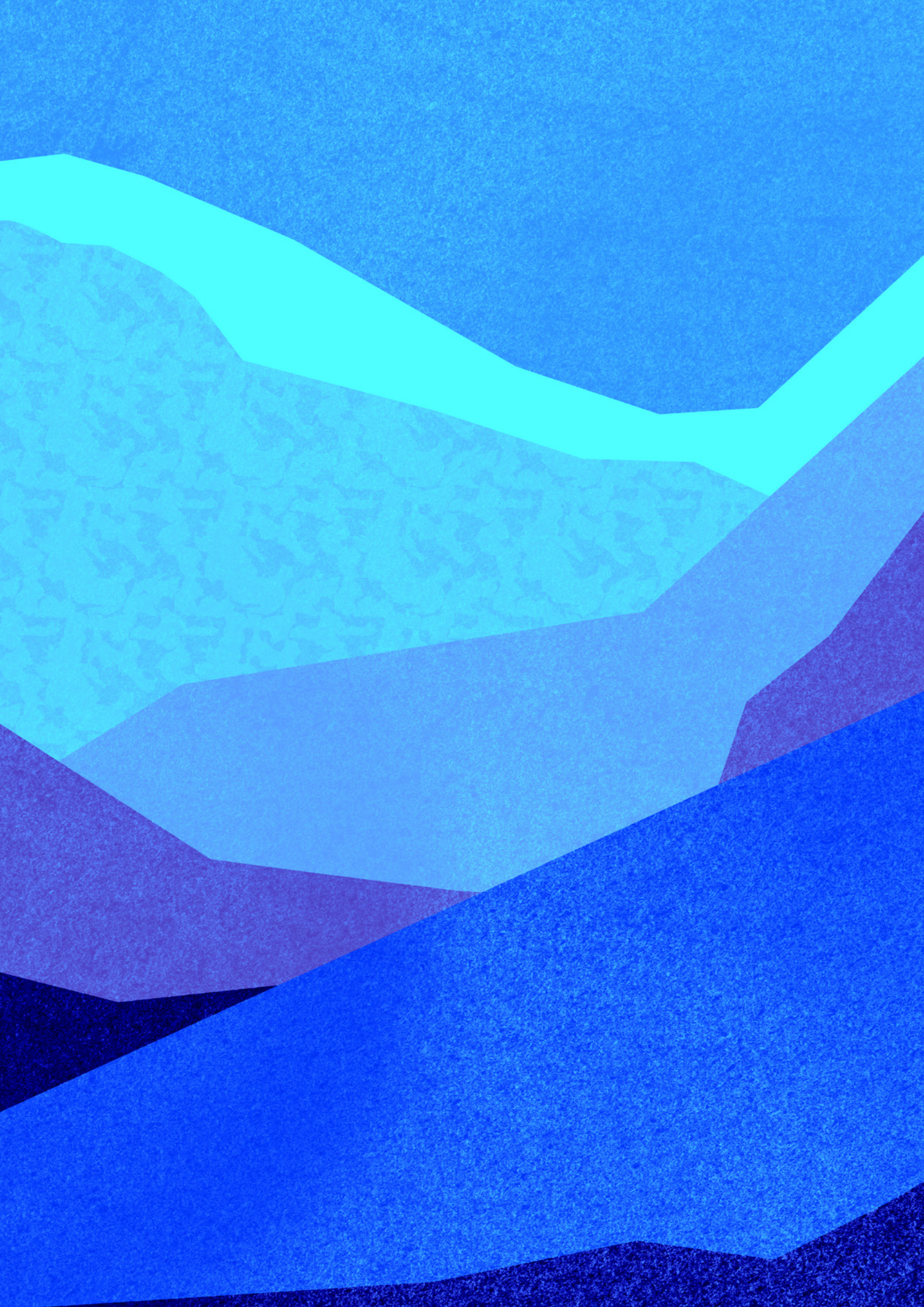
ThreatLabz. 'Darkside's Ransom Note.' ransomware_notes/darkside/darkside.txt, August 1, 2022. https://github.com/ThreatLabz/ransomware_notes/blob/main/darkside/darkside.txt.

———. 'Hello Kitty's Ransom Note.' ransomware_notes/hellokitty/[File_Name].README_TO_RESTORE, April 29, 2024. https://github.com/ThreatLabz/ransomware_notes/blob/main/hellokitty/%5BFile_Name%5D.README_TO_RESTORE.

———. 'Karma's Ransom Note.' ransomware_notes/karma/KARMA-ENCRYPTED.txt, January 4, 2023. https://github.com/ThreatLabz/ransomware_notes/blob/main/karma/KARMA-ENCRYPTED.txt.

Virtual Routes. 'ECCRI Holds the Oxford Cyber Forum.' *Virtual Routes* (blog), June 27, 2024. https://virtual-routes.org/eccri-holds-the-oxford-cyber-forum/.

White House. 'FACT SHEET: The Second International Counter Ransomware Initiative Summit.' The White House, November 1, 2022. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2022/11/01/fact-sheet-the-second-international-counter-ransomware-initiative-summit/.



virtual routes

For more information, please visit: **www.virtual-routes.org**

If you have any further queries, questions, or concerns, feel free to reach out via email at:
contact@virtual-routes.org