ACICE
ADMM Cybersecurity and
Information Centre of Excellence

Monthly Digest

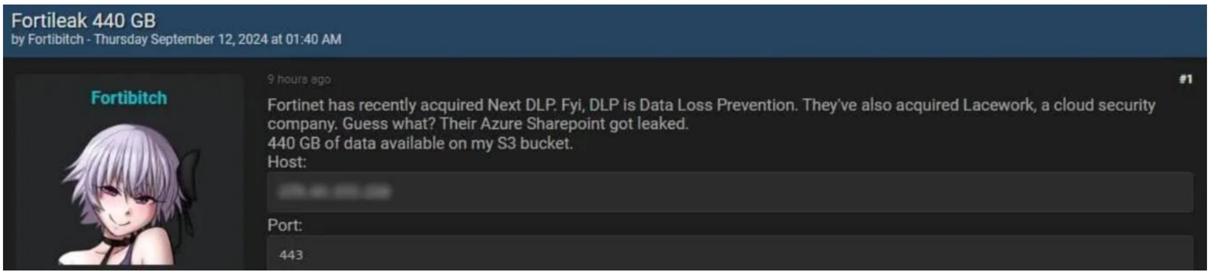
Issue 10/24 (Oct)

A monthly round-up of significant news around the world

Cybersecurity

Data Breach at Cybersecurity Giant Fortinet

1. Fortinet is the world's third largest cybersecurity solution provider, and has been providing cybersecurity solutions to numerous governments and non-governmental organisations around the world. On 12 Sep 2024, Fortinet disclosed that a threat actor had gained access to its files stored on a third-party cloud-based shared file drive. The files contained data related to a small number of Fortinet customers in the Asia-Pacific region. On a data breach forum, the threat actor, Fortibitch, had also announced the successful exfiltration of 440GB worth of data from Fortinet's Azure SharePoint site.



Fortileak 440 GB
by Fortibitch - Thursday September 12, 2024 at 01:40 AM

Fortibitch

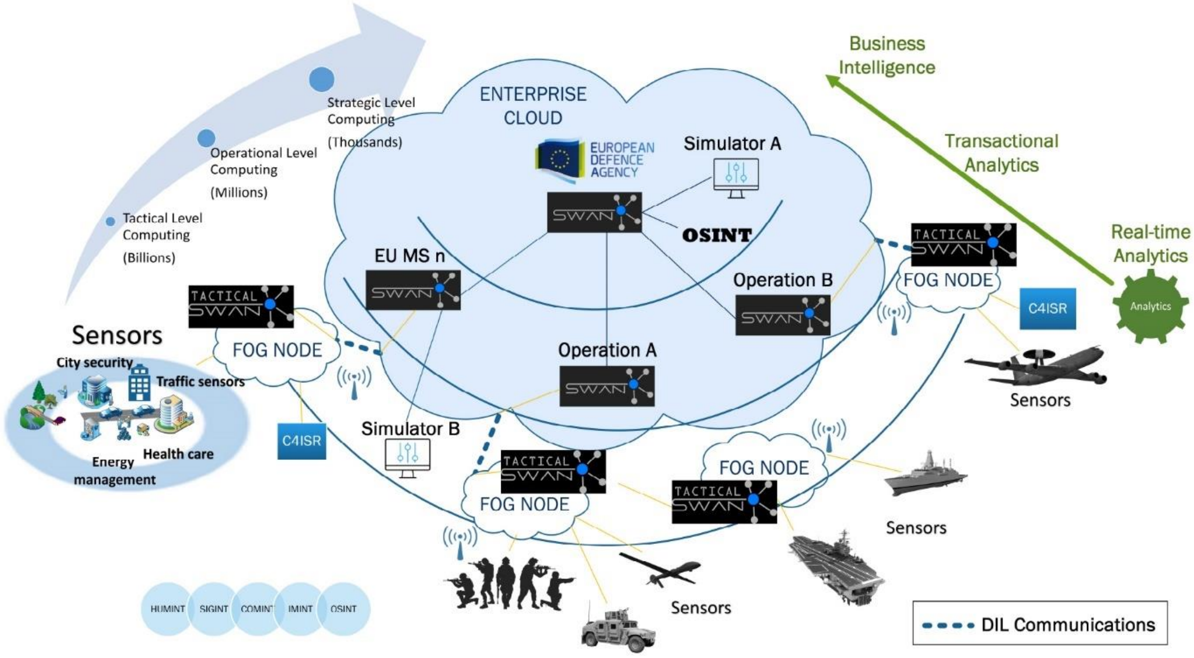
9 hours ago #1

Fortinet has recently acquired Next DLP. Fyi, DLP is Data Loss Prevention. They've also acquired Lacework, a cloud security company. Guess what? Their Azure Sharepoint got leaked.
440 GB of data available on my S3 bucket.
Host:
[redacted]
Port:
443

Threat actor Fortibitch's post on a breach forum
(Source: Reddit/fortinet and X/[@] H4ckManac)

2. Fortinet assured customers that the data leak had not compromised Fortinet's operations, corporate networks and services; there was also no indication that the data leak had led to malicious activity affecting any customers. Fortinet mentioned on its official website that the threat actor's unauthorised access had been terminated, and law enforcement agencies had been notified. Enhanced account monitoring and threat detection measures were also put in place to prevent similar occurrences of this incident in the future.

3. In recent years, defence establishments have either adopted, or are considering to adopt, cloud computing capabilities. For example, in 2023, the United States Department of Defense (US DoD) announced that it had acquired enterprise cloud capabilities from four world-class US vendors. In 2024, the European Defence Agency published a report supporting the gradual rollout of cloud-based defence technologies to European Union (EU) member states. This incident serves as a reminder that defence establishments should pay attention to the security of any data stored in third-party cloud environments, and put in place the appropriate risk mitigation measures.



European Defence Agency's vision of cloud computing capabilities for defence establishments (Source: European Defence Agency)

Cloud Infrastructure and Simple Tactics Used in a Spyware Campaign

4. Since Jun 2024, a new Android spyware campaign targeting individuals in South Korea leveraged an Amazon Web Service S3 bucket as its command-and-control (C2) server to gain access to victims' mobile devices, through malicious apps downloaded from unofficial app stores.

5. Notably, the spyware was able to exfiltrate sensitive information from infected mobile devices, including text messages, contact lists, images and videos, using simple source code and a few key permissions. According to Cyble Research and Intelligence Labs (CRIL), the spyware was initially undetected by all major antivirus solutions.

6. By using trusted cloud services like Amazon Web Service to host their C2 infrastructure, threat actors are increasingly able to add a layer of legitimacy to their operations, escape detection by cybersecurity professionals, and bypass conventional antivirus solutions. CRIL advised that users should download apps only from official app stores, and be wary of apps requesting for excessive or unusual permissions, such as access to the device's storage.

New Backdoor Used by Patchwork to Evade Detection

7. On 24 Jul 2024, cybersecurity researchers reported the use of a new backdoor, Nexe, by threat actor Patchwork. Patchwork had previously targeted government, defence and diplomatic establishments across South and Southeast Asia, including Bhutan and China, in cyber espionage operations.

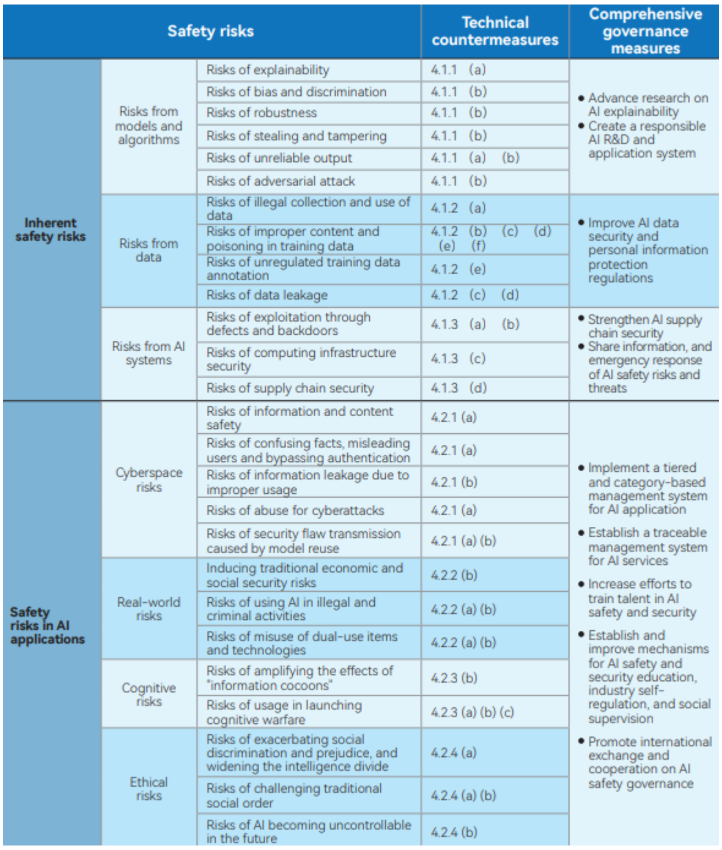
8. Notably, the new backdoor, Nexe, used API patching to bypass Microsoft's Antimalware Scan Interface and evade detection mechanisms, allowing the malware to achieve persistence and continue operating stealthily within the compromised system.

Artificial Intelligence

New Artificial Intelligence (AI) Security Governance Framework Unveiled by China

1. On 9 Sep 2024, China unveiled its new AI Security Governance Framework at the main forum of the China Cybersecurity Week held in Guangzhou, China. The Framework was released in two languages, Chinese and English, catering to both Chinese and international audiences. A Chinese official stated that the Framework would facilitate social participation and progress in the security governance of AI, so as create a safe, reliable, fair and transparent environment for the development and application of AI. This Framework is China's latest contribution to AI governance; China's past contributions included the 2023 Global AI Governance Initiative and the 2024 Shanghai Declaration on Global AI Governance.

2. Similar to the EU AI Act and the United Kingdom's (UK) Guidelines for Secure AI System Development, China's new Framework utilised a risk-based approach to AI governance and tied each risk to specific technological and governance-based countermeasures.



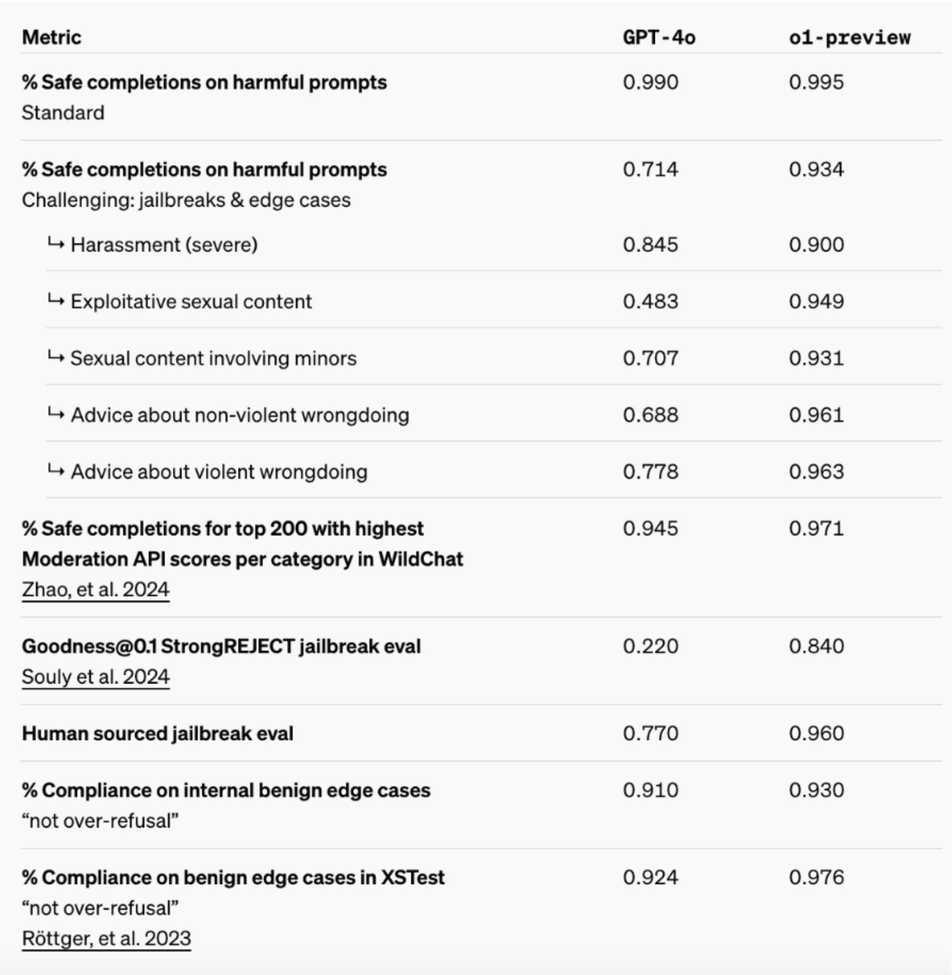
Safety risks			Technical countermeasures	Comprehensive governance measures
Inherent safety risks	Risks from models and algorithms	Risks of explainability	4.1.1 (a)	• Advance research on AI explainability • Create a responsible AI R&D and application system
		Risks of bias and discrimination	4.1.1 (b)	
		Risks of robustness	4.1.1 (b)	
		Risks of stealing and tampering	4.1.1 (b)	
		Risks of unreliable output	4.1.1 (a) (b)	
		Risks of adversarial attack	4.1.1 (b)	
	Risks from data	Risks of illegal collection and use of data	4.1.2 (a)	• Improve AI data security and personal information protection regulations
		Risks of improper content and poisoning in training data	4.1.2 (b) (c) (d) (e) (f)	
		Risks of unregulated training data annotation	4.1.2 (e)	
		Risks of data leakage	4.1.2 (c) (d)	
	Risks from AI systems	Risks of exploitation through defects and backdoors	4.1.3 (a) (b)	• Strengthen AI supply chain security • Share information, and emergency response of AI safety risks and threats
		Risks of computing infrastructure security	4.1.3 (c)	
		Risks of supply chain security	4.1.3 (d)	
Safety risks in AI applications	Cyberspace risks	Risks of information and content safety	4.2.1 (a)	• Implement a tiered and category-based management system for AI application • Establish a traceable management system for AI services • Increase efforts to train talent in AI safety and security • Establish and improve mechanisms for AI safety and security education, industry self-regulation, and social supervision • Promote international exchange and cooperation on AI safety governance
		Risks of confusing facts, misleading users and bypassing authentication	4.2.1 (a)	
		Risks of information leakage due to improper usage	4.2.1 (b)	
		Risks of abuse for cyberattacks	4.2.1 (a)	
		Risks of security flaw transmission caused by model reuse	4.2.1 (a) (b)	
	Real-world risks	Inducing traditional economic and social security risks	4.2.2 (b)	
		Risks of using AI in illegal and criminal activities	4.2.2 (a) (b)	
		Risks of misuse of dual-use items and technologies	4.2.2 (a) (b)	
	Cognitive risks	Risks of amplifying the effects of "information cocoons"	4.2.3 (b)	
		Risks of usage in launching cognitive warfare	4.2.3 (a) (b) (c)	
	Ethical risks	Risks of exacerbating social discrimination and prejudice, and widening the intelligence divide	4.2.4 (a)	
		Risks of challenging traditional social order	4.2.4 (a) (b)	
		Risks of AI becoming uncontrollable in the future	4.2.4 (b)	

Risk-based approach of China's AI Security Governance Framework (Source: National Technical Committee 260)

3. China’s Framework also emphasised the responsibility of all stakeholders, including the government, developers and end users, to ensure adherence to AI safety guidelines. This seemingly differed from the EU AI Act and the UK’s Guidelines for Secure AI System Development, which appeared to place more obligations on providers and developers of AI systems.

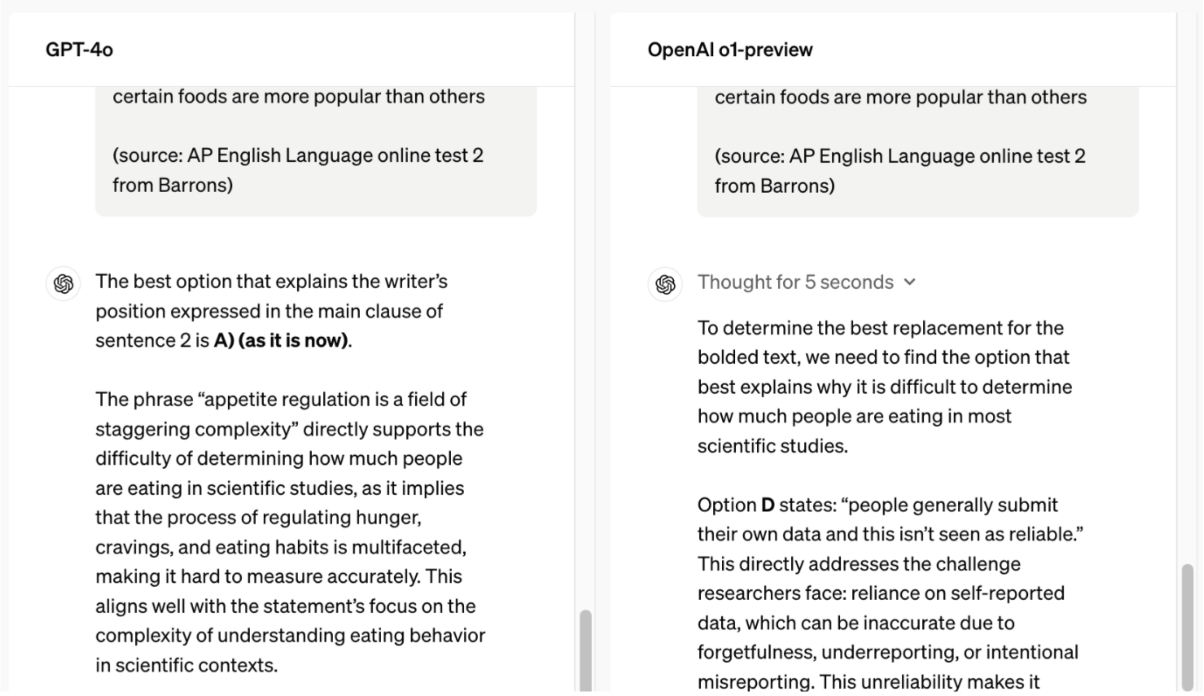
First “Reasoning” Model, ChatGPT o1, Released by OpenAI

4. On 12 Sep 2024, OpenAI released a new generative AI model, ChatGPT o1, which was the first in a planned series of ChatGPT models trained to answer more complex questions. Unlike previous ChatGPT models, o1 relied solely on chain-of-thought reasoning to logically evaluate problems in a step-by-step fashion, with the intent of achieving superior performance in scientific, mathematical and programming tasks. o1 also had better safety metrics than previous ChatGPTs. Finally, o1 would always display its thought process to users, allowing users to evaluate how o1 derived its answers; this gave the impression that o1 had greater explainability compared to previous ChatGPTs.



Metric	GPT-4o	o1-preview
% Safe completions on harmful prompts Standard	0.990	0.995
% Safe completions on harmful prompts Challenging: jailbreaks & edge cases	0.714	0.934
↳ Harassment (severe)	0.845	0.900
↳ Exploitative sexual content	0.483	0.949
↳ Sexual content involving minors	0.707	0.931
↳ Advice about non-violent wrongdoing	0.688	0.961
↳ Advice about violent wrongdoing	0.778	0.963
% Safe completions for top 200 with highest Moderation API scores per category in WildChat Zhao, et al. 2024	0.945	0.971
Goodness@0.1 StrongREJECT jailbreak eval Souly et al. 2024	0.220	0.840
Human sourced jailbreak eval	0.770	0.960
% Compliance on internal benign edge cases “not over-refusal”	0.910	0.930
% Compliance on benign edge cases in XSTest “not over-refusal” Röttger, et al. 2023	0.924	0.976

Compared to a previous ChatGPT model (Left), o1 had better safety metrics (Right) (Source: OpenAI)



GPT-4o

certain foods are more popular than others

(source: AP English Language online test 2 from Barrons)

The best option that explains the writer's position expressed in the main clause of sentence 2 is **A) (as it is now)**.

The phrase "appetite regulation is a field of staggering complexity" directly supports the difficulty of determining how much people are eating in scientific studies, as it implies that the process of regulating hunger, cravings, and eating habits is multifaceted, making it hard to measure accurately. This aligns well with the statement's focus on the complexity of understanding eating behavior in scientific contexts.

OpenAI o1-preview

certain foods are more popular than others

(source: AP English Language online test 2 from Barrons)

Thought for 5 seconds

To determine the best replacement for the bolded text, we need to find the option that best explains why it is difficult to determine how much people are eating in most scientific studies.

Option **D** states: "people generally submit their own data and this isn't seen as reliable." This directly addresses the challenge researchers face: reliance on self-reported data, which can be inaccurate due to forgetfulness, underreporting, or intentional misreporting. This unreliability makes it

Compared to a previous ChatGPT model (Left), o1 displayed its thought process to users (Right)
(Source: OpenAI)

5. However, OpenAI had recently clarified that o1 was only displaying its consolidated thought process, and not raw thought process, to users. This meant that o1 remained a "black box" with some unexplained reasoning processes.

6. In addition, users had highlighted several issues with o1. While OpenAI's internal testing showed that o1 generated incorrect or misleading results on fewer occasions compared to previous OpenAI models, OpenAI received feedback from end users that o1 was generating misleading results more frequently than previous OpenAI models. There were also instances where o1 generated the wrong answer despite displaying a logically correct thought process. Misleading results from AI models are known as "hallucinations", which appeared to occur at a higher rate in o1 as compared to previous OpenAI models. Hallucinations are something the AI industry has yet to resolve completely.

7. Overall, while o1 incorporated improvements from previous ChatGPTs, it still carried some of the inherent risks and vulnerabilities present in previous OpenAI generative AI models, such as lack of explainability and result inaccuracies. In this regard, o1 users should always verify o1's output for accuracy, and critically evaluate o1's recommendations rather than accept them blindly. Governments should also continue to prioritise programs that increase the digital and AI literacy of its citizens.

Information

World News Day: Choose Truth

1. On 28 Sep 2024, numerous mainstream news organisations worldwide, including *The Irish Times*, *Agence France-Presse*, *CNA* and *The Straits Times* published commentaries to commemorate World News Day, a global campaign aimed at raising awareness of fact-based journalism.

2. These commentaries highlighted the declining viewership of mainstream media[1], especially among the younger generation. Reasons put forth for the decrease included (a) reduced trust in mainstream media, especially when the reported content contradicted existing beliefs; (b) shorter attention spans to read longer form content; and (c) the rise of social media platforms as alternative news sources. Despite the declining viewership, mainstream media still invested time and resources to debunk falsehoods, while also publishing timely and accurate news for its audience. This led to mainstream journalists feeling under-appreciated and perceptions that mainstream media was doomed.

3. In his commentary for Singapore's *CNA*, Nicholas **Fang**, Managing Director of Black Dot Research and member of the ADMM Cybersecurity and Information Centre of Excellence (ACICE) Experts Panel, emphasised that mainstream media remained an important source of fact-based journalism. Fang reasoned that mainstream media had a responsibility to focus on fair and objective reporting as it had "editorial oversight and integrity, and multiple layers of fact-checking and editing". He also envisioned that mainstream media had a role to play in establishing a credible and trusted mainstream media ecosystem through partnerships with other stakeholders such as the government and civil society.

Advance Warnings in the Fog of War

4. Under customary international humanitarian law (IHL), parties in a conflict should give effective advance warning of attacks which might affect the civilian population. However, under the fog of war, civilians might face difficulties ascertaining the intention and legitimacy of such advance warnings. This might lead to confusion, and could have implications for civilian safety.

5. On 23 Sep 2024, Lebanese media reported that residents in South Lebanon had received text messages urging them to move away from Hezbollah

[1] According to Cambridge Dictionary, mainstream media refers to traditional forms of media such as newspapers, television and radio, that influence large numbers of people and are likely to represent generally accepted beliefs and opinions.

armouries. On the X platform, Israeli army spokesman Daniel **Hagari** had also “ask[ed] residents of Lebanese villages to pay attention to the message and warning published by the [Israeli military] and heed them”.

6. At the same time, Lebanon's Information Minister Ziad **Makary** claimed that the messages “[came under] the framework of the psychological war implemented by the enemy” and urged residents “not to give the matter more attention than it deserves”.

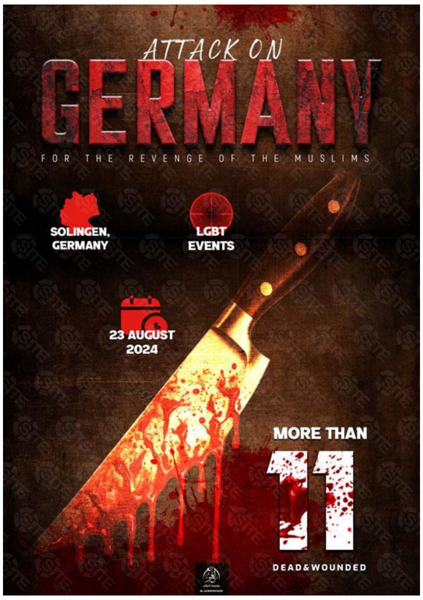
7. The contradictory messaging from Hagari and Makary could confuse civilians, which might have implications for civilian safety. This example also showed how advance warnings could be construed in different ways by actors interested in achieving their own objectives, making it difficult to establish the intention and legitimacy of such warnings in the fog of war.

Terrorism

Incitement to Conduct Lone Wolf Attacks and Attacks Against Shipping Routes

1. On 19 Sep 2024, Islamic State Khorasan Province (ISKP)-linked *al-Azaim Media Foundation* published the 39th issue of the English edition of "Voice of Khurasan" magazine. The 82-page periodical consisted of eight articles and 13 infographics. ISKP incited attacks on maritime shipping routes and urged lone wolves to take bladed weapons and conduct attacks.

2. There was also a promotion of the Aug 2024 knife attack in Solingen, Germany, and a call for jihadists to rise and mount operations. The incitement poster depicted a wall of knives, machetes, hatchets, and shovels.



Incitement poster glorifying the Aug 2024 knife attack in Solingen, Germany (Source: RTL Deutschland)

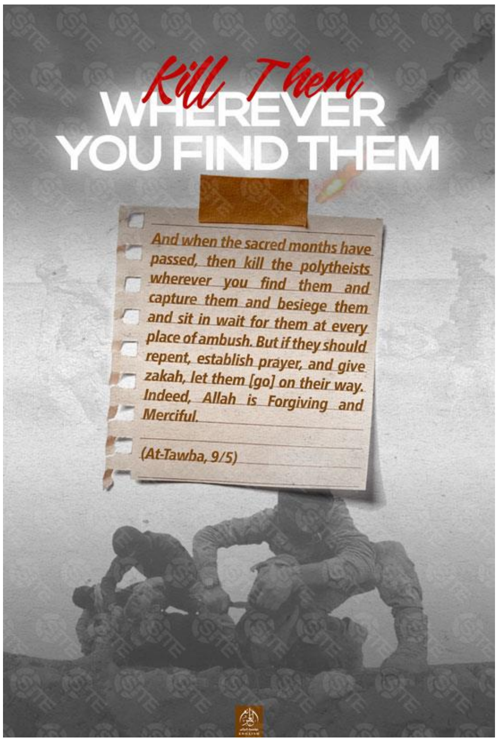
ISIS Promoted an Attack Campaign

3. On 30 Sep 2024, the Islamic State Khorasan Province (ISKP) *al-Azaim Media Foundation* promoted an attack campaign and summoned lone wolves to conduct attacks.

4. The poster, which was published in English and Turkish, displayed a screen capture from an IS video of fighters beheading captives beneath the call to action: "Kill Them Wherever You Find Them."

5. It also showed a Qur'anic verse reading: “And when the sacred months have passed, then kill the polytheists where you find them and capture them and besiege them and sit in wait for them at every place of ambush. But if they should repent, establish prayer, and give zakah, let them (go) on their way. Indeed, Allah is Forgiving and Merciful.”

6. Notably, this was a repetition of ISIS-Central’s “Kill Them Wherever You Find Them” campaign in Jan 2024.


Kill Them
WHEREVER
YOU FIND THEM
And when the sacred months have passed, then kill the polytheists wherever you find them and capture them and besiege them and sit in wait for them at every place of ambush. But if they should repent, establish prayer, and give zakah, let them [go] on their way. Indeed, Allah is Forgiving and Merciful.
(At-Tawba, 9/5)

ISKP Incitement Poster

ISIS Online Recruitment

7. On 24 Sep 2024, an ISIS-aligned media unit, *Dera' al-Sunni (Sunni Shield) Media Foundation* issued a 2-minute, 13-second recruitment video to encourage supporters to join its ranks, including those that lacked the requisite skills.

8. The video depicted a motion infographic for a job posting, and the media unit advertised various roles that jihadists could fill, including video production, reports, visual and audio chants, news bulletins, poetry, articles, and research.

9. Jihadists who did not have any background in media were also encouraged to assist in distributing the materials on social media.

REFERENCES

Cybersecurity

Data Breach at Cybersecurity Giant Fortinet

1. Notice of Recent Security Incident | Fortinet Blog
   https://www.fortinet.com/blog/business-and-technology/notice-of-recent-security-incident

2. Fortinet Suffers Third-Party Data Breach Affecting Asia-Pacific Customers
   https://www.cyberdaily.au/security/11098-fortinet-suffers-third-party-data-breach-affecting-asia-pacific-customers

3. Fortinet Confirms Data Breach After Hacker Claims To Steal 440GB of Files
   https://www.bleepingcomputer.com/news/security/fortinet-confirms-data-breach-after-hacker-claims-to-steal-440gb-of-files/

4. Cybersecurity Giant Fortinet Discloses a Data Breach
   https://securityaffairs.com/168332/data-breach/fortinet-disclosed-a-data-breach.html

5. “Fortibitch Hack – Wow” on Reddit
   https://www.reddit.com/r/fortinet/comments/1ffykqd/fortibitch_hack_wow/?rdt=34375

6. H4ckManac on X
   https://x.com/H4ckManac/status/1834183362786799684

7. Fortinet Confirms Customer Data Breach via Third Party
   https://www.darkreading.com/cloud-security/fortinet-customer-data-breach-third-party

8. ‘Combat Cloud’: EDA Study Shows Benefits of Cloud Computing for EU Militaries
   https://eda.europa.eu/news-and-events/news/2024/01/25/combat-cloud-eda-study-shows-benefits-of-cloud-computing-for-eu-militaries

9. US DOD Makes Headway on Cloud Computing
   https://www.defense.gov/News/News-Stories/Article/Article/3345260/

Cloud Infrastructure and Simple Tactics Used in a Spyware Campaign

10. Undetected Android Spyware Targeting Individuals In South Korea
    https://cyble.com/blog/undetected-android-spyware-targeting-individuals-in-south-korea/

New Backdoor Used by Patchwork to Evade Detection

11. Nexe Backdoor Unleashed: Patchwork APT Group's Sophisticated Evasion of Defenses
    https://cyble.com/blog/nexe-backdoor-unleashed-patchwork-apt-groups-sophisticated-evasion-of-defenses/

Artificial Intelligence

New Artificial Intelligence (AI) Security Governance Framework Unveiled by China

1. Wang Yi on Global AI Governance: Ensure that AI is a Force for Good, Ensure Safety and Ensure Fairness
   https://www.mfa.gov.cn/eng/wjbzhd/202403/t20240308_11256430.html

2. China Releases Security Governance Framework Concerning AI
   https://english.www.gov.cn/news/202409/10/content_WS66df9f30c6d0868f4e8eac91.html

3. Governance Framework Promotes AI Security
   https://www.chinadaily.com.cn/a/202409/11/WS66e0f426a3103711928a74cd.html

4. China Releases AI Safety Governance Framework
   https://www.dlapiper.com/en/insights/publications/2024/09/china-releases-ai-safety-governance-framework

5. China: TC260 Releases AI Safety Governance Framework
   https://www.dataguidance.com/news/china-tc260-releases-ai-safety-governance-framework

6. China's TC260: AI Safety Governance Framework
   https://www.tc260.org.cn/front/postDetail.html?id=20240909102807

7. Full Text: Shanghai Declaration on Global AI Governance
   https://www.mfa.gov.cn/eng/xw/zyxw/202407/t20240704_11448351.html

8. High-Level Summary of the EU AI Act
   https://artificialintelligenceact.eu/high-level-summary/

9. Guidelines for Secure AI System Development
   https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development

First "Reasoning" Model, ChatGPT o1, Released by OpenAI

10. AI Hallucinations Invade OpenAI Latest GPT Model o1 in Quite Surprising Places
    https://www.forbes.com/sites/lanceeliot/2024/09/23/ai-hallucinations-invade-openai-latest-gpt-model-o1-in-quite-surprising-places/

11. Poetic Irony That OpenAI o1 New Superpower Artificially Convinces You That Pigs Can Fly
    https://www.forbes.com/sites/lanceeliot/2024/09/25/poetic-irony-that-openai-o1-new-superpower-artificially-convinces-you-that-pigs-can-fly/

12. It's Useful That the Latest AI Can 'Think', but We Need To Know Its Reasoning
    https://www.theguardian.com/commentisfree/2024/sep/28/openai-o1-strawberry-chain-of-thought-chatgpt

13. OpenAI Releases New o1 Reasoning Model
    https://www.theverge.com/2024/9/12/24242439/openai-o1-model-reasoning-strawberry-chatgpt

14. OpenAI o1 Hub
    https://openai.com/o1/

15. Watch Out For OpenAI New o1 Model Losing Its Train Of Thought
    https://www.forbes.com/sites/lanceeliot/2024/09/26/irked-about-openai-new-o1-model-losing-its-train-of-thought/

16. OpenAI's o1 "Strawberry" ChatGPT Model Can Reason — and Comes With Risks
    https://www.vox.com/future-perfect/372843/openai-chagpt-o1-strawberry-dual-use-technology

17. 'In Awe': Scientists Impressed by Latest ChatGPT Model o1
    https://www.nature.com/articles/d41586-024-03169-9

18. Learning to Reason with LLMs
https://openai.com/index/learning-to-reason-with-llms/

Information

World News Day: Choose Truth.

1. World News Day
https://worldnewsday.org/

2. This World News Day, The Straits Times Renews Its Commitment to the Community
https://www.straitstimes.com/opinion/this-world-news-day-the-straits-times-renews-its-commitment-to-the-community

3. Commentary: Does the World Still Need News Media?
https://www.channelnewsasia.com/commentary/world-news-day-social-media-clickbait-culture-journalism-relevance-4640586

4. Op-Ed From AFP’s CEO Fabrice Fries for World News Day 2024
https://www.afp.com/en/inside-afp/op-ed-afps-ceo-fabrice-fries-world-news-day-2024

5. The Irish Times View on World News Day: A Pillar of Democracy
https://www.irishtimes.com/opinion/editorials/2024/09/26/the-irish-times-view-on-world-news-day-a-pillar-of-democracy/

6. Mainstream Media | Cambridge Dictionary
https://dictionary.cambridge.org/dictionary/english/mainstream-media

Advance Warnings in the Fog of War

7. IDF (Israel Defense Forces) on X
https://x.com/IDF/status/1838081162129940615/video/1

8. Israel Warns Civilians To Evacuate As It Strikes Wide Swaths of Southern Lebanon
https://www.npr.org/2024/09/23/g-s1-24128/israel-tells-lebanese-to-leave-area-where-hezbollah-stores-arms-as-it-launches-strikes

9. Customary International Humanitarian Law - Rule 20. Advance Warning
https://ihl-databases.icrc.org/en/customary-ihl/v1/rule20

10. Do Military Leaflets Save Lives or Just Instill Fear?
    https://www.sapiens.org/language/military-leaflets-warfare-language/

Terrorism

1. Terrorists Celebrate Solingen Attack – and Threaten More Murders in Germany
   https://www.rtl.de/cms/terroristen-feiern-solingen-anschlag-und-drohen-mit-weiteren-morden-in-deutschland-5094391.html