

Could AI Enable Catastrophic Cyberattacks on the US Power Grid?

Matthew van der Merwe



GovAI

Author and Affiliation

Matthew van der Merwe

GovAI

This work represents the views of its author rather than the views of the organization, and does not constitute legal advice. This work has been peer reviewed. Any remaining errors are the author's.

Acknowledgments

I'm grateful to Tom Alrich, Markus Anderljung, Adam Bales, Asher Brass-Gershovich, Rocco Casagrande, Rebecca Ceppas de Castro, Michael Chen, Christian Chung, Stephen Clare, Shaun Ee, Quindi Franco, Ben Garfinkel, Wesley Hurd, Tom Johansmeyer, Holden Karnofsky, Kamile Lukošiuūtė, Omer Nevo, Kevin B. Perry, Frank Rusco, Byron Tomes, Jessica Wang, and Bridget Williams for valuable conversations and comments on drafts. Thank you to Sangeet Brar and José Luis León Medina for editorial assistance. I'm especially grateful to John Halstead, Connor Aidan Stewart Hunter, Luca Righetti, and Elias Groll for their extensive contributions to the report. Finally, thank you to an anonymous peer reviewer, whose comments substantially improved this report. The views expressed in this report, and any remaining errors, are my own.

Contact

Matthew van der Merwe (matthew.vandermerwe@governance.ai)

Please Cite as

van der Merwe, Matthew. *Could AI Enable Catastrophic Cyberattacks on the US Power Grid?* GovAI, 2026.

About GovAI

GovAI is a 501(c)(3) non-profit organization. Our mission is to help decision makers navigate the transition to a world with advanced AI by producing rigorous research and fostering talent. Researchers at GovAI work on a wide range of topics, with a particular emphasis on the security implications of frontier AI.

Abstract

This report analyzes one concrete threat model of AI misuse – an AI-enabled cyberattack on the US power grid causing \$100 billion in economic damages – and asks how this scenario should inform AI risk assessment. The report estimates that such an attack would require a blackout affecting around 100 million people for roughly a week, which is four orders of magnitude more disruption than any grid cyberattack on record. Analysis of two attack pathways – physically damaging critical grid equipment and both triggering and sustaining a cascading blackout – reveals no single narrow technical bottleneck. Rather, the binding constraint appears to be the operational capacity to coordinate diverse capabilities in attacks on dozens to hundreds of targets. The report therefore concludes that the \$100 billion scenario is poorly suited to triggering costly, scenario-specific mitigations. AI systems capable of enabling lower-skilled actors to launch \$100 billion grid cyberattacks would likely also pose more serious risks in other domains. Lowering the barriers for top-tier states – which may already possess the required capacity – would have an unclear effect on risk, since their willingness to launch such attacks outside active conflict appears significantly constrained by the prospect of retaliation or escalation. \$10 billion grid cyberattacks appear qualitatively easier, and may provide more useful capability thresholds for triggering mitigations. The \$100 billion grid cyberattack scenario remains useful for understanding the broader range of cyber risks to critical infrastructure, including those at lower damage thresholds.

Executive Summary

Many frontier AI companies have committed to mitigating severe risks posed by their systems. Cyberattacks against critical infrastructure have been cited as one such risk. This report examines the extreme end of critical infrastructure risk through one concrete scenario: a single AI-enabled cyberattack on the US grid causing at least \$100 billion in economic damages, a threshold based on the definition of severe harm in OpenAI's Preparedness Framework.

The report draws two key conclusions:

1. An AI-enabled cyberattack on the US grid causing at least \$100 billion in damages appears to require capabilities and operational capacity beyond anything demonstrated in documented cyberattacks against infrastructure – though top-tier states may be capable of carrying out such attacks.
2. The \$100 billion scenario appears poorly suited to triggering costly, scenario-specific mitigations. A \$10 billion grid attack – comparable to the worst blackouts in US history – would represent a major regional emergency. Such an attack appears qualitatively easier to carry out, suggesting that lower-damage scenarios may provide more decision-relevant capability thresholds.

A \$100 Billion Blackout Would Be Unprecedented

Three cyberattacks have caused blackouts by manipulating grid control systems – all against Ukraine and attributed to state-linked actors. The worst of these attacks caused estimated economic damages on the order of \$1 million, while the worst *accidental* blackouts in the US have caused damages on the order of \$10 billion.

Reaching the \$100 billion threshold requires an outage affecting 100 million people for around a week – four orders of magnitude more disruption than any prior grid cyberattack has achieved. [Figure ES1](#) shows the combinations of outage duration and affected population required to meet the \$10 billion and \$100 billion thresholds.

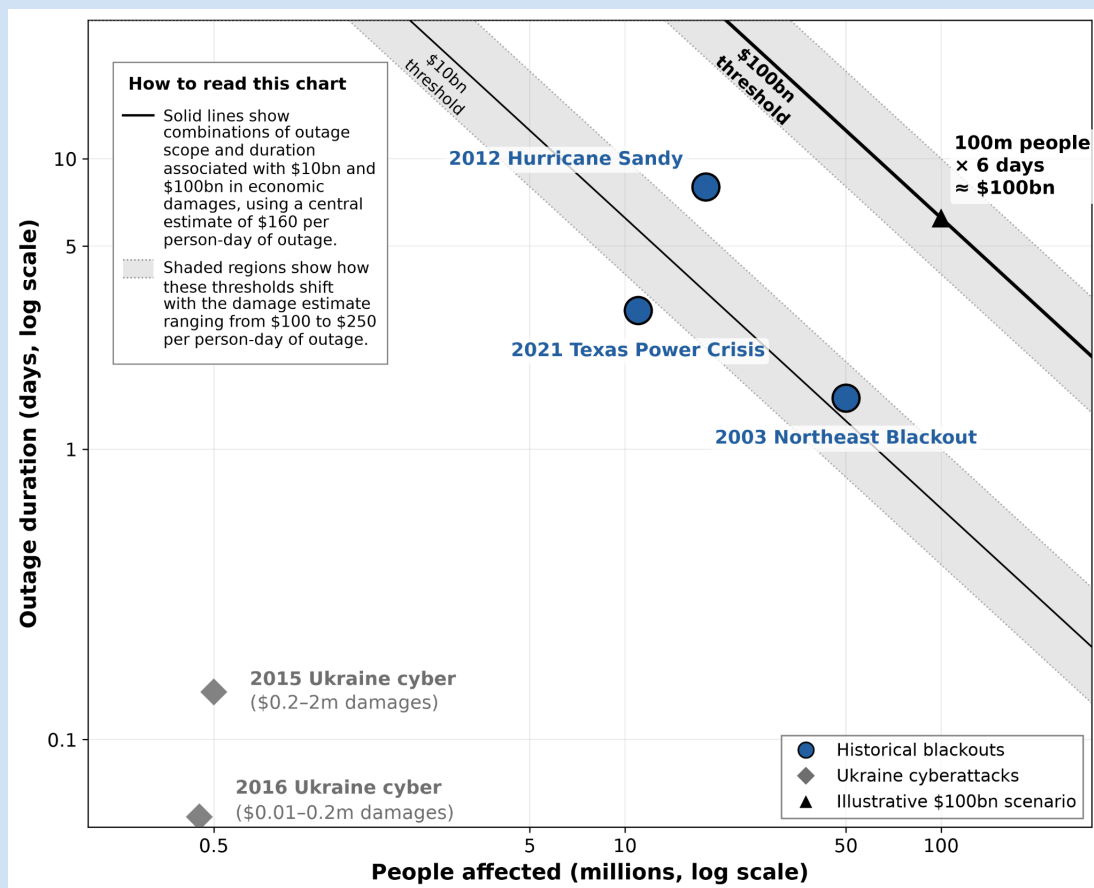


Figure ES1 | \$100 billion in economic damages would require a blackout affecting 100 million people for around a week (or equivalent combinations of scope and duration).
Historical blackouts and cyberattacks are shown for comparison.

In a pilot survey of 8 domain experts and 13 superforecasters – conducted with the Forecasting Research Institute – the median estimated risk of a \$100 billion grid cyberattack in 2026 was 0.1% in both groups, and 1% for a \$10 billion attack.

The Hard Problem Is Sustaining a Blackout

The challenges of sustaining a blackout make it difficult to cause \$100 billion in damages from a cyberattack on the US grid. Triggering a large cascading blackout via cyberattack, while unprecedented, could in principle be achieved via a handful of well-chosen disturbances under the right conditions. Yet blackouts resolve relatively quickly by default. Absent serious physical damage to equipment, power is restored within hours or, at most, a few days. Causing a week-long blackout of this scale would require delaying recovery, either by physically damaging enough critical grid equipment or by

thwarting efforts to restore power over several days. Either route would demand extreme scale – damaging dozens or more large transformers, at least 100 generators, or thwarting dozens of parallel restoration efforts. This would far exceed the effects demonstrated in any prior cyberattack. Notably, causing \$10 billion in damages via a grid cyberattack appears qualitatively easier, since this may be achievable by triggering a cascading blackout without the additional work involved in sustaining it for a week.

There Does Not Appear to Be Any Single Narrow Technical Bottleneck

The major documented cyberattacks on operational technology (OT), including Stuxnet and the Ukraine grid attacks, are all attributed to state actors. Each required months or years of work by skilled teams of operators, and achieved disruptive effects at one to three targets – an order of magnitude fewer than seem to be required for the \$100 billion grid cyberattack scenario.

What distinguishes the most capable state actors is not simply narrow technical skills. Indeed, many of the impressive technical elements in these documented OT attacks appear accessible to lower-skilled actors. Rather, top-tier states are distinguished by the operational capacity to integrate a diverse range of capabilities in complex, dynamic operations.

The Most Capable Actors Appear Least Willing

Were a state to launch a \$100 billion grid cyberattack against the US, this would invite swift and severe retaliation. Outside of conflict, strategic logic therefore limits the willingness of state actors to pursue such attacks. In contrast, individuals and extremist groups regularly attack the US grid, overwhelmingly via physical means rather than cyber. The sheer number of such lower-skilled actors suggests that, were they capable, at least some would attempt catastrophic grid cyberattacks.

Actors' motivation also depends on the availability of alternatives. To date, kinetic attacks against the grid have been considerably more effective at causing disruption than cyberattacks. A 2022 rifle attack on electrical substations in North Carolina caused a larger outage than Russia's 2015 and 2016 grid cyberattacks on Ukraine combined. That said, cyberattacks retain significant advantages, especially for state actors outside of wartime – they are generally more deniable, less escalatory, and can be conducted at a distance.

AI Could Raise the Risk of Grid Cyberattacks via Two Channels

The clearest route through which AI could increase the risk of catastrophic grid cyberattacks is by enabling lower-skilled actors. These actors are currently unable to launch such attacks but appear more willing to do so – if they could. Uplifting these actors would require AI systems capable of automating much of the work of a state-level offensive cyber operation.

This level of uplift, however, is an extremely high bar. AI systems meeting this threshold would plausibly also be capable of automating other complex, long-horizon endeavors currently requiring teams of experts. Such powerful systems could have consequences far more significant than any increased risk of grid cyberattacks, such as accelerating AI research and development, lowering barriers to catastrophic biological attacks, or enabling widespread labor displacement.

By making cyber operations cheaper, faster, more reliable, or more scalable, increasingly capable AI systems may lower the barriers to state actors carrying out catastrophic grid cyberattacks. However, since state actors appear constrained at least as much by strategic logic as by capability, it is less clear whether this uplift would translate into increased overall risk. The most concerning narrow capabilities may be those that would shift their strategic calculus more directly – for example, by reducing the risk of attribution.

To be sure, whether greater AI capabilities translate into increased risk from this scenario will depend on the relative gains to attackers and defenders, and the pace of adoption on each side. This report analyzes how AI could lower barriers for attackers, holding defensive capabilities fixed. More powerful AI could help defenders harden their systems, detect intrusions, and respond to incidents. The analysis is therefore only one input into an assessment of the overall marginal risk from AI cyber capabilities.

How This Analysis Should Inform AI Risk Assessment

When determining whether a model is safe to release, it does not appear particularly useful to ask whether that model would materially increase the risk of a \$100 billion grid cyberattack. However, lower-damage scenarios may provide more useful near-term thresholds. To understand how risk is developing, decision makers should track evidence of AI lowering barriers to grid attacks, such as evidence of AI-enabled OT intrusions in the wild. In the pilot survey, respondents updated risk estimates far more strongly on the basis of hypothetical real-world incidents than on model performance in

simulated attacks – suggesting that real-world warning shots are stronger evidence than evaluations.

Studying catastrophic grid attacks remains useful in other ways. By analyzing this scenario in depth, we identify OT-specific capabilities and lines of evidence that can feed into ongoing assessments of a broader range of cyber risks to critical infrastructure. Lower-cost interventions, such as monitoring real-world incidents, evaluating model performance on OT-specific capabilities, and implementing model-level safeguards, are likely worthwhile, especially insofar as they address this broader class of threat models. To specifically address catastrophic cyberattacks on the grid, the most robust mitigations may be investments in grid resilience, such as increasing stockpiles of critical equipment and strengthening response and recovery capacity.

Limitations of This Analysis

This report's conclusions are limited by several scoping decisions, in particular the \$100 billion threshold and the exclusion of non-economic damages from the analysis. For example, grid cyberattacks during military crises could have severe strategic consequences even when economic damage remains small.

The report's claims about the bottlenecks to grid cyberattacks rest on a limited evidence base. The public track record of OT cyberattacks is sparse, lagging, and biased in important respects, limiting the strength of conclusions about the capabilities of state actors. The analysis of attack pathways is heavily anchored in historical grid failures. Grid modernization – greater reliance on digital devices, and an increasing share of renewables – could introduce new attack surfaces and easier routes to disruption. The clearest disconfirming evidence would be a real-world attack succeeding with far lower levels of operational capacity than this report argues is required.

Contents

Executive Summary..... 3

Structure of the Report..... 10

1. Introduction..... 11

1.1 Background..... 11

1.2 Scope of the Analysis.....12

1.3 Motivation for Prioritizing This Threat Scenario..... 13

1.4 Methodology..... 16

2. Track Record and Baseline Risk of Grid Cyberattacks and Major Blackouts..... 20

2.1 Baseline Risk of Grid Cyberattacks.....20

2.2 Track Record of Major Accidental Blackouts.....23

2.3 Baseline Risk of \$100 Billion Grid Cyberattacks.....24

3. What Would It Take for a Blackout to Cause \$100 Billion in Damages?..... 26

3.1 Types of Blackout.....26

3.2 Relationship between Damages, Scope, and Duration..... 28

4. How Could a Cyberattack Cause Blackouts of This Scale?..... 33

4.1 Physical Damage.....34

4.2 Disrupting Grid Operations..... 42

4.3 Supply Chain Compromises.....48

4.4 Summary of Both Pathways..... 51

4.5 Implications for a \$10 Billion Damage Threshold..... 52

5. The Resources and Capabilities Required for a Catastrophic Grid Cyberattack54

5.1 Prior OT Cyberattacks Have Required Significant Resources..... 54

5.2 Prior OT Cyberattacks Relied on Diverse Capabilities..... 58

5.3 A \$100 Billion Grid Cyberattack Would Require Unprecedented Capabilities and Scale.... 60

5.4 Implications for a \$10 Billion Threshold..... 61

6. AI Uplift Analysis.....62

6.1 Methodology..... 62

6.2 Baseline Capability and Willingness of Threat Actors..... 64

6.3 AI Uplift Pathways..... 69

7. Implications..... 74

7.1 Risk Prioritization..... 74

7.2 Risk Monitoring..... 75

7.3 Mitigations..... 78

7.4 Summary of Implications..... 79

8. Limitations and Uncertainties..... 80

8.1 Scope..... 80

8.2 Evidence Base and Methodology..... 80

8.3 Modeling Assumptions..... 81

9. Conclusion..... 83

About the Author..... 85

Appendices..... 86

Appendix 1 | How the Grid Works..... 86

Appendix 2 | Estimating Damages from Ukraine Cyber Blackouts..... 88

Appendix 3 | European Blackout Data..... 89

Appendix 4 | Methodology for VOLL-based Estimate..... 90

Appendix 5 | Historical OT Cyberattacks..... 94

Appendix 6 | Resource Requirements for Prior OT Cyberattacks..... 95

Appendix 7 | Threat Actor Classification Methodology..... 97

Appendix 8 | Additional Survey Results..... 98

Appendix 9 | Kinetic vs. Cyber Grid Attacks..... 101

Appendix 10 | Other Attack Pathways Not Considered in Depth..... 102

Appendix 11 | Grid Cyberattack Case Studies..... 103

References..... 113

Structure of the Report

The main report proceeds as follows:

- [Section 1](#) provides background context and motivation, and outlines the scope of the report's analysis and the methodology.
- [Section 2](#) reviews the track record of cyberattacks against the grid, and of major accidental blackouts.
- [Section 3](#) estimates the scale of blackout required to exceed \$100 billion in economic damages.
- [Section 4](#) considers two broad pathways through which a cyberattack could cause a \$100 billion blackout:
 - [Section 4.1](#) considers attacks via physical damage to critical grid equipment
 - [Section 4.2](#) considers attacks via disrupting grid operations
 - [Section 4.3](#) considers the role of supply chain compromises across both pathways
- [Section 5](#) estimates the resources and capabilities required to successfully execute attacks on either of the above pathways, drawing heavily on historical case studies of cyberattacks against industrial targets.
- [Section 6](#) considers whether and how AI could meaningfully lower the barriers to catastrophic cyberattacks against the US grid.
- [Section 7](#) discusses implications for decision makers on risk prioritization, risk monitoring, and mitigations.
- [Section 8](#) discusses limitations of the analysis and major uncertainties.
- [Section 9](#) draws overall conclusions.

1. Introduction

1.1 Background

To address emerging cyber risks from AI systems, many frontier AI companies have defined capability thresholds in their safety policies. These are if-then commitments of the form: If a model passes some threshold of capability deemed to be indicative of unacceptable risk, then the company commits to mitigating that risk prior to releasing the model ([METR 2024](#); [Karnofsky 2024](#)). Companies have committed to testing whether these capability thresholds have been crossed using model evaluations and monitoring of real-world usage ([Rodriguez et al. 2025](#); [Anthropic 2025a](#); [Bhatt et al. 2024](#); [OpenAI 2025a](#)).

Capability thresholds in AI safety policies are designed to manage only a subset of particularly severe or catastrophic risks related to frontier AI systems ([Anthropic 2025b.1](#); [Meta 2025.12](#); [OpenAI 2025b.5](#); [Google DeepMind 2025b.2](#)). There may be other risks that AI companies ought to and do mitigate. For example, AI companies do not want their systems to incite hateful behavior, though these smaller-scale risks would not be covered by their safety policies, which focus on catastrophic risks specifically ([Anthropic 2025b.1](#)).

There is not yet agreement among AI companies on which cyber capability thresholds, if any, ought to be included in AI safety policies. However, the risk of AI-enabled cyberattacks on critical infrastructure has been highlighted by AI companies and the national security community ([Google DeepMind 2025a.5](#); [Department of Homeland Security 2024b.16–17](#); [Idaho National Laboratory 2025](#)). Nonetheless, no published threat model yet identifies the pathways by which AI cyber capabilities could increase the risk of critical infrastructure attacks, or by how much. As a result, it is difficult to know how current and future AI systems might affect the risks of cyberattacks on critical infrastructure, and it is difficult for companies to prioritize risk mitigations. For the same reason, policymakers lack foresight into emerging cyber risks.

This report aims to help fill this gap by exploring one critical infrastructure cyberattack threat model in depth. Specifically, it explores the plausibility of the following scenario:

A large-scale AI-enabled cyberattack on the US electricity grid causing a major blackout that leads to more than \$100 billion in economic damages.

The scope of this report is therefore limited: It focuses only on one type of critical infrastructure and uses a very high bar for economic damages. Nevertheless, it aims to shed light on other critical infrastructure cyberattack threat models.

1.2 Scope of the Analysis

There are many potential ways in which AI models could be misused for cyberattacks. Following [NIST \(2025, Appendix E\)](#), threat models can be categorized by: (1) the type of cyberattack; (2) the relevant capability that could enable such attacks; (3) the threat actor pursuing it; and (4) the potential negative outcomes that might result. [Table 1.1](#) summarizes which threat models are in and out of scope in this report.

	Type of Attack	Key Capability	Threat Actor	Outcome
In Scope	Single large coordinated cyberattack on the US grid: A cyberattack physically damaging grid equipment, or precipitating a large cascading blackout and preventing grid operators from restoring power for a week.	All steps in the cyber kill chain: This report’s analysis suggests that there is not a single narrow bottleneck to large-scale cyberattacks on the grid, and that a diverse range of tasks are necessary to execute an attack.	All	Severe economic damages (≥\$100 billion): The immediate costs to firms and consumers from economic disruption. This could include lost revenue, reduced consumption, remediation and costs borne by insurers.
Out of Scope (selected examples)	Multiple smaller grid attacks that might in aggregate have severe economic costs.	–	–	Long-run economic costs: Counterfactual GDP loss over months and years following an outage.
	Cyberattack on non-US grids.			Non-economic welfare costs: Costs on other determinants of human welfare, such as health. E.g. attacks on the grid might cause failures at hospitals leading to loss of life, but these effects are treated as out of scope.
	Cyberattack on non-grid critical infrastructure such as the water supply or telecoms networks.			
	State espionage: Cyberattacks by states in order to access sensitive information (e.g. 2020 SolarWinds). Targeted ransomware attacks: Attacks by cybercriminals to extort private businesses			Geopolitical costs: Attacks on the grid, especially those carried out by state actors during conflicts, could have important geopolitical effects, but these effects are treated as out of scope.

Table 1.1 | Different AI-cyber threat models.

Threat Actors: Differentiating by Operational Capacity

The analysis is structured around different threat actor classes, ranging from individual novices to the most well-resourced states. Actors are categorized by their operational capacity, following the framework introduced in [RAND \(2024\)](#). These actor classes are defined in more depth in [Section 6](#).

Threat Actor	Description	Examples	Est. population ¹
TA1	Single individual, limited infosec expertise, <\$1k budget for the specific operation, no infrastructure.	Hobbyist hacker	~1m
TA2	Single individual, professional infosec capability, ~\$10k budget for the specific operation, personal infrastructure.	Individual professional hackers	10k–100k
TA3	Team of ~10 experienced professionals, ~\$1m budget and months of effort for the specific operation.	Criminal hacking groups	100–1k
TA4	Team of ~100 state-level experts, ~\$10m budget and a year or more of effort for the specific operation, state resources.	Second-tier states (e.g. Iran, North Korea)	10–50
TA5	Team of ~1,000 top experts, ~\$1bn budget and years of effort for the specific operation, state-level resources, infrastructure, and access.	Top-tier states (e.g. US, China)	~5

Table 1.2 | Threat actor definitions. See [Appendix Z](#) for a more detailed explanation.

1.3 Motivation for Prioritizing This Threat Scenario

It is important to stress the narrow scope of the scenario analyzed in this report. Four points are particularly noteworthy. First, the report focuses only on cyberattacks on the electricity grid, not other types of critical infrastructure. Second, it focuses on scenarios involving a single coordinated attack on the grid, rather than multiple smaller attacks which may in aggregate have severe economic costs. Third, the report focuses on particularly severe economic damages of more than \$100 billion. Cyberattacks on the grid could cause lower levels of economic damage. The plausibility of \$10 billion attacks is briefly discussed, though these are not the primary focus of the analysis. Fourth, the report focuses only on immediate economic damages rather than other types of negative effects of grid cyberattacks such as loss of life or geopolitical consequences.

¹ These estimates are from the survey of experts and forecasters, discussed in [Section 1.4](#).

Why Focus on Grid Cyberattacks Rather Than Other Critical Infrastructure Attacks?

The grid cyberattack scenario falls within the broader class of cyberattacks against critical infrastructure. Such attacks have been highlighted by governments and AI companies as a pathway by which AI could enable significant harms via cyber misuse ([Department of Homeland Security 2024a](#); [UK AISI 2024](#); [Google DeepMind 2025a, 5](#)).²

The scope of this report was narrowed to grid cyberattacks for two main reasons. First, sustained loss of power appears particularly damaging among critical infrastructure disruptions, since most other critical services themselves depend on electricity. For example, the 2025 UK National Risk Register deems a failure of the UK transmission network one of three “accidents or systems failures” with catastrophic impact ([Cabinet Office 2025, 16](#)).³ Second, the analysis can be empirically grounded in real-world events, including several grid cyberattacks that have caused blackouts, other failed grid cyberattacks, and a large number of accidental blackouts.

While this report does not directly analyze attacks on other critical infrastructure, it may shed light on this broader class. The barriers identified are clearest for attacks that, like grid attacks, seek to achieve physical effects via manipulating operational technology (OT). Other OT-dependent sectors (e.g. water, oil and gas, rail) share many of the same obstacles, though they differ in architecture, defensive posture, recovery dynamics, and cross-sectoral dependencies.

However, this report’s conclusions are less applicable to critical infrastructure attacks that do not depend on disrupting OT. This includes attacks where upstream disruption of IT has downstream effects on infrastructure services (e.g. NotPetya, Colonial Pipeline), and attacks on infrastructure that is itself primarily digital, such as cloud providers and telecommunications.

Why Focus on a Single Coordinated Attack Rather Than Multiple Smaller Attacks?

The analysis focuses on scenarios in which a threat actor causes a \$100 billion blackout through a single coordinated attack on the grid with disruptive effects over a short period of time (i.e. several days), rather than multiple grid attacks over the course of weeks to months. This focus reflects one criterion used in several safety frameworks to demarcate catastrophic risks: whether the resulting harms are “instantaneous or irremediable” ([Meta 2025, 12](#); [OpenAI 2025b, 5](#)). A single instantaneous attack could occur with little to no warning. In contrast, several smaller-scale “warning shot” attacks would create opportunities to adapt and respond.

² Version 2 of Google DeepMind’s Frontier Safety Framework defines a capability threshold around AI uplifting threat actors to carry out “high impact” attacks, citing, as an example, “the creation of predictable and large-scale effects on critical national infrastructure” ([Google DeepMind 2025a, 5](#)). However, version 3 does not mention risks to critical infrastructure ([Google DeepMind 2025b](#)).

³ Alongside accidental radiation release from UK or overseas nuclear facilities.

Such incidents would likely prompt efforts to reduce the risk of further attacks – for example, by hardening grid security, and applying political or law-enforcement pressure to the actors responsible. If future AI could enable large-scale instantaneous and severe attacks, there is a stronger case that AI developers should mitigate the risk prior to deployment.

Why Use the \$100 Billion Damage Threshold?

This report's primary focus is on grid cyberattacks causing at least \$100 billion in economic damages, a threshold chosen for two initial reasons. When this project began, the published safety policies of two leading AI companies used “hundreds of billions” as an explicit quantitative threshold for catastrophic harms – OpenAI's Preparedness Framework ([OpenAI 2023, 2](#)) and Anthropic's Responsible Scaling Policy ([Anthropic 2023, 1](#)). Moreover, the most prominent and widely cited scenario analysis of grid cyberattacks found that damages could exceed \$100 billion. [Lloyd's and CRS \(2015b\)](#) estimate direct economic damages of \$80 billion to \$300 billion from a hypothetical cyberattack on the Eastern US grid.⁴ Together, these suggested that \$100 billion in damages was both *prima facie* plausible and a salient threshold for policy.⁵

The landscape of risk thresholds in AI policy has since evolved. While OpenAI has retained the \$100 billion threshold ([OpenAI 2025b, 2](#)), Anthropic no longer provides an explicit quantitative threshold ([Anthropic 2025b](#)). A lower threshold of \$1 billion is now used by at least one AI company ([xAI 2025, 3](#)), and in California's SB53 ([California State Legislature 2025](#)). In general, the question of how and where companies should set risk thresholds remains under-researched and unresolved ([Koessler et al. 2024](#)).

The \$100 billion threshold is significant for reasons independent of policy. As [Section 3](#) will establish, the worst accidental blackouts in the US have caused damages on the order of ~\$10 billion, and \$100 billion in damages would require an outage of unprecedented magnitude. As [Section 4](#) will argue, reaching this scale of damages via cyberattack would require not merely causing a large blackout but sustaining it for several days – either through widespread physical damage or prolonged disruption to restoration efforts. While the primary focus throughout is the \$100 billion threshold, the capability requirements for \$10 billion attacks are explored in [Sections 4.5](#) and [5.4](#).

⁴ Lloyd's also estimates cumulative GDP impacts of \$300 billion to \$1.3 trillion over five years. (Nominal damage estimates converted to \$2025). Risk analysis by the UK National Risk Register provides more ambiguous support for the plausibility of this scenario. [Cabinet Office \(2025, 45\)](#) discusses the prospect of a terrorist cyberattack on the electricity grid, leading to a total failure of the National Grid. In the section on grid cyberattacks, the report does not directly state the likelihood and impact of this specific scenario. However, in a section on accidental failure of the National Grid, which seems to have very similar if not identical effects, the potential damages are given a score of 5 in terms of impact ([Cabinet Office 2025, 90](#)), which equates to damages of tens of billions of pounds ([Cabinet Office 2025, 14](#)). The potential damages of an attack on the US grid would be larger.

⁵ Though as discussed below, Lloyd's scenario analysis appears implausible in several respects.

The \$100 billion threshold choice also shapes this report's conclusions. As Sections 6–8 will argue, the capabilities required for such attacks are so advanced and so broad that the \$100 billion scenario is ill-suited to grounding near-term decisions on costly scenario-specific mitigations. Threat models anchored to lower damage thresholds may serve this purpose better.

To be clear, the use of the \$100 billion threshold should not be taken to imply that less damaging attacks are unimportant or fall outside the scope of responsible AI risk management. In analyzing the extreme end of plausible harms, this report hopes to shed some light on a broad range of scenarios. The approach developed here can be adapted to lower thresholds as policy discussions evolve, and further work to extend this report's analysis would be welcomed.

Why Focus on Economic Damages Only?

This report focuses only on the immediate economic damages caused by disruption to the private and public sectors from lost power, since these are most straightforward to model and likely represent a significant fraction of overall costs. This includes effects like reduced consumer power consumption, lost industrial production, and lost business revenue. Long-run economic costs, such as counterfactual GDP loss over the years following an immediate output shock, are not considered, both because of the measurement challenges of forecasting baseline economic activity and to align this report's approach with that used in the literature on the economics of power outages. Nor are the health and mortality costs from power outages.

Grid cyberattacks may have other effects that are harder to quantify. Most notably, grid attacks, particularly those launched by states amid conflict, could plausibly have important geopolitical consequences. Large-scale or targeted grid attacks could materially affect the chances of success during military operations and could lead to military escalation. These sorts of effects could plausibly exceed economic impacts but are much harder to quantify. An adequate treatment of these consequences would likely require national security expertise and access to classified information.

1.4 Methodology

This report aims to understand the marginal or “net new” risk posed by future hypothetical AI capabilities, compared to existing technologies ([Kapoor et al. 2024](#); [NIST 2025](#); [Meta 2025.12](#)). The analysis focuses on how AI could lower barriers for attackers, while holding defensive capabilities fixed. It is therefore one input into a complete estimate of AI's net effect on risk. For AI cyber capabilities to be concerning, they must make a counterfactual difference to whether threat actors can carry out cyberattacks. For example, if a state-level actor can

already launch a grid cyberattack, then a capable AI system might have limited marginal effect on the risk of attacks by this actor.

Thus, this report tries to estimate the baseline risk of catastrophic grid cyberattacks by different threat actors, assuming no further progress in AI. It then analyzes the marginal effect of AI on attackers, assuming that AI models gain certain cyber capabilities. The analysis is structured around a simple model consisting of two parameters – a threat actor’s capability to successfully launch a catastrophic grid cyberattack and their willingness to do so if they were able (see [Section 6](#)). This report uses a simple model with few parameters, as these are more transparent and have demonstrated superior forecasting performance ([Green and Armstrong 2015](#); [Forster and Sober 1994](#); [Stock and Watson 2002](#); [Makridakis and Hibon 2000](#); [Brighton and Gigerenzer 2015](#); [Morgan and Henrion 2012, Ch. 11](#); [Tetlock and Gardner 2015, Ch. 5](#)).

The ultimate aim of this report is to provide an analysis that is well-supported by evidence. To that end, the approach involved:

1. **Empirical research:** Extensive review of case studies of historical cyberattacks and accidental blackouts to ground the analysis in real-world evidence.
2. **Expert feedback:** Expert feedback from cyber, grid, and national security experts who provided comments on earlier drafts.
3. **Expert survey:** A survey of 8 subject-matter experts and 13 credentialed “superforecasters” on the model’s parameters and other questions relevant to the report.

1.4.1 Survey of Experts and Forecasters

A survey of experts and forecasters was conducted in partnership with the Forecasting Research Institute.⁶ The survey covered two threat models – the one examined here and one involving AI-enabled data-damaging computer worms ([Halstead and Righetti 2026](#)). Only the results on grid cyberattacks are presented here.⁷

Forecasts can be highly sensitive to the precise wording of a question and its resolution criteria. To develop the survey questions, we undertook an iterative process in which we drafted initial versions of the question, a small sample of experts and superforecasters answered the questions, and we then revised the questions in light of how they were interpreted by participants. We conducted two rounds of this process. We also provided

⁶ [Singer et al. \(2023\)](#) surveyed 18 grid security experts on the likelihood and impact of four grid cyberattack threats, and demonstrated how differences in modeling assumptions – including grid topology and assumed attacker capabilities – can produce divergent risk assessments.

⁷ See [Ceppas de Castro et al. \(2026\)](#) for comprehensive detail on the survey, including questions on the worm threat model.

participants with an abridged version of an early draft of this report, which excluded our own estimates and rationales to avoid anchoring.

As this was a pilot study, we used a convenience sampling method and aimed for a sample of around 15–30 people.⁸ We invited two groups of respondents:

1. People with expertise in cybersecurity and AI impacts on cybersecurity (henceforth, “experts”).⁹
2. High-performing generalist forecasters, or “superforecasters”, who are people who have previously scored highly in geopolitical forecasting tournaments.

We included generalist forecasters because a well-established body of literature has shown that it is possible to identify skilled generalist forecasters, or “superforecasters”, who reliably make accurate predictions spanning a wide range of domains ([Mellers et al. 2015](#); [Mellers et al. 2014](#); [Tetlock and Gardner 2015](#)). Indeed, in some cases, forecasters have produced comparable or superior performance to subject-matter experts at forecasting, while in other cases experts have outperformed forecasters (see for example [Kučinskas et al. 2025](#)).

A total sample of 33 people was invited to participate via email. To incentivize engagement, we paid participants for their time spent completing the survey.¹⁰ Participants provided responses to the main survey between 23 July and 29 August 2025.

We surveyed participants on questions including:

- **Baseline and conditional probabilities:** Participants estimated the probability of at least one cyberattack against the US electrical grid in 2026 resulting in a blackout causing economic damages of $\geq \$100$ billion and $\geq \$10$ bn respectively. They provided both unconditional baseline forecasts and forecasts conditional on hypothetical AI capability scenarios. We also asked participants to estimate the total expected damages from cyberattacks against the US grid in 2026.
- **Actor capability and willingness:** For each threat actor category and damage threshold, participants estimated the probability that a randomly selected actor in that category could launch such an attack with six months of effort; and, assuming capability, the probability that at least one actor in that category would spend six or more months actively attempting to launch such an attack in 2026. We also asked

⁸ In convenience sampling, participants are recruited on the basis of being available and relatively easy to access.

⁹ Cybersecurity is a particularly broad and varied domain, and the specific expertise relevant to the grid cyberattack threat model is particularly concentrated in subdomains like industrial control systems, grid cybersecurity, and offensive cyber operations. Invited experts were tilted towards some specialty in these subdomains, but also included those with more generalist expertise.

¹⁰ The average payment was \$700.

participants to estimate the probability that, were such an attack to occur, it was caused by each category of threat actor.

This survey has several important limitations. First, the sample size for this pilot study was small: 21 participants completed the survey, including 13 superforecasters, 6 experts who completed the full survey, and 2 experts who completed shorter versions of the survey. This makes the results sensitive to individual forecasts, and the reported aggregate statistics fragile. Second, our convenience sample of experts may be biased in some respects and should not be taken as representative of cybersecurity experts. Third, some of the experts who participated in the survey also provided feedback on earlier drafts of this report, which may also introduce bias. These problems are not a concern for the recruitment of superforecasters.

Fourth, the questions in this study differed from typical forecasting exercises, as they focused on hypothetical AI evaluations and extreme, low-probability events without clear resolvability. Many were conditional questions, meaning forecasters cannot expect to be scored on the accuracy of their predictions. This absence of performance incentives or feedback, combined with the unusual nature of some questions and the lack of past data to guide judgments for certain scenarios, could impact how accurate and useful the forecasts are. Importantly, this may limit the extent to which superforecasters' calibration track record is predictive of accuracy on our survey.

Lastly, we did not share the full report, including our own quantitative estimates and reasoning, with survey respondents. This has the advantage of avoiding anchoring respondents. The main disadvantage is that sharing the full report would have shared more information with the respondents, and, if there were disagreement, it would have been easier to understand why the respondents did not agree with our estimates. There was also no chance for respondents to update their estimates following discussion with each other or with the authors.

Despite these limitations, the survey provides estimates from a wider range of perspectives than our own estimates alone. A larger and more systematic version of the survey would address the limitations mentioned above.

The remainder of the report reviews evidence on the grid cyberattack threat model in depth.

2. Track Record and Baseline Risk of Grid Cyberattacks and Major Blackouts

To better understand this threat model, it is useful to first examine the track record of grid cyberattacks and major accidental blackouts. (For a basic overview of how the grid works and a glossary of key terms, see [Appendix 1](#).)

There appear to have been only three instances of blackouts caused by cyberattacks, all in Ukraine. These blackouts have all been relatively low severity, lasting a few hours, affecting a few hundred thousand people, and causing economic damages on the order of \$1 million. Accidental blackouts have caused considerably greater harms, with economic costs amounting to over \$10 billion in some cases.

2.1 Baseline Risk of Grid Cyberattacks

Blackouts via Cyberattack Are Rare (~0.3/year)

[Table 2.1](#) collects historical grid cyberattacks and their impacts, drawn from open-source reporting. (See [Appendix 11](#) for detailed methodology and sources.)

This search confirmed that cyberattacks causing blackouts are extremely rare, and identified only three cyberattacks in which the manipulation of operational technology resulted in a blackout, all targeting Ukraine.¹¹ In addition, there are two edge cases in which cyberattacks on IT billing systems prevented customers from obtaining electricity.¹²

There are many more cases of cyberattacks against the grid that have not resulted in blackouts. Some of these have resulted in other disruptions, such as grid operators losing visibility or control over equipment. US utilities have reported about four disturbances per year due to cyberattacks, none of which have caused power outages.¹³

Altogether, the baseline frequency of cyber-induced blackouts caused by manipulation of OT is low – 3 in the 10 years between December 2015 and December 2025 (~0.3 per year). Since this count relies on open-source information, it may undercount the true frequency. However,

¹¹ The search revealed one large blackout – in Mumbai, India in 2020 – initially attributed to a cyberattack, but which was subsequently determined to have had a non-cyber cause. See [Sanger and Schmall \(2021\)](#) reporting a cyber link, and [Indian Express \(2021\)](#) on investigation concluding it was accidental.

¹² Incidents in Ghana 2022 and South Africa 2019. These are interesting edge cases, where cyberattacks on billing infrastructure resulted in outages due to features of electricity markets in the developing world, where a large fraction of customers use on-demand meters and were unable to purchase electricity. These are excluded from the base rate, since they did not disrupt grid operations or equipment.

¹³ Cyber events are reported by utilities to the US Department of Energy in Form OE-417 filings ([DOE n.d.](#)). The average number of cyber events across annual reports is 3.92 between 2011 and 2023.

given the inherent visibility of blackouts, it seems unlikely that there is a significant number of additional nonpublic cases.

Target	Date	Actor	Impact	Details
Ukraine	Oct 2022	State	Outage via OT	Outage in one city coinciding with missile strike (Mandiant 2023)
Ukraine	Dec 2016	State	Outage via OT	1.25h blackout across northern Kyiv (INL 2020 ; Polityuk 2016)
Ukraine	Dec 2015	State	Outage via OT	Blackout for 225k customers for 3.5h (INL 2020)
Ghana	2022	Unknown	Incidental outage via IT	“5+ days of power outages” (Machtemes et al. 2025)
South Africa	2019	Unknown	Incidental outage via IT	250k customers lost power, delayed restoration (Machtemes et al. 2025)
Poland	2025	State	Process compromise	OT devices bricked; no outage (Midnight Blue 2026 ; MITRE 2025)
Ukraine	Apr 2022	State	Process compromise	Malware in OT network; foiled attack (ESET Research 2022 ; O’Neill 2022)
US	2014	Unknown	Process compromise	Utility’s control system software accessed (ICS-CERT 2014)
US	2013	State	Process compromise	Dam’s SCADA accessed during maintenance (DOJ 2016)
US	2012	Unknown	Process compromise	Malware delayed plant restart for 3 weeks (ICS-CERT 2012)
Brazil	2011	Unknown	Process compromise	Conficker worm downed power plant industrial control systems (ICS) (Branquinho 2011)
US	2003	Unknown	Process compromise	5h nuclear plant safety system loss by Slammer worm (Poulsen 2003)
UK	1999	Non-state	Process compromise	Guard trips access control/electronic door lockdown (Miller et al. 2021)
Denmark	2022	Unknown	Loss of view/control	Operator lost visibility into 3 remote assets (SektorCERT 2023)
Germany	2022	State	Loss of view/control	Loss of control/view of 5,800 wind turbines (Willuhn 2022)
US	2019	Unknown	Loss of view/control	Brief loss of view/control for remote assets (NERC 2019 ; Sobczak 2019)
US	2018	Unknown	Loss of view/control	Control center offline for 12–24h out of caution (Alrich 2020)
Germany	2022	Unknown	Loss of view/control	Cautionary 1–2 day loss of view of wind turbines (Deutsche Windtechnik 2022)
Montenegro	2022	Unknown	Loss of view/control	Switched to “manual handling” out of caution (Reuters 2022b)
Norway	2019	Non-state	Loss of view/control	Took “months” to restore full capability (Jeffries et al. 2022)

Ukraine	2017	State	Loss of view/control	Forced to “paper standards” for 10 days (Borys 2017)
US	2003	Unknown	Loss of view/control	Control center SCADA downed by Slammer worm (Owens 2009)
US	2003	Unknown	Loss of view/control	Slammer worm downed SCADA traffic via telecom (Poulsen 2003)
UK (BVI)	2024	Unknown	Operational impact	IT attack slowed hurricane blackout recovery (Machtemes et al. 2025)

Table 2.1 | Historical grid cyberattacks by impact. Note: See [Appendix 11](#) for detailed methodology.

Severity: Cyber Blackouts Have Caused Only Modest Impacts (~\$1m)

The documented instances in which cyberattacks have caused power outages resulted in modest impacts. The 2015 and 2016 blackouts in Ukraine were relatively brief and narrow, lasting a few hours and affecting several hundred thousand people. This report estimates that the 2015 blackout caused \$0.2m–2m in economic damages, and that the 2016 incident was around 10x less damaging – see [Appendix 2](#) for calculation details.¹⁴ Detailed information on the third blackout in October 2022 could not be found.

Blackout	Scope	Duration (~100% restored)	Est. damages (\$2025)
Dec 2015 ¹⁵	500k people	3.5 hours	\$0.2–2m
Dec 2016 ¹⁶	<500k people	1.3 hours	\$0.01–0.2m
Oct 2022 ¹⁷		Unknown	

Table 2.2 | Severity of cyber blackouts in Ukraine.

It is worth noting that while the 2016 attack ultimately caused less economic damage than the 2015 attack, it is regarded as having been more technically complex. The attack was significantly more automated, and the attackers appear to have been aiming for considerably more disruptive effects than were achieved.¹⁸ More detailed case studies of the Ukraine grid attacks can be found in [Section 4](#) and [Box 4.3](#).

¹⁴ The 2016 blackout was at least 3x smaller in magnitude. Notably it also took place at around midnight, resulting in meaningfully less economic impact per unit of outage.

¹⁵ Duration, scope and load impact via [INL \(2020, 1\)](#). This source reports the scope as 225,000 electricity customers, which is converted to people. Ukraine has ~2.3 people per electricity customer – 17.7m electricity customers ([IEC 2023, 12](#)) and population of 41.2m ([UA Gov n.d.](#)) – so this amounts to ~520,000 individuals affected.

¹⁶ Duration, scope and load impact via [INL \(2020, 2\)](#). Regarding scope, this source says “fewer individual customers experienced a power outage” [than the 2015 attack].

¹⁷ Blackout reported in [Mandiant \(2023\)](#), without detail on scope or duration.

¹⁸ In particular, it targeted high-voltage transmission networks rather than distribution-level infrastructure, and leveraged ICS-specific malware ([Dragos 2019](#)). [Slowik \(2019a\)](#) suggests the ultimate goal of the attack was to achieve physical damage to equipment.

This suggests economic damages from the worst cyber blackout to date have fallen more than five orders of magnitude short of the \$100 billion threshold. Notably, however, economic damages from equivalent blackouts in the US would plausibly be larger given the significantly higher levels of economic activity.¹⁹

The observed track record of disruptive cyberattacks is an imperfect window into the current capabilities of state-level actors. The track record is lagging, reflecting capabilities demonstrated several years ago rather than those of today. States generally prefer covert operations to avoid attribution, and disruptive attacks that become public may represent a biased sample, including operations that failed or in which attackers were principally seeking to demonstrate capabilities rather than achieve disruptive effects. Finally, while the Ukraine blackouts are central to this report's analysis, generalization to a US context is limited by Ukrainian operators' more extensive experience of manual operations and readiness following earlier attacks. These issues with interpreting the track record are discussed further in [Section 8.2](#).

2.2 Track Record of Major Accidental Blackouts

To empirically ground the potential damages from major grid failures caused by cyberattacks, the track record of major accidental blackouts was also reviewed. [Table 2.3](#) summarizes several of the most significant US blackouts in recent decades for which credible estimates of economic damages could be found.²⁰

¹⁹ US GDP per capita is ~15x higher than Ukraine's ([World Bank 2025](#)). On the other hand, the welfare effects of a unit of economic damages in the US would be lower than for Ukraine due to the diminishing marginal utility of money. It is unclear how these two effects net out.

²⁰ This table includes blackouts drawn from [National Academies \(2017, Appendix E\)](#) and [Reuters \(2018\)](#) for which estimates of economic damages from the outage itself are available. The 2021 Texas Power Crisis is also included, as a particularly salient recent example.

Event	Cause	Scope m people	Duration 90% restored	Damages \$2025
2003 Northeast Blackout ²¹	Accident Cascading failures	50	1.5 days	\$13bn
2012 Hurricane Sandy ²²	Accident Storm damage	18	8 days	\$23bn
2021 Texas Power Crisis ²³	Accident Extreme cold⇒supply crisis	11	3 days ²⁴	\$5bn

Table 2.3 | Major accidental blackouts in the US.

The historical record shows that even the most significant accidental blackouts in the US, varying widely in their scope and duration, have caused economic damages on the order of ~\$10 billion, an order of magnitude below this report’s threshold.

Note that for the blackouts caused by severe weather – Hurricane Sandy and the Texas Power Crisis following Storm Uri – the cited estimates are for the economic damages attributed to the power outages specifically. Natural disasters also cause significant damages via property destruction and broader impacts on economic activity. For example, headline figures of \$80–130 billion are sometimes quoted for damages from the 2021 Texas Power Crisis, but these include broader impacts of the extreme weather event.²⁵

Altogether, this suggests that for a grid cyberattack to reach catastrophic levels of economic damage, it would have to exceed the level of disruption from the worst accidental blackouts.

2.3 Baseline Risk of \$100 Billion Grid Cyberattacks

To complement the analysis of the historical track record, the survey of 8 domain experts and 13 superforecasters asked them to estimate the probabilities that, in 2026, at least one cyberattack on the US grid causes \$100 billion in economic damages, with a corresponding

²¹ **Scope** via [DOE \(2004, 9\)](#). **Duration** using NERC data via [ICF \(2003, 2\)](#), suggesting that 89% of load was restored after 32 hours (6,600 MW shortfall), and 97% after 42 hours (2,000 MW shortfall). **Damages:** [Elcon \(2004\)](#) reports separate estimates of \$4.5bn–8.2bn and \$7–10bn (\$2003). The table cites the mean of these bounds (\$7.4bn), inflation adjusted using the [US Inflation Calculator](#).

²² **Scope** of 8.5m customers via [DOE \(2013\)](#), which is converted to 18m people via the 2012 US ratio of population to electricity customers (314.7m/145.3m) ([EIA 2012, 11](#); [FRED n.d.](#)). **Duration** estimated from chart in [DOE \(2013, 11\)](#) – suggesting that 90% customers were restored 8 full days from the start of disruptions on 29 October (see [spreadsheet](#)). **Damages** from the outage have been estimated at \$16.3 billion ([Kunz et al. 2013](#)), inflation adjusted using [US Inflation Calculator](#).

²³ **Scope** via [Dallas Fed \(2021\)](#). **Damages:** [Dallas Fed \(2021\)](#) estimates outage damages at \$4.3bn, inflation adjusted using [US Inflation Calculator](#).

²⁴ **Duration** via [Dallas Fed \(2021\)](#). Note that this is the total duration of outages during 2–3 weeks of rolling blackouts, and is not strictly comparable to the other values. Rolling blackouts are discussed in [Section 3.1](#).

²⁵ Several respectable sources misattribute these combined estimates to the power outage specifically. For example, an official report by FERC-NERC claims “... the Federal Reserve Bank of Dallas estimated that the outages caused direct and indirect losses to the Texas economy of between \$80 to \$130 billion.” ([FERC et al. 2021, 11](#)); see also [Brookings \(2021, 9\)](#). A careful reading of the original source indicates this estimate includes damages from the severe weather event [Author’s emphasis]: “Early estimates indicate that the freeze and outage may cost the Texas economy \$80 billion–\$130 billion in direct and indirect economic loss.” [Dallas Fed \(2021\)](#).

question using a \$10bn threshold. For a \$100 billion attack, the median estimate among both experts and forecasters was 0.1% (interquartile range: 0.002–0.15% for experts and 0.05–0.12% for superforecasters). At the \$10bn damage threshold, the median estimate among both groups was 1% (IQR: 0.48–1.5% for experts and 0.5–2% for superforecasters). See [Figure 2.1](#).

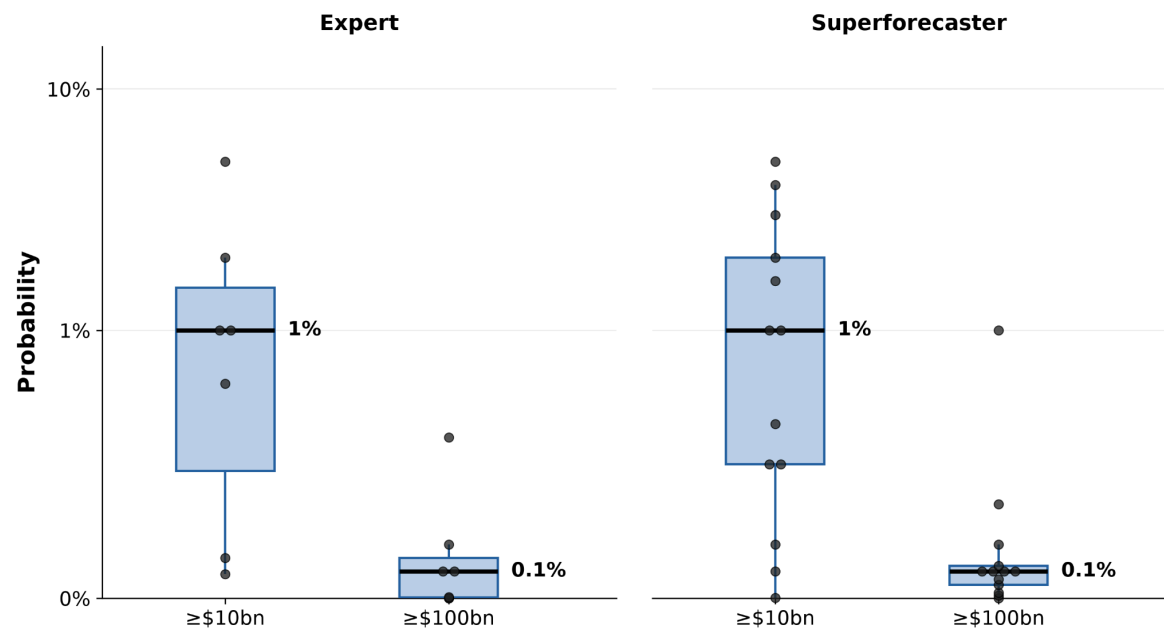


Figure 2.1 | Survey estimates of baseline risk for ≥\$100 billion and ≥\$10 billion grid cyberattacks.
Note: Estimated probability of at least one cyberattack against the US grid causing a blackout with at least \$X billion in economic damages in 2026. Boxes span the interquartile range of estimates within each group; whiskers extend to the furthest estimate within 1.5× the interquartile range; black bars indicate medians. Source: [Ceppas de Castro et al. \(2026\)](#)

3. What Would It Take for a Blackout to Cause \$100 Billion in Damages?

The historical track record suggests that a \$100 billion grid cyberattack would be unprecedented – requiring impacts orders of magnitude greater than prior grid cyberattacks and exceeding the worst accidental blackouts in the US.

To assess whether such unprecedented damage is plausible, this section works backwards from the \$100 billion damage threshold. This section addresses two key questions:

- What are the primary types of blackout, and how do their mechanisms determine their impacts?
- What combination of scope (people affected) and duration (length of outage) would be required to reach \$100 billion in economic damages?

Understanding these relationships makes it possible to estimate, in [Section 4](#), the effects that a cyberattack would have to achieve in order to cause such a blackout.

3.1 Types of Blackout

The mechanism by which a blackout occurs determines both how many people lose power and for how long.

There are three primary blackout types – cascading blackouts, blackouts caused by widespread physical damage, and rolling blackouts – each with characteristic patterns of scope and duration that will inform the analysis of attack pathways. (For a basic overview of how the grid works, see [Appendix 1](#).)

Cascading Blackouts

- **Examples:** 2003 Northeast Blackout; 2025 Iberian Blackout.²⁶
- **Mechanism:** One or more isolated faults (e.g. generators or transmission lines tripping offline) cause significant grid instability, which causes more generators and lines to

²⁶ In April 2025, a cascading blackout affected much of Spain and Portugal. The cascade was precipitated by several large generation sites disconnecting unexpectedly. 90% of power was restored within 14 hours (calculated in this [spreadsheet](#)). See [ICS Investigation Expert Panel \(2025\)](#).

disconnect, further destabilizing the grid. Absent timely intervention, this can rapidly cascade as more equipment is automatically disconnected from the grid.²⁷

- **Restoration:** Power can be restored relatively quickly, since there is typically limited damage to physical equipment ([National Academies 2017, 144](#)). After tripping offline, most generation capacity can be brought online within a few hours.²⁸ [Section 4.2](#) will consider the recovery dynamics following cascading blackouts in more detail.

Blackouts Due to Widespread Physical Damage

- **Examples:** 2012 Hurricane Sandy; 2017 Hurricane Maria.
- **Mechanism:** Extreme weather (e.g. hurricanes) causes widespread damage to grid equipment, particularly distribution infrastructure like utility poles and lines.²⁹ Previous hurricanes have damaged tens of thousands of utility poles, thousands of transformers, and taken hundreds of substations offline.³⁰
- **Restoration:** Typically much more time-consuming, since damage is spread across very large numbers of distribution-level assets (poles, lines, local transformers) which require extensive field crews to locate and repair – e.g. 70,000 workers were involved in restoring power after Hurricane Sandy.³¹

Rolling Blackouts

- **Examples:** 2021 Texas Power Crisis, where customers experienced ~70 hours of rotating outages over 2–3 weeks after extreme cold led to a substantial loss of generation amidst unexpectedly high demand.³²
- **Mechanism:** Electricity supply and demand need to be finely balanced to maintain grid stability. When the grid experiences a drastic and unexpected shortfall in supply (for example, due to unplanned generator outages) operators can use emergency load shedding to keep demand in line with supply and prevent uncontrolled (cascading) blackouts. This often involves rotating, localized outages.

²⁷ In advanced grid systems like the US, sophisticated protection schemes, redundant transmission paths, and better-trained operators limit propagation. Cascades can propagate across an entire interconnected grid, as seen in LMICs where cascading failures have caused nationwide blackouts – for example, Sri Lanka experienced a nationwide blackout in 2025 for 5–6 hours, after a monkey came into contact with a transformer. [EconomyNext \(2025\)](#); [BBC \(2025\)](#).

²⁸ More than half of capacity can be started from cold within 12 hours ([EIA 2020](#)). Thermal generators which have recently tripped and are still warm can be restarted considerably more quickly e.g. combined cycle gas turbines can restart within 1–4 hours from hot vs. 8–12 hours from cold ([O'Brien et al. 2022, 14](#)).

²⁹ One expert told us “the majority of damage from a natural event like a hurricane or an ice storm is destruction of distribution infrastructure.”

³⁰ Figures from Table 2 in [DOE \(2008, 12\)](#) summarizing impacts of 2005 and 2008 hurricanes.

³¹ See, e.g., [Chondrogiannis et al. \(2017\)](#) for discussion of the challenges in power restoration after natural hazards. Hurricane Sandy figure from [FEMA \(2020, 4\)](#).

³² [NERC \(2021, 7\)](#) notes that ~1,000 generating units experienced unplanned outages and derates due to the cold weather (representing 200 GW capacity).

- **Relevance:** In contrast to the above types, rolling blackouts are an intentional, controlled response to emergency conditions. They are relevant because they represent a potential outcome of an attack that damages a significant portion of generation or transmission capacity.

Empirical analyses confirm that blackouts caused by widespread physical damage are typically of much greater duration than cascading blackouts. An exhaustive study of US outages finds those caused by natural events last approximately 10 times longer than those from other causes.³³

In contrast, the blackouts with the largest scope are more commonly cascading blackouts. The three largest North American blackouts between 1984 and 2006 were all driven by cascading failures rather than physical damage ([Hines et al. 2009](#)).³⁴

These distinct blackout mechanisms have significant implications for cyberattack feasibility. Cascading blackouts can affect tens of millions through failures at just a few points but typically resolve quickly. Blackouts caused by physical damage last far longer, but producing them requires damage across vast amounts of widely distributed equipment. As described in [Section 4](#), causing \$100 billion in damages would require an attacker to navigate this trade-off between achievable scope and sustainable duration.

3.2 Relationship between Damages, Scope, and Duration

Having established how blackout mechanisms shape scope and duration, we can now consider how these variables together determine economic damages. The aim here is not to produce a precise estimate, but to bound the scale of disruption that would be required to reach catastrophic damage levels. To this end, economic damages can be approximated as a function of outage scope and duration, anchored on damage estimates for historical blackouts and standard outage-cost methodologies to obtain rough order-of-magnitude estimates. As discussed in [Section 1](#), this analysis focuses on direct economic damages incurred during the outage period itself – lost economic output, spoiled goods, emergency response costs – rather than cumulative GDP effects over subsequent months or years.

This analysis allows setting concrete targets for attack success: **\$100 billion in economic damages would require a blackout affecting approximately 100 million people for around a week** (or equivalent combinations of scope and duration, e.g. 50 million people for around 2 weeks).

³³ Calculated from Figure 3 in [Ankit et al. \(2022\)](#) – see [spreadsheet](#). The analysis of European outage data in [Appendix 3](#) finds a similar relationship.

³⁴ Table 1 of [Hines et al. \(2009\)](#) shows the three largest US blackouts, by customer count, have been caused by cascading failures. The largest blackout in history, which struck much of India in 2012, was caused by cascading failures ([Reuters 2012](#); [CERC 2012](#)).

The required scale of disruption is estimated using two approaches. First, by extrapolating from the economic damages caused by historical blackouts, expressed as damages per person-day of outage, as shown in [Table 3.1](#). Next, this estimate is cross-checked using estimates of the value of lost load, which measures the economic cost of unserved electricity.

Event	Damages \$2025	Scope m people	Duration ~90% restored	Damage per unit disruption = damages / [scope × duration] \$2025 per person-day
2003 Northeast Blackout	\$13bn	50	1.5 days	\$170
2012 Hurricane Sandy	\$23bn ³⁵	18	8 days	\$160
2021 Texas Power Crisis	\$5bn ³⁶	11	3 days	\$150
Average				\$160
Author's 90% confidence interval				\$100–250

Table 3.1 | Damages per unit disruption.

The rate of damage per unit disruption is between \$150 and \$170 per person-day across this small sample. This consistency is somewhat surprising given the events were caused by distinct blackout mechanisms.

This analysis suggests that a major blackout would cause damages per person-day of \$100 to \$250, with the wide range reflecting the major uncertainties and limitations of this extrapolation.

The most fundamental limitation is using a very small sample of three historical blackouts to estimate a per-person-day damage rate that is extrapolated to a blackout of unprecedented scope and duration. The absolute level of this damage rate is the dominant source of uncertainty in the estimate. The range of \$100–250 is skewed upwards relative to the sample mean of \$160 to reflect the expectation that, when extrapolating far beyond historical precedents, there are more plausible mechanisms for underestimation than overestimation.

Another limitation is that the estimate treats damages as scaling linearly with respect to disruption (scope × duration), which is a significant simplification. Economic losses may scale sublinearly if households and firms adapt (for example, with backup generation, load shifting, relocation of activity), or superlinearly if longer outages trigger broader supply chain disruption, business failure, or other macroeconomic effects. Recent economic modeling finds

³⁵ This estimate includes only damages from power outages (see footnotes to [Table 2.3](#) for sources)

³⁶ This estimate includes only damages from power outages (see footnotes to [Table 2.3](#) for sources)

that damages may scale sublinearly over multi-day horizons, with additional days of outage causing progressively smaller marginal losses ([Wing et al. 2025](#)).³⁷ However, evidence remains sparse for widespread long-duration blackouts. In any case, this is a less significant source of uncertainty than uncertainty over the absolute damage rate.

Several other caveats apply. First, damages will depend heavily on the characteristics of the affected area. For example, outages affecting energy production or data centers may cause more damages per person-day than those affecting only residential customers. Second, the reported damage estimates are themselves uncertain, their underlying methodologies have not been examined in depth, and differences in modeling assumptions may limit their comparability. Altogether, this report's wide confidence interval is intended to capture these limitations. The goal of this analysis is to establish rough bounds rather than a precise estimate.

The estimate of \$100 to \$250 damages per person-day implies that \$100 billion in economic damages would require between 0.4 and 1 billion person-days of outage – for example, a blackout spanning 100 million people and lasting 4 to 10 days.

To visualize what combinations of scope and duration could reach this threshold, [Figure 3.1](#) plots the three historical cases against scope alongside contour lines representing \$10 billion and \$100 billion. This helps to illustrate that reaching the \$100 billion threshold would require unprecedented combinations of scope and duration.

³⁷ See Table 4 p.8 (ComEd total). Interestingly, the direct costs of curtailment increase with duration (costs of 14-day outage are ~18x greater than 1 day), but this is outweighed by the opposing non-linearity in the ancillary costs. See [spreadsheet](#) for calculation.

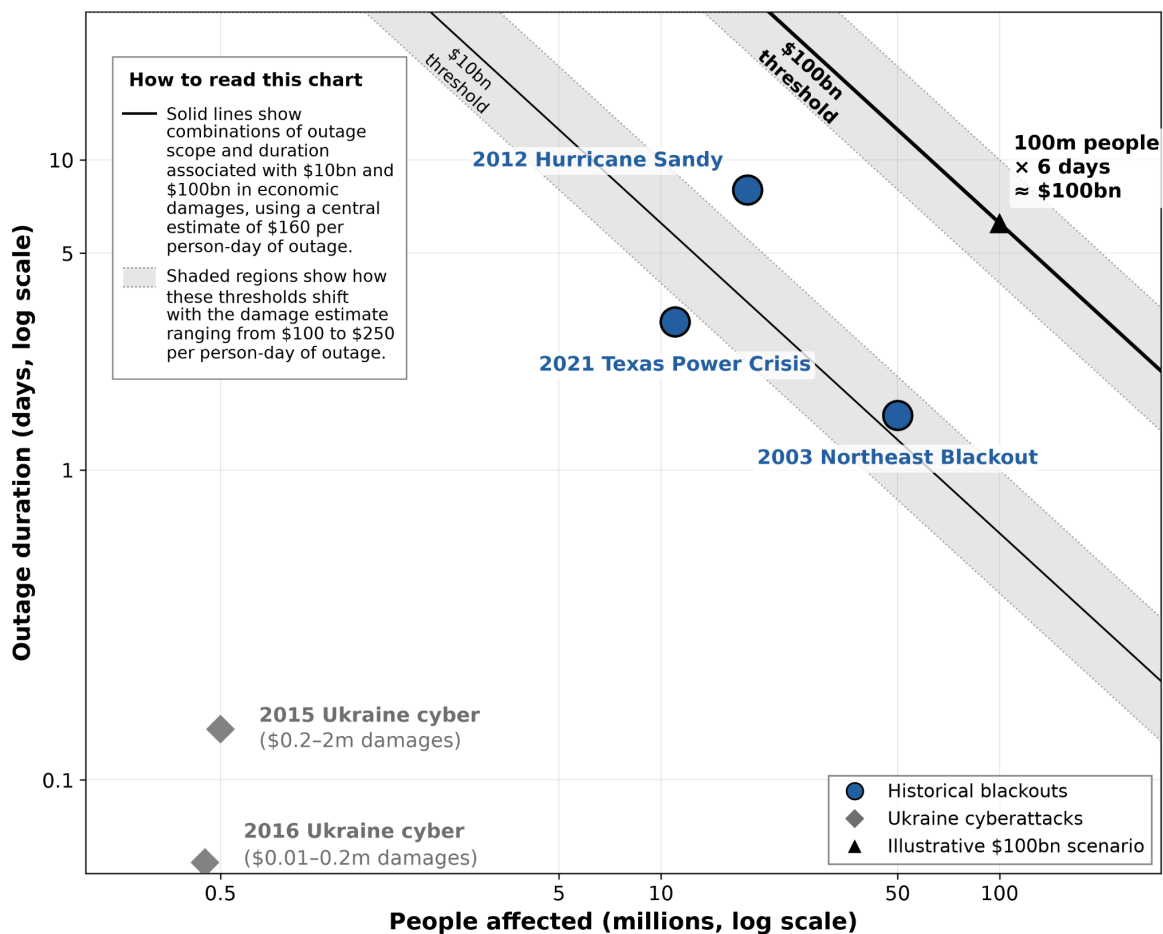


Figure 3.1 | \$100 billion in economic damages would require a blackout affecting 100 million people for around a week (or equivalent combinations of scope and duration). Historical blackouts and cyberattacks are shown for comparison.

This approach was cross-checked using a more conventional economic method based on the value of lost load (VOLL), a measure of the cost of a unit of power interruption, typically measured in \$/kWh. VOLL estimates are derived from surveys of willingness to pay to avoid further outages and econometric studies of actual outage costs.³⁸ Estimates vary widely by customer type, outage duration, timing, and economic context and should not be treated as a single universal constant. The VOLL estimate therefore serves primarily as a cross-check on orders of magnitude, rather than as a precise estimator of total economic loss.

This approach finds that \$100 billion in economic damages would require approximately 7 TWh of lost load, equivalent to an outage for 100 million people lasting between 3.5 and 9.5

³⁸ See [Schröder and Kuckshinrichs \(2015\)](#) and [Gorman \(2022\)](#) for a discussion of different methodological approaches and the range of extant estimates.

days. This VOLL-based estimate overlaps closely with the estimate from historical case studies, though the two approaches are not wholly independent, since the Texas damage figure is itself VOLL-derived. The calculation methodology is detailed in [Appendix 4](#) and relies on recent VOLL estimates for long-duration outages and model power restoration schedules based on historical blackouts.³⁹

Altogether, \$100 billion in damages would require a blackout affecting 100 million people for 4 to 10 days (henceforth, generally shorthand as “roughly a week”), or equivalent combinations of scope and duration. Having established the level of disruption required, we can now evaluate what a cyberattack would need to accomplish to cause such a blackout.

³⁹ The central VOLL estimate is drawn from [Gibbons and Sergici \(2024, 52\)](#), which estimates the system-wide (Texas) VOLL for 16 hour outages at \$13.6/kWh, with a 95% confidence interval of \$10.6–\$18.7.

4. How Could a Cyberattack Cause Blackouts of This Scale?

Having established the unprecedented scale of disruption required, we now turn to what effects a cyberattack would need to achieve to cause such a blackout.

A key challenge to any attack is that, in the absence of significant physical damage to grid equipment, power can typically be restored within hours or at most a few days after major blackouts. This creates an important bottleneck for attackers seeking to achieve an outage of roughly a week. Achieving a blackout this long would likely require either physically damaging equipment that cannot be quickly replaced or actively interfering with restoration efforts to delay recovery.

We will examine two pathways to catastrophic-scale grid disruption:

- a) **Physical damage:** Damaging critical grid equipment (generators or large transformers) that can take weeks or months to repair or replace.
- b) **Disrupting grid operations:** Precipitating a large cascading blackout and substantially delaying restoration through coordinated interference with recovery efforts.

The analysis focuses on these two pathways because they capture the fundamental mechanisms by which a cyberattack could produce a prolonged outage: widespread physical damage to critical equipment, or sustained disruption of grid operations and recovery. This does not exhaustively characterize the routes to achieving either of these outcomes. For example, demand-side attacks targeting distribution loads, hybrid cyber-kinetic operations, or combined attacks on the grid and other infrastructure networks such as telecommunications fall outside the scope of the analysis.⁴⁰ A distinct question is whether supply chain compromises could enable either attack pathway, overcoming some of the key challenges of achieving disruptive effects at scale. [Section 4.3](#) takes up this question.

These two pathways are also not mutually exclusive, since an attack could combine elements of each. In the 2016 Ukraine attack, attackers caused an outage and disabled protective equipment, apparently in the hope of causing physical damage during power restoration, though without success (see Boxes [4.3](#) and [4.4](#)).

⁴⁰ See [Lakshminarayana et al. \(2025\)](#), [Acharya et al. \(2020\)](#) and [Soltan et al. \(2018\)](#) on demand-side attacks. Russia's late 2022 attack on Ukraine's grid appears to have been a hybrid attack, coinciding with missile strikes ([Mandiant 2023](#)).

4.1 Physical Damage

The first pathway involves physically damaging critical grid equipment – specifically generators and large power transformers – whose replacement timelines of months or more could sustain outages long enough to cause catastrophic-scale economic damages.⁴¹ The analysis considers both:

- The feasibility of physically damaging the equipment via cyberattack;
- And the number of pieces of equipment that would need to be damaged to cause a catastrophic blackout.

Physical damage of grid equipment is distinct from “bricking” of the digital devices that are used to monitor and control grid equipment such as protective relays. Bricking refers to rendering devices inoperable by deleting or corrupting firmware or configuration files, such that devices must be manually restored or replaced. While bricking certain OT devices could disrupt grid operations – for example, compromising operators’ ability to view and control equipment remotely – it does not directly damage the underlying generation and transmission equipment.⁴² This would leave the grid’s physical capacity to deliver power largely intact, with recovery depending on operators’ ability to restore service under degraded conditions. Bricking therefore falls under the second pathway (disrupting grid operations). [Box 4.4](#) discusses the implications of bricking in grid cyberattacks.⁴³

4.1.1 Physical Damage to Generators

Feasibility: Damaging a Generator via Cyberattack Appears Possible

In 2007, researchers at the Idaho National Laboratory carried out a controlled demonstration in which a small generator was damaged via cyberattack, in what is known as the Aurora Test. This demonstrated that malicious manipulation of control systems can physically damage a generator, though experts disagree about the feasibility of reproducing the effect in a real-world cyberattack (See [Box 4.1](#)). Given the demonstration was almost two decades ago, at the dawn of modern cyberoffense research, it seems unlikely that there are no other

⁴¹ [EMP Commission \(2008. 64–66\)](#): “Replac[ing] irreparably damaged large transformers: One to two years plus production backlog plus transportation plus transportation backlog... Replac[ing] damaged furnace, boiler, turbine, or generator: one year plus production backlog plus transportation backlog.”

⁴² Bricking safety equipment, such as protective relays, could plausibly enable physical damage of equipment, as discussed in [Box 4.4](#).

⁴³ Following the cybersecurity literature, “physical damage” is used here to include effects such as overheating, exploding, or breaking, rather than wiping data or corrupting firmware. However the boundary is not clear-cut. “Hard bricking” – irreparable firmware-level corruption requiring a device to be replaced rather than restored (as with the serial-to-ethernet converters bricked in the 2015 Ukraine attack [Dragos 2017.10](#)) – is an interesting edge case, since it arguably approaches physical damage. This report treats hard bricking as disruption rather than damage, since causing physical damage to generation and transmission equipment requires materially different capabilities and attack chains than corrupting the firmware of digital monitoring and control devices. See [Box 4.4](#) for further discussion.

vulnerabilities that could in principle be exploited to damage generators.⁴⁴ Nonetheless, no real-world examples of generators being physically damaged via a cyberattack could be found.

Box 4.1. The Aurora Test

In 2007, the US Department of Homeland Security commissioned Idaho National Laboratory to conduct a live demonstration in which a 2.25 MW diesel generator was deliberately exposed to a rapid sequence of malicious breaker operations. The budget for the test was \$3.4–4.5m (\$2025).⁴⁵

The generator’s protection relay was misconfigured, enabling the generator to be repeatedly reconnected to the grid while its frequency and voltage had fallen out of sync, resulting in extreme mechanical stresses that destroyed the generator.

Expert opinion on the demonstration is mixed:

- [Zeller \(2011\)](#) points out that many obvious protection measures were not in place during the 2007 demonstration but says the attack “can easily target systems that have little or no security”.
- [Alrich \(2024\)](#) notes that “there has never been any recorded attack using the Aurora vulnerability. Indeed, the attack can only be executed by someone onsite, which makes it hard to conduct remotely from say China”.
- [Slowik \(2019c\)](#) says “direct manipulation of equipment to achieve an Aurora-like impact is either extremely difficult, or outright impossible” but suggests that there might be other more effective ways of causing damage via manipulation of protective relay behavior (see Boxes [4.3](#) and [4.4](#)).
- Another expert said they “would not rule it out” and that the question is “whether a malicious actor can gain remote electronic access to the relay”.

⁴⁴ Notably, the increasing share of renewables in the grid may present new opportunities for attackers seeking to damage generators. See, for example, [Freeman et al. \(2024\)](#), [Rieger et al. \(2022\)](#).

⁴⁵ Budget estimate \$2.9m (or \$2.2m without contingency). Inflation-adjusted: \$3.4–4.5m (\$2025) [Aurora FOIA \(2015, 57\)](#). See [spreadsheet](#) for cost breakdown – only ~20% of budget was for procurement of the generator and other equipment, with a further ~20% on installation and cleanup.

Scale: Many Generators Would Have to Be Damaged to Cause Significant Outages

Destroying a single generator would be extremely unlikely to result in any outages, let alone an outage of the scale envisioned in this report. This is because the US grid is designed to be very resilient to loss of generator capacity, in part because generators regularly come offline due to faults or maintenance.

US generation capacity is highly fragmented, with around 14,000 utility-scale power plants ([EIA 2024](#)). The number of damaged generators required to cause significant disruption can be estimated by considering an optimally targeted attack – that is, an attack directed at the largest generators. (This is an unrealistic scenario in many respects.)⁴⁶

An attack on 100 generators could damage at most 9% of US capacity.⁴⁷ Such an attack would be unlikely to result in any sustained outages, since US generation capacity exceeds expected peak demand by more than 20% – more than enough to absorb such a loss ([Thunder Said Energy 2024](#)). Damaging 100 generators at random would remove a negligible fraction of US generation capacity (0.4%).

Generators damaged	% US capacity damaged	
	Optimal targeting	Random targeting
1	~0.1%	<0.01%
10	~1%	0.04%
100	~9%	0.4%
1,000	~39%	4%

Table 4.1 | How many generators would an attack have to damage? Note: See [spreadsheet](#) for data and calculation.

A 2011 energy crisis in Cyprus offers an interesting case study in the consequences of losing a much larger fraction of generation capacity.⁴⁸ An explosion near the island’s main power plant destroyed 50–60% of the island’s generation capacity, equivalent to losing the 2,000 largest generators in the US.⁴⁹ Despite the plant taking two years to fully repair, the resulting outages were relatively modest: rolling blackouts of 2–4 hours per day over one month, affecting only

⁴⁶ For example, many of the largest generators are hydro or nuclear plants, which are not straightforwardly vulnerable to Aurora-style attacks. Our analysis here is partially inspired by [Lee et al. \(2018\)](#), which attempts to quantify the number of generators an attacker would have to target to cause a cascading blackout once factoring in targeting and success rates, as part of a critical analysis of Lloyd’s 2015 scenario.

⁴⁷ Calculated [here](#) using EIA data.

⁴⁸ Detail via [Zachariadis and Poullikkas \(2012\)](#).

⁴⁹ It was caused by the self-detonation of a large quantity of military munitions, which were being stored on the island having been seized in transit between Iran and Syria in 2009 ([BBC 2011](#)).

residential customers. Within a month, roughly half the capacity shortfall was addressed through imports and emergency generators, after which voluntary conservation measures were sufficient to avoid outages.

While the US grid operates on a vastly different scale, this case study demonstrates two important points. First, the relationship between capacity destruction and outage severity is not straightforward. Second, even extreme capacity loss need not result in catastrophic outages, as emergency measures and demand management can significantly mitigate impacts.⁵⁰

Box 4.2. Lloyd's grid cyberattack scenario

This report was motivated in part by Lloyd's 2015 analysis that a massive cyberattack on the US grid could cause from \$80 billion to \$300 billion in economic damages ([Lloyd's and CRS 2015b](#)).⁵¹ In their scenario, attackers damage 50–100 generators using an Aurora-style attack, triggering cascading failures affecting 93 million people.

Lloyd's assumptions relating outage duration to damages appear plausible. Their estimate of direct losses amounts to ~118% regional output loss per outage-day,⁵² which appears reasonable. However, their overall damage estimate is driven primarily by assuming extremely slow restoration times:

- In their two main scenarios it takes 1–2 days to restore 50% of power, and 14–21 days to restore 90% of power.
- They attribute the slow recovery to operator caution: utilities refusing to reconnect generators until understanding what caused the damage.
- However, this appears dramatically out of line with the empirical track record. After natural disasters (which cause far greater physical damage than envisioned in Lloyd's scenario), restoring 90% of power typically takes 2–3x as long as restoring 50% – rather than the >10x as long assumed in Lloyd's analysis.⁵³

The analysis provides no clear mechanism for why their cyberattack would produce restoration dynamics >5x slower than natural disasters, attributing this primarily to operator caution. Notably, the scenario assumes operators retain full control of their monitoring and control systems, whereas real-world grid attacks have typically involved bricking or otherwise compromising these systems – see Boxes [4.3](#) and [4.4](#). In practice, operator caution would be weighed against the scale of economic harms

⁵⁰ A complex attack would likely target backup and restoration capabilities, as demonstrated in the Ukraine grid attacks – see [Box 4.3](#).

⁵¹ Nominal damage estimates converted to \$2025.

⁵² They estimate the regional GDP of the affected area at \$5 trillion ([Lloyd's and CRS 2015a, 11](#)) (\$6.8tn in \$2025). This suggests a regional GDP of ~\$19bn/day (\$2025) – using figures for outage duration (in outage-days) from [Lloyd's and CRS \(2015a, 22\)](#). Their S1 scenario sees ~\$83bn damages over 3.8 outage-days – so \$damages per day of regional GDP are (~\$83bn/[3.8 × \$19bn]=)118%. Presumably this incorporates spillovers outside of the affected region and beyond the blackout duration

⁵³ Ratio of 90% restoration time to 50% restoration time: Hurricane Maria: 2.8x (187 days, 60 days) via [Reuters \(2018\)](#); Hurricane Wilma: 2.8x (11 days, 4 days) via [Reuters \(2018\)](#); Hurricane Irma: 2x (6 days, 3 days) via [Reuters \(2018\)](#); Hurricane Sandy: 3.2x (~8 days, ~2.5 days) estimated from chart in [DOE \(2013, 11\)](#)

associated with long-duration blackouts, with significant pressure from government and regulators to reconnect functioning equipment quickly.

With recovery assumptions aligned to historical patterns, their scenario would likely produce a substantially shorter blackout. (See [Section 4.2.2](#))

It remains uncertain how many generators would have to be destroyed to cause a \$100 billion outage in the US. Given the fragmentation of US capacity and significant reserve margins, catastrophic effects would require attacks on a massive scale. And the Cyprus case study illustrates that even extreme levels of capacity loss might be manageable without prolonged, nationwide blackouts. Altogether, an attack would likely have to damage 100 or more generators to achieve the required effects.

We now turn to attacks on large transformers, which might require fewer targets to cause significant outages.

4.1.2 Physical Damage to Large Transformers

Feasibility: It Is Unclear Whether Damaging a Transformer via Cyberattack Is Possible

Cyberattacks damaging transformers remain a more speculative threat. No demonstrations nor real-world examples of such attacks could be found, and experts disagree on whether they are possible in principle. Large transformers do sometimes fail catastrophically under certain conditions (for example, due to power surges or mechanical failures).⁵⁴ One source suggests attackers could cause overheating by manipulating digitally controlled tap changers: transformer components that regulate voltage.⁵⁵ One factor making deliberate damage difficult is the presence of non-digital protective systems, such as Buchholz relays, which cannot be hacked directly but may have to be disabled or compromised to achieve damage.⁵⁶ It is challenging to ascertain how pervasive these devices are today, with some sources suggesting they have mostly been replaced with digital systems.⁵⁷ Overall, while damaging a transformer via cyberattack would be unprecedented, the possibility cannot be ruled out.

Scale: Many Transformers Would Have to Be Damaged to Cause Significant Outages

The US has approximately 2,000 of the highest-voltage transformers installed.⁵⁸ The largest of these, used in the transmission grid, cost millions of dollars to replace, have unique designs

⁵⁴ [Bartley \(2003\)](#) on large transformer failures generally; [Hooley et al. \(2017\)](#) on transformer fires and explosions.

⁵⁵ [Baker et al. \(2021, 5\)](#); see also [Koelemij \(2020\)](#).

⁵⁶ One expert pointed out that mechanical devices could in theory still be compromised via cyberattack, e.g. if an attacker were able to interfere with the manufacturing process to introduce faults.

⁵⁷ E.g. [Lee et al. \(2018, 16\)](#): "In the past, protection from damage was provided by electromechanical devices not subject to cyberattack. Now, nearly all protective devices are small special-purpose computers that accept commands (some of them, remote commands) to control their operation."

⁵⁸ Transformers rated 345kV or higher or EHV (extra high-voltage) transformers. [DOE \(2014b, 28\)](#).

and very long lead times, and are mostly imported from overseas.⁵⁹ They are so large that special railcars are required to transport them.⁶⁰ Several credible sources suggest that replacing some transformers after damage could take months or years.⁶¹

It is therefore plausible that an attack severely damaging a significant number of critical transformers could cause long disruptions. As with generator damage, such an attack may result in a phased recovery rather than prolonged total blackouts. Initial emergency measures could likely restore partial service within days or weeks through rerouting power flows and temporary stopgaps (e.g. using smaller transformers). This may be followed by a longer period of demand reduction through rolling blackouts, prioritization of critical use cases, and voluntary conservation measures until full service could be resumed.

Estimates of the number of large transformers that would have to be destroyed to cause sustained outages were not found in the literature; nonetheless, indirectly relevant data points include:

- Contested FERC scenario:** A 2013 memo by the Federal Energy Regulatory Commission suggested that, were a physical attack to destroy 9 interconnection substations and a transformer manufacturer, the entire US grid could be offline for 18 months or more.⁶² Since each substation will have 3–4 individual transformers, this probably amounts to around 30 transformers, in addition to a manufacturing facility. The analysis was heavily criticized by the Department of Energy, which argued that it relied on “highly unlikely assumptions”, and that, even granting these, the loss of the substations “would not result in the consequence described in the analysis or any other consequence that could be reasonably expected to result in damage to national security” ([DOE 2015, 6](#)).
- Metcalf sniper attack:** A 2013 sniper attack on a major transmission substation in California resulted in damage to 17 transformers. Notably, this did not result in any outages, as power was rerouted after the substation was taken offline ([DOE 2016b, 19–20](#)).

⁵⁹ [GAO \(2023, 14\)](#) for lead times, [GAO \(2023, 10\)](#) for design specifications.

⁶⁰ [GAO \(2023, 17\)](#). Footnote 25 states that (as of April 2023) the US had only 10 of the custom railcars required to transport the largest transformers in service. The need for transportation can be partly mitigated by storing spare transformers at their intended location, though this provides limited protection against physical attacks ([DOE 2014a, 16](#)).

⁶¹ - GAO: “DOE estimates that LPTs could take years to replace, especially following certain severe low-probability, high-impact events” [GAO \(2023, 5\)](#).

- 2012 National Academies report said terrorist attacks targeting large transformers could result in restoration taking “months to years” ([National Academies 2012, 79](#)).

- 2008 EMP Commission report on recovery times for different pieces of infrastructure after EMP ([EMP Commission 2008, 64–66](#)) estimated that replacing irreparably damaged transformers would take at least one to two years.

⁶² The nonpublic memo was reported in [WSJ](#): “Destroy nine interconnection substations and a transformer manufacturer and the entire United States grid would be down for at least 18 months, probably longer.”

- **Geomagnetic storm literature:** A 2008 modeling study found that a severe geomagnetic storm could put over 300 extra high-voltage transformers at risk of permanent damage, which the authors estimated would result in prolonged outages for 130 million people.⁶³
- **Spare transformer stockpile:** US utilities maintain inventories of spare large transformers ([EEI 2023, 4](#); [GAO 2023, 23](#)). Public figures on the overall stockpile of spare transformers could not be found. As of 2007, one major operator held spares equivalent to ~15% of its large transformers⁶⁴ – if representative of overall stockpiles today, this would suggest roughly 300 spares in total.

Using these sparse data points, it is possible to set some rough bounds. Damaging hundreds of large transformers, and therefore overwhelming the stockpile of spares, appears more than sufficient to cause a \$100 billion outage. It is difficult to determine a precise lower bound. Note, however, that the Metcalf incident, in which 17 transformers were damaged, did not result in any outage. Altogether, it seems reasonable to estimate that an attack would have to damage at least dozens of critical transformers to achieve a \$100 billion blackout.

4.1.3 Key Challenges of Either Physical Damage Pathway

Whether targeting generators or transformers, causing \$100 billion in damages would require damaging dozens to hundreds of critical grid assets. This introduces a number of key challenges for any such attack.

The first major challenge is scaling the attack across dozens or hundreds of targets. Given the decentralized nature of the US grid, these targets are likely to be distributed among many different companies, requiring relatively individualized attacks. Each intrusion creates an additional opportunity for detection. Maintaining covert access across so many targets for long periods would therefore require extreme stealth.⁶⁵ As discussed in [Section 4.3](#), OT environments are heterogeneous. Differences in equipment, configurations, and network architecture mean that attack capabilities often cannot be directly reused across targets.

The second challenge is the need for simultaneity. In the event of a massive attack on grid equipment, operators and automatic controls would react quickly to protect their assets.

⁶³ Kappenman via [National Academies \(2008, 78–79\)](#) modeled a disturbance with the intensity of the 1921 geomagnetic storm ([Silverman and Cliver 2001](#)). This modeling is focused primarily on estimating the numbers of transformers damaged, with very little information given on how they estimate the associated outages. See [Roodman \(2015\)](#) for a comprehensive and critical evaluation of Kappenman's modeling of geomagnetic storm impacts.

⁶⁴ [Glover et al. \(n.d., 2\)](#) cited in [CRS \(2014, 12\)](#).

⁶⁵ Operations aimed at physical effects will create detection opportunities distinct from IT attacks. Reconnaissance of control systems, including testing and experimentation, can generate observable activity that defenders may notice. For example, the Triton attackers' interaction with safety instrumented systems at a petrochemical facility inadvertently triggered a plant shutdown, leading to the discovery of the intrusion ([Slowik 2019b, 5–6](#)).

Consequently, an attacker seeking to damage dozens or hundreds of targets would likely have to do so within a narrow window, posing major additional technical challenges:

- Controlled physical effects.** Synchronous damage requires significant control over the timing of damage. Causing a predictable physical effect via cyberattack seems dramatically more challenging than causing any physical effect at all, such as by causing a piece of equipment to fail randomly at an undetermined time or to wear down at an accelerated rate.⁶⁶
- Narrow window for coordinated damage.** Damaging many generators or transformers would likely cause grid instability and trigger protective systems to start disconnecting equipment from the grid. Since routes to physical damage typically rely on equipment being connected to the grid, disconnections would reduce the number of targets available to the attacker. Moreover, the window of opportunity could be extremely short. The cascade sequence in the 2003 Northeast Blackout, for example, lasted only 90 seconds, and grid protection systems operate with millisecond latency.⁶⁷ As an enabling step, an attacker may have to hack protective systems ahead of time to modify their behavior. Doing so at large scales would add significant additional complexity and effort.⁶⁸
- Testing limitations.** Testing coordinated damaging attacks on multiple pieces of equipment would be very costly. The Aurora test cost several million dollars and involved a single small diesel generator.⁶⁹ Given the cost and availability of large transformers, even well-resourced actors may be reluctant to use them for testing. An attacker may therefore have to conduct an attack without realistic testing.
- Operational challenges.** Coordination across dozens of simultaneous attacks would present significant operational challenges, particularly given the narrow time window for impacts and inherent communication latencies. This could in principle be overcome via automation, though this would likely trade off against reliability. For example, Russia's April 2022 grid attack was scheduled to activate at a predetermined time, but this created a window allowing defenders to intervene.⁷⁰

⁶⁶ E.g. Figure 3 in [Assante and Lee \(2015, 11\)](#) labels "high confidence process and/or equipment effect" as "extremely difficult". [Conway et al. \(2020, 8\)](#): "it's not hard to deny service to a PLC but it is incredibly difficult to cause a specific directed physical event or denial of safety especially if you want to do it in an engineered and repeatable manner."

⁶⁷ [NERC \(2004, 73\)](#): "Between 16:10:36 and 16:13, a period of less than a minute and a half, a chain reaction of thousands of events occurred on the grid, driven by physics and automatic equipment operations. When it was over, much of the Northeast was in the dark."

⁶⁸ [Slowik \(2019c, 19\)](#) discusses something similar: "a much more effective (if difficult) attack vector lies in modifying breaker logic or functionality to create subtle changes in behavior that weakens protection ... such impacts would include modifying tolerances for automated responses ..."

⁶⁹ \$3.4–4.5m (\$2025). See footnotes to [Box 4.1](#) for additional detail and sources.

⁷⁰ [ESET \(2022\)](#): "The destructive actions were scheduled for 2022-04-08 but artifacts suggest that the attack had been planned for at least two weeks."

Altogether, achieving \$100 billion in damage by physically damaging grid equipment would be exceptionally difficult. Damaging dozens or hundreds of pieces of grid equipment is not only significantly harder than damaging one, but represents a major jump in complexity due to the challenges of scale and coordination.

A natural question is whether supply chain compromises, which have historically compromised hundreds or thousands of targets in single campaigns, could overcome these challenges. [Section 4.3](#) examines this in detail and concludes that while supply chain compromises could substantially assist with gaining access at scale, they face fundamental limitations in enabling the physical or operational effects required for catastrophic damage.

4.2 Disrupting Grid Operations

The challenges of scale, coordination, and achieving unprecedented cyber effects raise the question: Could attackers bypass the need for physical destruction entirely? Next, we turn to whether and how attacks aimed at disrupting grid operations, without damaging equipment, could achieve catastrophic effects.

This pathway has a critical two-step requirement. First, an attacker must trigger a large enough outage, most likely by precipitating a cascading blackout affecting 100 million people. But as noted earlier, cascading blackouts without physical damage typically resolve within hours to a few days. This means attackers must then interfere with restoration efforts to sustain the outage for roughly a week. As we will see, while the first step would be very difficult, the second would be dramatically harder.

Russia's cyberattacks against Ukraine's power grid are the primary empirical basis for much of the analysis that follows. [Box 4.3](#) provides further technical details on each attack, which will be referred to in subsequent sections.

Box 4.3. Ukraine grid cyberattacks

2015 (BlackEnergy/KillDisk): Attackers compromised three Ukrainian regional electricity companies via spear-phishing, then moved laterally through IT networks to reach OT systems. In December 2015, 19 months after the initial intrusion attempts, attackers used legitimate remote access tools to manually open breakers at 53 distribution substations, causing an outage for 225k customers. In addition, attackers deployed wiper malware against control center workstations, bricked serial-to-ethernet converters at substations (see [Box 4.4](#)), and launched a DoS attack on the utility's call center to prevent customers from reporting outages. Power was restored manually within around 3 hours, though SCADA systems were impaired for several months.⁷¹

2016 (CrashOverride/Industroyer): In December 2016, attackers deployed ICS-specific malware at a transmission substation in Kyiv, causing a blackout for around 1 hour. The multi-stage attack was

⁷¹ Attack details from [E-ISAC \(2016\)](#). SCADA remediation timeline from [Slowik \(2019c, 10\)](#).

designed to (a) manipulate breakers via protocol commands to cause an outage; (b) wipe SCADA systems and ICS configuration files; and (c) disable protective relays via denial-of-service.⁷² The attack aimed to exploit operators' willingness to manually restore power – observed in 2015 – by encouraging manual reconnection in the absence of protective relay safeguards, creating conditions for physical equipment damage ([Slowik 2019a](#)). Ultimately, the attack caused less disruption than 2015, and the protective relay component appears to have failed due to a software error. Rapid restoration was aided by operators' familiarity with manual procedures and heightened readiness following the 2015 attack.⁷³

2022 (Industroyer2): In early 2022, during Russia's full-scale invasion of Ukraine, a new variant of Industroyer was deployed against grid infrastructure and scheduled to execute in April. The attack was detected and blocked by CERT-UA and ESET before the OT payload could execute. This again demonstrates the role of defender readiness, though the wartime context – with heightened alert and international support – limits how far this can be generalized to peacetime conditions.⁷⁴

Some caution is warranted in generalizing from the cyberattacks in Ukraine to similar events in the United States. Ukrainian grid operators had significant experience with manual operations and operated at heightened readiness amid the conflict with Russia – particularly following the 2015 attack. These factors plausibly reduced restoration times relative to what might be expected on a more automated grid operating in peacetime.

4.2.1 Precipitating a Massive Blackout

In contrast to the physical damage pathway, there are real-world precedents of attackers causing outages via disrupting grid operations. In the 2015 Ukraine attack, attackers hijacked legitimate operator controls to disconnect 53 distribution substations, in parallel with efforts to delay recovery. This resulted in an outage for ~0.5m people lasting around 3 hours ([Zetter 2016](#); [INL 2020, 5](#)).

To cause catastrophic damages, the natural route for an attacker would be to trigger a cascading blackout (see [Section 3.1](#)). The interconnectedness of the grid means that a handful of local disturbances can, in certain circumstances, cascade into a system-wide outage. For example, in the 2003 Northeast Blackout, events in a few control centers, in combination with system-wide stress, led to an outage spanning 50 million people.⁷⁵ In the 2025 Iberian Blackout, a major fault at a single generation site initiated a rapid sequence of dozens of

⁷² Attack details from [ESET \(2017\)](#), [Dragos \(2017\)](#), and [Slowik \(2019c\)](#).

⁷³ Ukraine's head of cyberdefense: "Since 2014 we've been under constant aggression, and our expertise is unique in how to rebuff this aggression. We're stronger. We're more prepared." – quoted in [CRS \(2024\)](#), which discusses Ukraine's preparedness in some depth.

⁷⁴ Attack details from [ESET \(2022\)](#); [CERT-UA \(2022\)](#).

⁷⁵ Four entities were involved in the initiating events (Two utilities: First Energy and American Electric Power; Two reliability organizations: MISO and PJM) ([DOE 2004, 20](#)). The blackout was triggered in Ohio, and spread to 8 states and Canada ([NERC 2004, 63](#)).

generation disconnections, precipitating a country-wide blackout.⁷⁶ This suggests that, in the right conditions, an attacker could precipitate an extremely large blackout via compromising a small number of targets.⁷⁷

Attackers seeking to trigger such cascading failures would need to disrupt grid operations in ways that destabilize the system. Potential approaches might include some or all of the following:

- **Disconnecting transmission lines or generators.** Carefully timed disconnections of sufficient supply during high-load conditions could cause cascading grid instability.
- **Manipulating protective relay settings.** Protection relays automatically disconnect equipment in abnormal conditions to prevent damage and cascading failures. Digital relays could, in principle, be manipulated to cause unnecessary disconnections or to prevent stabilizing responses during a cascade.⁷⁸
- **Manipulating automatic generation control systems.** Regional grid operators use automated control systems to adjust generation output to balance supply and demand across their control area. Feeding false data to these systems could create artificial imbalances that trigger protective disconnections of loads or generators.⁷⁹
- **Manipulating grid operators.** Attackers could manipulate grid operators themselves into unwittingly taking destabilizing actions, rather than – or in addition to – directly manipulating grid control systems. For example, Russia’s 2016 attack on Ukraine’s grid involved using tailored malware to open breakers at a transmission substation, while also disabling protective relays in an apparent effort to cause grid instability and physical damage to equipment when operators manually restored power.⁸⁰

Attackers seeking to precipitate a cascading blackout would face several key challenges, including:

- **External conditions for success.** Fairly specific grid conditions are required for isolated faults to cause cascading failures. These conditions include high load levels,

⁷⁶ The initiating major fault was the trip of a 355 MW generation transformer in Granada at 12:32:57, which triggered dozens of further generation disconnections across multiple regions, totaling 2.5 GW by 12:33:18, with the Iberian grid disconnecting from the rest of Europe at 12:33:19 ([ICS Investigation Expert Panel 2025, 10–12](#)).

⁷⁷ It is worth noting that the US is divided into three largely independent grids, though the largest of these – the Eastern Interconnection – serves over 200 million people in the US ([DOE 2016a](#)). This plausibly represents the maximum scope of a single cascading blackout in the continental US.

⁷⁸ Protection relays typically monitor individual substations, transmission lines, generators, and transformers. Discussed in [Mandiant \(2025\)](#): “A coordinated attack against multiple critical relays can lead to a cascading failure across the grid, potentially causing a large-scale blackout.”

⁷⁹ AGC operates at the level of control areas – typically spanning an entire region or balancing authority – and adjusts multiple generators based on frequency measurements and power flows. See [Tan et al. \(2015\)](#).

⁸⁰ Discussed at length in [Slowik \(2019a\)](#).

low reserve margins, and/or frequency instability. Attackers may need to wait for, or create, these vulnerable conditions to maximize their chances of success.

- **Access to simulation tools and models.** Without access to power models and simulation tools similar to those used by grid operators, attackers may struggle to identify the specific disturbances sufficient to trigger widespread cascading failures. To be sure, attackers may be able to exfiltrate analysis and modeling tools from compromised utilities or use open-source versions.
- **Precise timing.** Disturbances would have to be fairly synchronized across multiple sites to destabilize the grid sufficiently, which would pose similar obstacles to those considered in [Section 4.1](#).

Triggering a cascading blackout via cyberattack would be extremely challenging, though possible in principle by hitting a set of well-chosen targets. As with physically damaging generators and transformers, there appear to be no real-world cases of cyberattacks causing cascading blackouts.

4.2.2 Sustaining a Massive Blackout

Even if attackers successfully triggered a cascading blackout affecting 100 million people – an unprecedented achievement – this would likely fall short of \$100 billion in damages. The track record shows that such blackouts resolve within hours or, at most, a few days. A week-long outage of this size would require sustained interference with dozens of parallel recovery efforts to substantially delay restoration.

The core difficulty is an asymmetry between how cascading blackouts are triggered and how they are resolved. While a handful of well-targeted disruptions can, under the right conditions, cascade into a system-wide outage, power restoration in the US is a decentralized and dynamic process, involving the coordinated efforts of dozens of relatively independent operators. Sustaining a blackout therefore requires thwarting these distributed efforts over several days, which is a far more demanding operation than the already difficult task of triggering the blackout in the first place.

Prior grid cyberattacks have attempted to delay recovery through the use of destructive malware on systems and devices used by operators to monitor and control grid equipment. For example, in the 2015 Ukraine incident, attackers deployed wiper malware and bricked communications equipment (see Boxes [4.3](#) and [4.4](#)). Grid operators were unable to restore power from control centers and sent field crews to manually reconnect the affected

substations. Operators were able to restore power within a few hours, though full remediation of SCADA systems took significantly longer.⁸¹

Box 4.4. Bricking in grid cyberattacks

What is bricking? Bricking is rendering digital devices inoperable by destroying their software, firmware, or configuration data. A bricked device typically cannot be restored remotely and must be manually reimaged, reconfigured, or physically replaced. A “soft brick” is recoverable with physical access to the device (e.g. factory reset via a local engineering interface); a “hard brick” requires returning the device to the manufacturer or using invasive recovery procedures ([Midnight Blue 2026](#)). Bricking is typically accomplished through wiper malware or malicious firmware, both of which are technically simpler than OT-specific payloads, since they operate at the level of file systems and firmware without requiring detailed knowledge of industrial control protocols.

Bricking has featured in several cyberattacks against the grid:

- **2015 Ukraine:** Attackers hard-bricked serial-to-ethernet converters at substations. Power was restored within hours through manual intervention, but full SCADA remediation took months. Operators could not communicate remotely with field equipment until the converters were physically replaced.
- **2016 Ukraine:** The attack included a denial-of-service on Siemens SIPROTEC protective relays by placing them into firmware update mode. The devices were not bricked, since the denial-of-service was recoverable by power cycling.
- **2025 Poland:** Attackers hard-bricked Hitachi remote terminal units (RTUs) via malicious firmware and soft-bricked Hitachi protection relays, disabling their protection and control functionality in a manner not recoverable remotely or via the device’s local interface ([Midnight Blue 2026](#)). The UK and EU attribute the attack to Russia’s FSB Centre 16 ([ECDO 2026](#)).

Delaying recovery: Bricked embedded OT devices cannot be reimaged remotely the way IT systems can. Importantly, stockpiles of spare RTUs, protection relays, or serial converters are limited – online distributors typically hold fewer than a dozen units of a given model, with lead times of a month or more for larger orders ([Midnight Blue 2026](#)). If an attacker bricked hundreds of devices simultaneously across many sites, supply constraints could significantly slow recovery. A prolonged loss of communications equipment can potentially be managed by reverting to manual controls, though this might be challenging at scale.

Unsafe conditions: Beyond slowing recovery, bricking protection equipment can create actively unsafe conditions. As described in [Box 4.3](#), the 2016 attack aimed to induce manual reconnection in the absence of relay protections. The soft-bricking of protective relays in 2025 suggests a more disruptive version of this attack: Unlike a recoverable denial-of-service, bricked protection devices must be physically replaced before normal safeguards can be restored, forcing operators to choose between restoring power without functioning protection or delaying restoration until devices can be replaced or substituted. The 2025 attack exploited default credentials and outdated firmware, suggesting that basic security hygiene would make such an attack significantly more challenging; it also targeted distributed generation sites rather than transmission infrastructure.

⁸¹ [Slowik \(2019c\) p.10](#): “in 2015, Ukrainian operators quickly moved to restore service through manual intervention, even while SCADA equipment was essentially disabled for months due to the wiper deployed after interrupting distribution.” [Slowik \(2022b\)](#) cites anecdotal evidence that “the system wiping and effective destruction of serial-to-ethernet converters produced damage taking years to effectively correct.”

Beyond direct technical interference, attackers might seek to extend outages by exploiting operator behavior. [Table 2.1](#) lists prior grid cyberattacks and includes several instances of operational impacts from precautionary measures taken by grid operators. In the 2021 Colonial Pipeline incident, pipeline operations were halted for several days as a precautionary response to a ransomware attack on the company's business IT networks ([Blount 2021](#)). Similarly, an attacker who compromises grid control systems might induce operators to preemptively disconnect equipment or delay restoration while they verify safety instrumentation and control systems.

However, the pressure to restore power during a major blackout would likely compress this timeline significantly. The economic impacts of the Colonial Pipeline shutdown were mitigated by reserve buffers and alternative supply routes – resulting in a modest 4¢/gallon increase in gasoline prices for several days.⁸² In contrast, a blackout affecting 100 million people would create immediate, much larger, and more visible harm. Economic damages would be on the order of \$10 billion per day, and there would be intense pressure on grid operators and cyber defenders to restore service.⁸³ So while operator caution might plausibly extend restoration by hours, or in extreme cases a day or two, it appears very unlikely that this mechanism could produce week-long blackouts with such wide scope.

As noted above, the decentralized nature of grid restoration poses significant challenges to sustaining a blackout. US restoration plans dictate that transmission owners and operators are responsible for bringing up their own systems independently, restoring load in stable “islands” before synchronizing with neighboring areas, and, eventually, the wider grid (See [Figure 4.1](#)).⁸⁴ Altogether, the US has around 370 individual transmission owners and operators. However, restoration efforts are coordinated through a much smaller number of reliability coordinators, and there is significant overlap in control systems, communications infrastructure, and operator training, which plausibly introduce common points of failure. Still, restoring power after a blackout affecting ~100 million people would likely involve dozens of largely independent restoration efforts.⁸⁵

Importantly, these restoration efforts may depend significantly on telecommunications infrastructure for coordination over long distances between control centers, field crews, and other restoration efforts. A combined attack on grids and telecommunications systems, analysis of which is beyond the scope of this report, could significantly hamper restoration.

⁸² See [Tsvetanov and Slaria \(2021\)](#).

⁸³ Per our estimate in [Section 3](#), a blackout affecting 100m would result in damages of approximately \$10–25 billion per day.

⁸⁴ [Electric Power Research Institute \(2019, 127–128\)](#): “each Transmission Operator is responsible for restoring its own system. Aid from external systems, in general, is to be minimized or avoided altogether.” See also [PJM \(2023\)](#).

⁸⁵ As of September 18, 2025, NERC reports 373 individual entities that are either transmission owners or operators – see [spreadsheet](#), using data via [NERC \(2025\)](#).

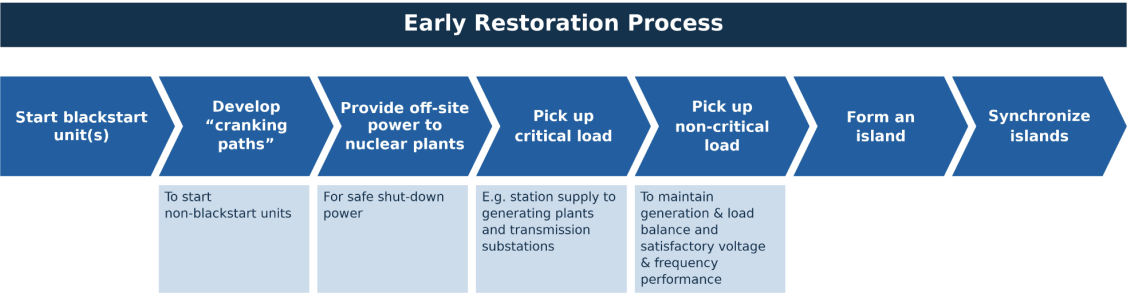


Figure 4.1 | Typical grid restoration process. Source: Adapted from [Electric Power Research Institute \(2019, 127\)](#)

Delaying restoration for a week would therefore require replicating intensive, sustained interference across dozens of independent targets simultaneously, with attackers maintaining active operations over days rather than the minutes-to-hours window needed to trigger the initial blackout. Each additional target would compound the risk of detection, the difficulty of maintaining access while defenders respond, and the challenge of coordinating an attack. Sustaining a massive blackout by disrupting grid operations appears extraordinarily challenging. [Table 4.2](#) summarizes the gap between triggering a cascading blackout and sustaining one long enough to cause catastrophic damages.

Scenario	Outcome				Attack Requirements	
	Precedents	Scope	Duration	Damages	Targets	Active duration
Cascading blackout	2003 Northeast; 2025 Iberian	≤100m	≤1–2 days	≤\$10bn	Handful	Minutes to hours
Cascading blackout with delayed recovery	No precedent	≤100m	Days or more	\$10–100bn	Dozens or more	Multi-day interference

Table 4.2 | Comparison of attack pathways via disrupting grid operations. Note: Active attack duration refers to the time required to execute the final attack sequence (triggering a cascading blackout, or delaying restoration for a week).

4.3 Supply Chain Compromises

To cause a catastrophic blackout via either physically damaging equipment or disrupting grid operations, attackers must achieve disruptive effects across many distinct targets. Because the number of targets represents a key difficulty of both pathways, it is worth considering

whether supply chain compromises, which involve gaining access to multiple targets via a shared upstream dependency, could help overcome this challenge. While supply chain compromises could meaningfully assist the access phase of a grid attack, they face fundamental limitations in enabling the effects required to achieve catastrophic blackouts.

Notable examples of supply chain attacks include:

- **SolarWinds:** 18,000 organizations received a compromised software update.⁸⁶
- **NotPetya:** Initially spread via a malicious update to Ukrainian accounting software (and subsequently spread rapidly from initially infected devices to other devices on shared networks).⁸⁷
- **Dragonfly campaign:** Compromise of three European ICS software vendors to gain access to energy sector targets.⁸⁸

The ICS Cyber Kill Chain, a framework to describe the attacks against industrial systems, distinguishes between gaining access to and moving within a network (Stage 1) and developing and executing an attack that achieves a physical or operational effect (Stage 2) ([Assante and Lee 2015](#)). Supply chain compromises could substantially assist with Stage 1 – for example, a compromised software update could provide an initial foothold across multiple target networks and potentially bypass some defenses.

However, for a cyberattack to achieve sustained operational or physical disruption, Stage 1 is insufficient ([Slowik 2022a](#)). Having gained persistent access to IT networks, an attacker must understand the specific OT environment, develop and test a capability tailored to the specific control systems and equipment present, and deliver it to achieve a precise physical or operational effect ([Assante and Lee 2015](#)).

The most plausible supply chain vector for grid attacks is software – specifically, compromised updates to IT or OT management software used across multiple utilities. This has been demonstrated in the Dragonfly campaign, which targeted European vendors of ICS-specific software ([Slowik 2021, 5–6](#)). Such a compromise could, in favorable cases, provide initial access to networks at multiple utilities, simplifying the access problem from dozens of independent intrusions to a single upstream compromise. However, this shortcut should not be overstated. OT software updates may require testing, staged deployment, and planned downtime, meaning that a compromised update channel may not translate into rapid or simultaneous access across many operational environments ([NIST 2023, 29](#)).

⁸⁶ See [Slowik \(2020\)](#); [NY DFS \(2021\)](#).

⁸⁷ See [Halstead and Righetti \(2026\)](#).

⁸⁸ See [Slowik \(2021, 5–6\)](#).

Several factors limit how much supply chain access can reduce the difficulty of the second stage of such operations:

OT environment heterogeneity. Even if a compromise delivers simultaneous access to dozens of utility networks, the attacker cannot deploy a single payload across all of them. OT environments are highly heterogeneous: Utilities use different vendors, equipment configurations, network architectures, and SCADA/EMS platforms.⁸⁹ While a supply chain compromise can reduce the access problem, the challenge of exploitation will still scale with the number of targets.

Targeting trade-off. Supply chain attacks face an inherent trade-off between reach and targeting. Broad compromises (e.g. major software vendors) can reach many targets but are unlikely to provide OT-specific capabilities ([Slowik 2022a, 11–12](#)). On the other hand, narrow compromises, such as Dragonfly’s attack on niche ICS vendors, offer more operationally relevant access but many fewer targets ([Slowik 2021, 5–6](#)). In order to provide a meaningful shortcut to catastrophic grid cyberattacks, a supply chain compromise would need to achieve both broad reach to dozens or more targets and OT-specific access to enable operational effects.

Capability execution. For a supply chain-enabled attack to achieve simultaneous effects across many targets, the deployed capability must either operate autonomously or rely on command-and-control (C2) communication with attackers. Both options create important challenges for OT cyberattacks.

- *Autonomous execution* might involve, for example, pre-programmed logic that identifies the local environment, develops an appropriate attack, and executes it without further instruction. There are numerous examples of crude and indiscriminate capabilities being delivered autonomously (e.g. NotPetya and WannaCry ransomware). However, these appear ill-suited to achieving complex disruptive effects on OT. In terms of targeted capabilities, [Slowik \(2022a\)](#) points out that Stuxnet is the only clear precedent for a tailored ICS capability being executed autonomously ([Slowik 2022a, 9](#)). Importantly, however, this targeted a single, well-characterized target, and nonetheless required an unprecedented investment of time and effort (discussed in [Section 5](#)).
- *Command and control.* Maintaining C2 communication with each target would necessarily involve network traffic between the attacker and victim networks, which

⁸⁹ See [Slowik \(2019b\)](#).

would reintroduce significant detection risk in proportion to the number of targets, removing a key advantage of supply chain compromises for access.⁹⁰

The empirical track record appears consistent with this picture. There appear to be no examples of supply chain compromises being used to achieve physical damage or sustained operational disruption in OT environments. The most relevant precedent, the Dragonfly campaign, appears to have been focused principally on espionage and reconnaissance.⁹¹ Moreover, Russia appears not to have used supply chain compromises in any of its targeted grid attacks against Ukraine aimed at causing outages, despite having demonstrated relevant capabilities in SolarWinds and NotPetya.⁹²

Altogether, supply chain compromises could substantially reduce the difficulty of gaining initial access at scale. However, they do not appear to address the core difficulty of developing and executing tailored effects against heterogeneous OT environments at scale, which is a necessary condition for achieving the disruption required for the \$100 billion scenario.

4.4 Summary of Both Pathways

We have considered two broad pathways to causing a \$100 billion blackout – by physically damaging equipment and by precipitating and sustaining a cascading blackout via operational disruption.

	# Targets required	Proofs of concept	Key challenges
Physical damage to generators	At least a hundred	1 lab demo (Aurora test); no real-world precedents	• Dramatic scale required • Expensive/challenging to test a single attack; full-scale testing impractical
Physical damage to large transformers	Dozens	No demos; no real-world precedents	• Coordinated physical damage at scale
Disrupting grid operations	Dozens	2015 Ukraine attack (four orders of magnitude smaller)	• Unprecedented real-world attack • Defensive response • Narrow time window • Warning shots

Table 4.3 | Summary of both pathways.

⁹⁰ [Conway et al. \(2020, 9\)](#), discussing the prospects of a hardware supply chain attack on transformers: “If the capability required a call out to be remotely activated or utilized, then it makes even less sense to take advantage of the supply chain in the first place as the call out would be as easily detected as the intrusion to place the capability.” See also [Slowik \(2022a, 11–12\)](#).

⁹¹ To be sure, this sort of exploratory work may nonetheless be useful in developing OT-specific capabilities and conducting operations against similar targets.

⁹² While NotPetya was delivered via supply chain and did cause disruption to grid operators’ IT systems, it was indiscriminate and caused no outage.

Both pathways share a fundamental challenge: unprecedented scale. The physical damage pathway would require destroying dozens to hundreds of critical assets; the operational disruption pathway would require interfering with dozens of independent restoration efforts over days.

This scale magnifies the challenge of evading detection, with each additional target creating opportunities for defenders to notice and intervene. For physical damage attacks, defenders discovering the campaign could disconnect equipment from the grid, eliminating routes to cyber-physical damage. For operational disruption, operators discovering interference could shift to manual restoration procedures. Maintaining covert access across dozens or hundreds of independent entities over the extended timeline required would be extremely difficult.

Beyond scale, both pathways would require achieving effects that have not been demonstrated in the empirical track record. No real-world attack has physically damaged a generator or transformer, nor has any cyberattack triggered a cascading blackout. While individual components of such attacks have been demonstrated in the wild, no single operation has come close to achieving coordinated effects of this nature at the required scale.⁹³

4.5 Implications for a \$10 Billion Damage Threshold

Having established the requirements for a \$100 billion attack, it is worth briefly considering the implications for reaching a lower damage threshold of \$10 billion in economic damages.

A \$10 billion blackout would require one-tenth the disruption: for example, a blackout affecting 100 million people for 10–24 hours, or 50 million people for 1–2 days (see [Section 3.2](#)). This magnitude would be comparable to major accidental blackouts without physical damage to equipment or deliberate efforts to delay recovery, such as the 2003 Northeast Blackout.

This meaningfully changes the feasibility relative to \$100 billion blackouts. It is plausible to achieve \$10 billion in damages by precipitating a cascading blackout of sufficient scale, without causing physical damage to equipment and without sustained interference with power restoration. An attacker may therefore be able to circumvent the two most challenging aspects of achieving \$100 billion in damages. This suggests a qualitative shift in difficulty between the \$10 billion and \$100 billion thresholds. Nonetheless, precipitating a cascading blackout of this scale remains an unprecedented and extremely challenging objective.

This analysis suggests that the precipitate/sustain asymmetry – more so than the scale of initial disruption – drives much of the increased difficulty in achieving the more damaging attack. A \$10 billion attack requires achieving a severe but time-limited disruption, whereas a

⁹³ As described in [Section 6](#), many relevant techniques appear accessible to relatively low-resource actors. This provides some evidence that the key bottleneck to executing such attacks is less specific technical skills, and more the ability to integrate a diverse range of skills across many domains in a complex, dynamic real-world attack.

\$100 billion attack requires either physical destruction at scale or sustained interference with independent restoration efforts to extend outages for many days. Analysis of resources and AI capabilities in Sections [5](#) and [6](#) focuses primarily on the \$100 billion threshold, though [Section 5.4](#) revisits the implications for a lower threshold.

5. The Resources and Capabilities Required for a Catastrophic Grid Cyberattack

Having established the effects a cyberattack would need to achieve to cause \$100 billion in economic damages, we next turn to estimating the capabilities and labor required, drawing on case studies of prior cyberattacks targeting the grid and other infrastructure. A \$100 billion grid cyberattack would likely require a skilled team with diverse offensive cyber and OT capabilities working for months or more. The key requirement appears not to be any one narrow technical skill, but the sustained operational capacity to coordinate and integrate many such skills across complex, multi-target operations.

5.1 Prior OT Cyberattacks Have Required Significant Resources

The track record of significant OT cyberattacks provides some empirical bounds on the effort required to cause a given level of disruption.

5.1.1 Total Effort Required for Historical OT Cyberattacks

This analysis relies on open-source reporting of prior incidents, which comes with important limitations.

First, the public track record is an inherently limited evidence base, consisting of operations that were detected or disclosed, and may understate the capabilities of the most capable actors (see [Section 8.2](#) for further discussion).

Second, this section's estimates of the effort required for historical cyberattacks rely primarily on the timing of visible signatures such as network intrusions, attacker movements within target networks, and disruptive effects. This evidence can yield high-level timelines of attacks and the broad capabilities involved but provides limited visibility regarding the time and effort required for various components of the attack, such as the person-hours devoted to reconnaissance or malware development. Given this, requirements are characterized at a relatively high level, combining attack timelines with evidence on approximate team size from indictments and open-source reporting to estimate total labor requirements.

[Table 5.1](#) summarizes quantitative estimates of the effort required for these attacks, along with their scale and effects. Only attacks with sufficient open-source reporting to estimate team size and attack duration are included. The list is not exhaustive, but is sufficient to establish

order-of-magnitude labor requirements. ([Appendix 5](#) provides a more detailed list of historical OT cyberattacks.)⁹⁴

These labor estimates are uncertain and are intended to give a rough approximation. Since these estimates capture only the marginal labor of offensive cyber operators, they likely undercount the true resources required, which would include supporting staff, overheads, and infrastructure. Moreover, they do not capture the accumulated expertise, access, and tooling developed over prior operations.

Attack	Scale (targets disrupted)	Calendar time (years)	Team size (FTE)	Est. person-years [90%]	Effects
Stuxnet	1 enrichment facility	0.5–4.5	20–50	30 [5–85]	Enrichment program delayed
2015 Ukraine	3 control centers	1.6	3–20	4 [0.7–11]	3.5h outage for 0.5 million
2016 Ukraine	1 substation	1	3–20	3 [0.5–7]	1.3h outage for 0.5 million
Triton	1 plant	0.3–3	1–10	2 [0.3–8]	2 unplanned shutdowns
\$100 billion blackout	Dozens–hundreds	?	?	?	~1 week outage for 100 million

Table 5.1 | Estimated labor inputs to prior OT cyberattacks. *Note:* See [Appendix 6](#) for detailed methodology and sources.

There are several key upshots of this analysis:

- **Scale gap.** A key difference between historical OT attacks and a \$100 billion grid attack is the number of targets. Prior attacks have targeted between 1 and 3 facilities, whereas a \$100 billion grid attack would require achieving effects across dozens or more independent targets. As discussed in [Section 4.3](#), supply chain compromises could reduce the challenge of gaining initial access at scale. However, access is arguably not the principal bottleneck: Achieving disruptive effects at scale requires tailored capabilities across heterogeneous OT environments, which supply chain access does not provide. Even granting significant efficiencies in the access phase, the effort required to develop, test, and deploy OT payloads across dozens or more targets would likely substantially exceed the historical precedents above.

⁹⁴ Notable omissions from this table include: **2014 German steel mill incident** – cyberattack resulted in physical damage to a blast furnace. The incident has been linked to the Dragonfly/Havex campaign by Russia-linked actors, with the damage speculated as having been unintentional (see [Slowik 2021, 7–8](#); [Lee et al. 2014](#)); **2022 Iranian steel plant incident** – deliberate cyberattack with similar consequences, claimed by an Israeli hacktivist group “Predatory Sparrow”, discussed in [Section 6.2](#) – see [BBC \(2022\)](#) and [Times of Israel \(2022\)](#).

- **Resource requirements.** These labor estimates reveal a consistent pattern – attacks achieving relatively modest effects required months or years of sustained effort by multi-person teams.
- **Stuxnet:** Stuxnet is an outlier, having required an order of magnitude greater effort than the other prior attacks. The attack is regarded as one of the most complex state cyberattacks to date and was the first real-world case of cyber-physical damage to equipment. A number of factors make it challenging to compare directly to grid cyberattacks.⁹⁵ Stuxnet is included as the upper end of publicly documented state capabilities and investment in a single operation, but this report makes no definitive claim about whether a \$100 billion grid cyberattack would require more resources.

Even the most impressive historical OT attacks achieved effects far short of the \$100 billion threshold, while targeting far fewer facilities. The order-of-magnitude increase in targets, relative to operations requiring months-to-years of team effort, suggests that a \$100 billion grid cyberattack would require extreme levels of resource and organizational capacity.

5.1.2 Allocation of Effort within OT Cyberattacks

Cyberattacks are complex, multi-stage operations that can be characterized as a sequence of steps to achieve the attacker's ultimate objective. As introduced in [Section 4.3](#), OT cyberattacks have been characterized as having a distinct two-stage form.⁹⁶ In Stage 1, attackers conduct reconnaissance, gain access to enterprise IT networks, and gather intelligence about the target's OT environment. In Stage 2, attackers develop, test, and execute capabilities specific to the target OT environment. See [Figure 5.1](#).

Analysis of historical OT cyberattacks has found that attackers typically use relatively common tooling and capabilities during the first stage of operations and more customized tooling, often tailored to the target environment, in the second stage ([Slowik 2019b, 2–4](#)).

⁹⁵ Important disanalogies: (a) The target network (Natanz enrichment facility) was an extremely hardened military site, with an air-gapped network – in contrast to grid assets (see [Langner 2013b](#)); (b) the attack involved multiple zero-day exploits (see [Langner 2013b](#)); (c) testing the attack chain involved building a replica centrifuge cascade with real uranium hexafluoride ([Langner 2013a](#)); (d) the attack relied heavily on state-level human intelligence, e.g. to intercept centrifuge designs and deliver the malware to the air-gapped network – see [Volkskrant \(2024\)](#). Recent reporting suggests Israeli operatives had physical access to the facility as early as 2000 ([ProPublica 2026](#)).

⁹⁶ See [Assante and Lee \(2015\)](#) on the ICS attack chain.

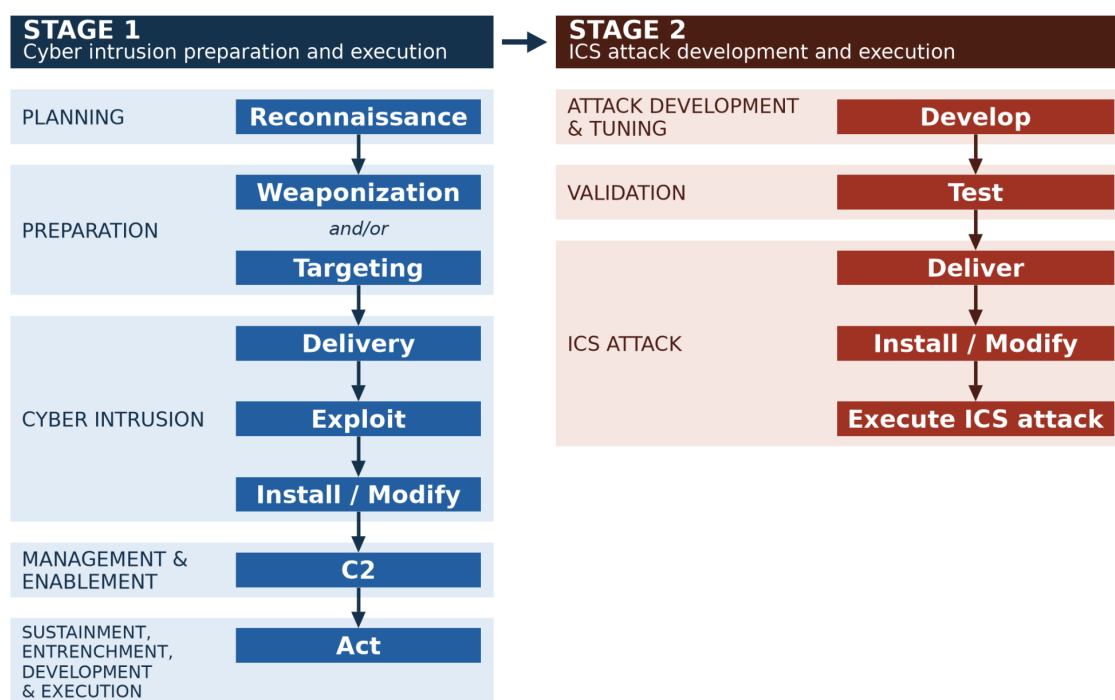


Figure 5.1 | The ICS attack chain. Source: Adapted from [Assante and Lee \(2015\)](#)

As mentioned above, open-source reporting on historical OT cyberattacks provides only limited evidence regarding the time attackers invest in different stages of an operation:

- **2015 Ukraine attack:** There is evidence that preparations were underway in May 2014 – 19 months prior to the ultimate attack – with attempted intrusions on a number of Ukrainian infrastructure targets ([Cys Centrum 2016](#)).
- **2016 Ukraine attack:** The intrusion phase of the 2016 attack appears to have been underway 11 months prior to the attack, while the OT phase was executed in several weeks ([Slowik 2018](#)).
- **Dragonfly campaign:** This intrusion campaign, targeting critical infrastructure, appears to have spanned several years, from as early as 2010 until 2014 ([Slowik 2021](#)). While the campaign has not been linked to any deliberate disruptive effects, it suggests a significant investment in preparatory work.

Broadly, this evidence suggests a majority of calendar time is spent on the preparatory phase of OT cyberattacks – conducting reconnaissance and intrusions, establishing persistent access, gathering intelligence about the target OT environment, and developing tailored OT

capabilities.⁹⁷ The tooling and expertise required in each stage appear to differ qualitatively. [Slowik \(2019b\)](#) describes a bifurcation between the two phases. In the first phase, attackers rely primarily on commodity tools and standard intrusion techniques; in the second phase, they rely on custom tooling tailored to the target environment. The tooling in the second phase tends to be developed with significant input from domain experts and likely draws significantly on reconnaissance and learnings from prior operations. The relatively short calendar time observed for this latter phase should not be taken as evidence that it is straightforward; rather, it reflects the degree to which it depends on expertise, intelligence, and operational capabilities accumulated over months of preparatory work and prior operations against similar targets.

Having established the level of overall effort required in these prior attacks and reviewed the limited evidence on how this effort is allocated between attack stages, we turn next to the skills and capabilities demonstrated in this track record.

5.2 Prior OT Cyberattacks Relied on Diverse Capabilities

Beyond the significant resource and time investments, prior OT cyberattacks have required a diverse range of capabilities (summarized in [Table 5.2](#)).⁹⁸ These operations involved intensive work over months or more to penetrate networks, establish persistent access, and develop custom attack chains. They required significant planning, the execution of a large number of distinct steps, and the integration of diverse capabilities. They also required extensive reconnaissance and intrusion campaigns, deep knowledge of target OT environments, custom malware tailored to specific industrial systems, physical testing infrastructure, and advanced operational security.

Long-term reconnaissance	● 2015 Ukraine: First intrusion attempt was 19 months prior to the attack. ● Triton: Reconnaissance and intrusion campaign almost 3 years prior to the plant shutdown. ● German steel mill: The Dragonfly campaign began at least 4 years prior to the (likely unintentional) German steel mill incident.⁹⁹
Knowledge of OT environments	● 2015 Ukraine: “The strongest capability of the attackers was not in their choice of tools or in their expertise, but in their capability to perform long-term reconnaissance operations required to learn the environment and execute a highly synchronized, multistage, multisite attack.” ● Stuxnet: “The detailed pin-point manipulations of these sub-controllers indicate a deep physical and functional knowledge of the target environment”¹⁰⁰

⁹⁷ This may be true of sophisticated cyber operations more broadly – the UK Ministry of Defence’s cyber primer states that “while [the] preparatory phase can take years, the execution phase may only take seconds” ([UK MoD 2022, 67](#)).

⁹⁸ This appears true of the most significant historical cyberattacks generally ([Halstead and van der Merwe n.d.](#)).

⁹⁹ **2015 Ukraine:** Cys Centrum via ([Maschmeyer 2021](#)); **Triton:** ([DOJ 2021, 5](#)); **German steel mill:** ([Slowik 2021, 3](#)).

¹⁰⁰ **2015 Ukraine:** Quoting [E-ISAC \(2016, 8\)](#); **Stuxnet:** Quoting [Langner \(2013b, 10\)](#).

Process integrity and protection logic	• Triton: Targeted safety instrumented systems designed to keep processes in safe states. • Stuxnet: Manipulated operators' process telemetry to obscure the cause of centrifuge damage. • 2016 Ukraine: Industroyer malware included functionality for manipulating substation protective relays.¹⁰¹
Tailored malware	• Triton malware targeted the specific PLC used in the petrochemical plant and the firmware version installed. • Stuxnet malware targeted the specific PLCs used in the Natanz enrichment facility. • 2015 Ukraine is an interesting mixed case. The disruptive effect was achieved via hijacking legitimate operator control, but other aspects of the attack involved custom tailoring, e.g. the malicious firmware to brick serial-to-ethernet converters.¹⁰² • 2016 Ukraine Industroyer malware was designed to disrupt the working processes of ICS used in electrical substations.¹⁰³
Test environments	• 2015 Ukraine: Adversaries "likely had systems ... that they were able to evaluate and test their firmware against". • Triton: Obtaining the PLC hardware and software "was essential for recreating the full working environment needed for experimenting and writing malicious code".¹⁰⁴ • Some mistakes in prior OT cyberattacks may be explained due to insufficient testing – e.g. Triton malware inadvertently causing plant shutdowns or the failure of the protective relay DoS attack in the 2016 grid attack.¹⁰⁵
Stealth	• Triton attackers interacted with target controllers during times when there were fewer workers on-site to react to alarms. • Stuxnet malware had a range of stealthy features e.g. using stolen digital certificates, hiding binaries in Windows rootkit, checking for and disabling antivirus software, hiding modified code on PLCs via rootkit.¹⁰⁶
Operational orchestration and adaptation	• 2015 Ukraine: Attackers coordinated disruptive actions across three control centers alongside wiper malware and call-center DoS to impede recovery. • Stuxnet: Reportedly involved cooperation between multiple state intelligence agencies, with malware refined across multiple versions during attack preparation.¹⁰⁷

Table 5.2 | Types of capability demonstrated in prior OT cyberattacks.

¹⁰¹ [Slowik \(2019c\)](#) argues that Stuxnet, the 2016 Ukraine attack, and Triton are best understood as attacks on process integrity, protection, or safety. Undermining these systems may be a prerequisite for causing physical damage via cyberattack.

¹⁰² [\(E-ISAC 2016, 7–8\)](#).

¹⁰³ **Triton:** [Slowik \(2019c, 13–14\)](#); **Stuxnet:** [Hemsley and Fisher \(2018, 11\)](#); **2016 Ukraine:** [ESET \(2017, 3\)](#).

¹⁰⁴ **2015 Ukraine:** quoting [E-ISAC \(2016, 13\)](#); **Triton:** quoting [Di Pinto et al. \(2018\)](#).

¹⁰⁵ [Mandiant \(2017\)](#): "... Mandiant ran trilog.exe in a lab environment with a valid Triconex controller and discovered a conditional check in the malware that prevented the payload binary from persisting in the environment. Mandiant confirmed that, after correcting patching [sic] the attack script to remove this check, the payload binary would persist in controller memory, and the controller would continue to run." [Slowik \(2019a, 4\)](#) "One possibility for this error may be test environment. Protocol emulators in software, such as the publicly available IEC Server project,¹⁷ do not enforce statefulness and associated communication timeouts. Physical hardware will employ such items in line with the vendor-specific implementations of the relevant communication standard."

¹⁰⁶ **Triton:** [Mandiant \(2017\)](#); **Stuxnet:** [Hemsley and Fisher \(2018, 11\)](#).

¹⁰⁷ **2015 Ukraine:** [E-ISAC \(2016\)](#); **Stuxnet:** [Langner \(2013b\)](#); [NYT \(2011\)](#); [Volkskrant \(2024\)](#).

The distinct attack stages identified in [Section 5.1.2](#) reinforce the diversity of capabilities required, insofar as intrusion and OT-specific phases draw on different skills and expertise ([Slowik 2019b](#)).

Altogether, evidence from historical OT cyberattacks suggests that attacks have relied on a diverse range of capabilities across the attack chain and have involved extended timelines from reconnaissance to ultimate effects. Importantly, no evidence could be found that attacks are bottlenecked by any one narrow technical capability, such as powerful exploits or malware development.¹⁰⁸ (Indeed, the use of zero-day or patched vulnerabilities in recent OT cyberattacks appears rare.)¹⁰⁹ This is consistent with characterizations of real-world offensive cyber operations as complex and dynamic, requiring attackers to continuously adapt to specific contexts and defensive actions ([Gennari et al. 2024](#)).

The capabilities required might be better characterized as sustained institutional capacity, including structures for coordinating diverse teams, years of preparatory investment in infrastructure and access, and deep operational experience that cannot be rapidly acquired. For example, Smeets characterizes state offensive cyber capability as being composed of people, exploits, tooling, infrastructure, and organizational structure. He argues that the utility of known exploits and tools depends heavily on testing and target-specific knowledge, which is costly and hard to assemble.¹¹⁰ Similarly, UK Ministry of Defence doctrine notes that successful cyber campaigns require a “high degree of integration and cooperation between units and organisations”.¹¹¹

Having established both the resource scale and capability diversity required for historical attacks, we can now assess what a \$100 billion grid attack would require.

5.3 A \$100 Billion Grid Cyberattack Would Require Unprecedented Capabilities and Scale

The capability diversity documented in [Section 5.2](#), together with sustained team efforts shown in [Section 5.1](#), suggests that the barrier for disruptive OT cyberattacks is not any one narrow tool or technique, but the coordination and integration of diverse capabilities over extended periods.

[Section 4](#) established that a \$100 billion grid cyberattack would need to achieve one of the following: (a) physically damage dozens or more critical grid assets in a coordinated manner,

¹⁰⁸ See [Buchanan \(2017\)](#) for a broader argument that sophistication in cyber operations is driven less by narrow capability than by the integration of diverse skills in complex operations.

¹⁰⁹ ([Slowik 2019b, 10](#)). Stuxnet is a clear exception, involving multiple zero-day exploits. The 2016 Ukraine attack made use of a patched exploit as part of the denial-of-service of Siemens SIPROTEC protective relays.

¹¹⁰ See [Smeets \(2022\)](#), particularly Chapter 5.

¹¹¹ ([UK MoD 2022, 68](#)).

or (b) precipitate a cascading blackout and sustain the outage for roughly a week via coordinated interference across dozens or more restoration efforts.

The scale required to achieve \$100 billion in damages via a grid cyberattack presents a major hurdle. The prior attacks reviewed above targeted 1–3 facilities and required months or years of effort by skilled teams. A \$100 billion attack would require coordinating effects across 10–100x more targets. Given the capability diversity required even for single-target operations, this scaling presents unprecedented challenges in customization (adapting techniques to heterogeneous systems and environments), coordination (synchronizing effects), and stealth (maintaining covert presence across many networks).

Moreover, each of the attack pathways would require achieving effects without clear precedent. No real-world cyberattack has physically damaged a large transformer or generator. And there are no documented examples of a cyberattack triggering a cascading blackout, much less sustaining a widespread outage for many days. A \$100 billion grid cyberattack would therefore demand resources and organizational capacity substantially exceeding the historical precedents examined above.

5.4 Implications for a \$10 Billion Threshold

The above analysis of \$100 billion attacks identifies two key challenges relative to historical attacks – achieving much greater scale and achieving unprecedented effects. At the \$10 billion threshold, these challenges are reduced. In principle, an attacker might reach this threshold by precipitating a large cascading blackout during stressed grid conditions, without also physically damaging grid equipment at scale or interfering with independent restoration efforts over several days. Triggering a major cascade via cyberattack would still be an unprecedented effect, far beyond those realized in prior grid cyberattacks. But an attacker would not have to solve the additional problem of sustaining a nationwide blackout for roughly a week. Altogether, it cannot be ruled out that a \$10 billion grid attack is achievable with resources and capabilities closer to those demonstrated in prior OT cyberattacks by states.

6. AI Uplift Analysis

In this section, we assess the AI capabilities that could meaningfully increase the risk of catastrophic grid cyberattacks by uplifting threat actors. This analysis distinguishes two pathways. The clearest route to a large increase in marginal risk would be AI systems enabling lower-skilled actors to execute attacks that currently require state-level operational capacity. This appears to require an extremely high level of capability, approaching full automation of complex offensive cyber operations against OT. A second pathway is partially uplifting state-level actors, which may require considerably lower levels of capability. However, since these actors appear significantly more constrained by strategic factors than by capability, it is much less clear that such uplift would translate directly into significant risk.

6.1 Methodology

Defining Threat Actors

Following [NIST \(2025, Appendix E\)](#), misuse threat models should specify not only the type of attack and relevant AI capabilities, but also the threat actors who might pursue such attacks. The analysis is therefore structured around five threat actor classes, ranging from individual hobbyist hackers to the world’s most capable states. This classification draws on prior work by [RAND \(2024\)](#). This categorization is heavily stylized – threat actors do not fall neatly into discrete classes, and the boundaries between categories (particularly TA3 and TA4) are contestable.

Threat actor	Description	Examples	Est. population ¹¹²
TA1	Single individual, limited infosec expertise, <\$1k budget for the specific operation, no infrastructure.	Hobbyist hacker	~1m
TA2	Single individual, professional infosec capability, ~\$10k budget for the specific operation, personal infrastructure.	Individual professional hackers	10k–100k
TA3	Team of ~10 experienced professionals, ~\$1m budget and months of effort for the specific operation.	Criminal hacking groups	100–1k
TA4	Team of ~100 state-level experts, ~\$10m budget and a year or more of effort for the specific operation, state resources.	Second-tier states (e.g. Iran, North Korea)	10–50
TA5	Team of ~1,000 top experts, ~\$1bn budget and years of effort for the specific operation, state-level resources, infrastructure, and access.	Top-tier states (e.g. US, China)	~5

Table 6.1 | Threat actor definitions. See [Appendix 7](#) for a more detailed explanation.

¹¹² These estimates are from our survey of experts and superforecasters, and are inherently uncertain.

AI Uplift

In the analysis that follows, “AI uplift” is used to refer to changes in threat actors’ ability to execute an attack. Such uplift could occur via several routes: enabling actors to do tasks they were previously incapable of doing; reducing the resources required to do some task; or enabling them to perform tasks better.

[Section 5](#) concluded that the bottleneck to \$100 billion grid cyberattacks is not narrow technical capabilities, but rather the ability to integrate a diverse range of capabilities across complex, dynamic operations.

This analysis focuses on attacker uplift, holding defensive capabilities fixed. Yet the same capabilities could help defenders harden their systems, detect intrusions, and respond to incidents. The overall effect of AI on risk will depend on the balance between offensive and defensive gains, and the relative pace of adoption, which this report does not attempt to model.

A Simple Risk Model

We can model the annual probability of at least one attack from a given class of threat actors as the product of two parameters:

- **Capability:** The probability that a randomly selected actor in this class could launch a cyberattack on the US grid causing at least \$100 billion in economic damages with six months of effort.
- **Conditional willingness:** The annual probability that at least one actor in this class would spend six or more months actively attempting to launch such an attack, if capable.¹¹³ This incorporates both the likelihood that a given actor would attempt an attack given strategic and operational conditions, and the size of the actor class.¹¹⁴

For a given class of threat actors i :

$$\text{Annual probability of at least one attack from class } i = \text{Capability}_i \times \text{Conditional willingness}_i$$

For simplicity, each threat actor class is treated as having a single baseline capability level and a single conditional-willingness level, abstracting away from variation among actors within the class. For the purposes of this simplified model, we can represent AI uplift as increasing capability while holding fixed conditional willingness ([Frontier Model Forum 2025](#)). This should be understood as a stylized model to organize the analysis, rather than a substantive

¹¹³ The *ex ante* probability that at least one threat actor in each category would be willing to launch an attack probably increases with the number of actors in that category. In retrospect, for this reason, it may have been preferable to construct a risk model that includes a separate parameter for the number of actors in each class.

¹¹⁴ Willingness therefore captures all-things-considered disposition – if an actor would strongly *desire* to conduct an attack, but ultimately would not do so, for fear of retribution, they are treated as not willing.

claim about the nature of AI uplift. In reality, capability and willingness are difficult to disentangle. For example, capabilities that reduce the risk of attribution or that reduce the cost of operations might make a given actor both more capable and willing to carry out an attack.

Distinguishing between capability and willingness is most straightforward for TA1 and TA2 actors, where the key question is whether AI could sufficiently lower barriers to enable willing-but-currently-incapable actors to carry out cyberattacks on the grid. It is much less clear for the TA4 and TA5 actors, where the primary effect of AI may be making attacks that are already feasible less costly, less attributable, or more reliable. In such cases, AI uplift may operate through capability, conditional willingness, or both.

6.2 Baseline Capability and Willingness of Threat Actors

There appears to be a roughly inverse relationship between threat actor willingness and cyber capability. The track record of disruptive OT cyberattacks is dominated by state-linked actors. As we will see, less capable actors have repeatedly attacked grid infrastructure through physical means, while states' willingness to launch catastrophic grid cyberattacks during peacetime appears significantly constrained by the risk of escalation and retaliation.

Survey Results

As noted in [Section 1.4](#), in collaboration with the Forecasting Research Institute, we conducted a small survey of 8 experts and 13 forecasters to inform this analysis. We asked respondents to estimate baseline capability and willingness for each threat actor class. The table below presents the results.¹¹⁵

¹¹⁵ We asked equivalent questions for \$10bn attacks, and present these results in [Appendix 8](#). The broad pattern is similar – capability increases and willingness decreases with threat actor level – though baseline capability estimates are higher across all actor classes, consistent with the lower requirements discussed in [Section 4.5](#), and the difference in baseline probability estimates of \$100 billion and \$10bn attacks in [Section 2.3](#).

Threat actor	Capability		Conditional willingness	
	Probability a given actor is capable of launching a \$100 billion grid cyberattack with six months of effort		Annual probability at least one actor spends six or more months actively attempting the attack, if capable	
	Experts	Forecasters	Experts	Forecasters
TA1	0%	0%	80%	60%
TA2	0%	0%	48%	54%
TA3	0.01%	1%	11%	33%
TA4	0.2%	13%	12%	5%
TA5	2%	58%	5%	4%
Relationship	Increases with TA level		Highest for lower TA levels	

Table 6.2 | Survey estimates for \$100 billion grid cyberattack capability and conditional willingness.

Experts and forecasters broadly agree that TA1, TA2, and TA3 actors are extremely unlikely to be capable today of launching a cyberattack causing \$100 billion in damages. The survey found substantial divergence in estimates of TA4 and TA5 capability between experts and forecasters, with forecasters estimating significantly higher likelihood that either class is currently capable. Experts and forecasters particularly disagreed on TA5 capability, which forecasters place at 58% and experts place at 2%. This divergence likely reflects the significant uncertainty about nation-state capabilities, the limitations of our small survey sample, and an ambiguity in the survey question discussed below. There is somewhat less divergence between groups on actor willingness. The largest disagreement concerns TA3 willingness, which experts place at 11% versus forecasters at 33%.

The pattern in capability and willingness has implications for the marginal risk posed by different threat actor groups given AI uplift. [Table 6.3](#) shows that, under the unrealistic assumption that all threat actors are uplifted to 100% capability, marginal risk would be dominated by lower-skilled actors, given their high baseline willingness and low baseline capability. This is not to say that such actors pose the greatest marginal risk in the near term. Rather, it illustrates that uplifting these actors would be particularly consequential.

Threat actor	Capability	Conditional willingness	Maximal risk from maximal uplift
TA1	0.0→100%	80%	+80pp
TA2	0.0→100%	48%	+48pp
TA3	0.01→100%	11%	+11pp
TA4	0.2→100%	12%	+12pp
TA5	2→100%	5%	+5pp

Table 6.3 | Illustrative marginal risk under maximal capability uplift. *Note: Experts’ baseline estimates.*

Note that the estimates in [Table 6.2](#) cannot be straightforwardly aggregated into an overall baseline risk estimate. If p_i is the annual probability of at least one attack from threat actor class i , then under an independence assumption the probability of at least one attack from any class is $1 - \prod_i (1 - p_i)$. Applying this aggregation to the forecaster estimates in [Table 6.2](#) yields an overall risk estimate that is much higher than forecasters’ median baseline-risk estimate in [Section 2.3](#) – about 3.3% per year, compared with a direct estimate of 0.1% per year.¹¹⁶ This may be due to an ambiguity in how capability estimates were elicited in the survey, detailed in the footnote below.¹¹⁷ The author places greater weight on the direct elicitation of baseline risk, which is not subject to the same ambiguity.

Explaining the Capability–Willingness Relationship

Historical evidence and strategic considerations support the relationship between capability and willingness for the most and least capable actors. The most capable actors appear the least willing, while the far more numerous lower-tier actors appear more willing but are far short of the required capability.

¹¹⁶ Aggregation of forecaster class estimates: $1 - \prod [1 - (58\% \times 4\%), 1 - (13\% \times 5\%), 1 - (1\% \times 33\%)] \approx 3.3\%$ (vs. median overall risk estimate of 0.1%). Experts’ aggregated estimate aligns more closely with their median overall risk estimate: $1 - \prod [1 - (2\% \times 5\%), 1 - (0.2\% \times 12\%), 1 - (0.01\% \times 11\%)] \approx 0.1\%$.

¹¹⁷ Willingness was elicited as “the probability that at least one actor in each threat actor class would spend 6 or more months actively attempting to launch [a \$100 billion grid cyberattack]”; capability, as “the probability that a randomly selected threat actor in that level is currently capable of launching [a \$100 billion grid cyberattack]”. If capability of launching an attack is read as including the ability to achieve the \$100 billion outcome, experts’ parameters are coherent with their direct estimate, whereas forecasters’ are not. If capability is instead read more loosely – ability to launch an attack without guarantee of success – the forecasters’ parameters are coherent (implying that around 1 in 30 attacks launched succeed), whereas experts’ imply that attacks launched will approximately always succeed, which may be at odds with the historical evidence in [Sections 4 and 5](#). The author suspects that respondents resolved this ambiguity differently, in which case the divergence in TA5 capability (2% vs. 58%) may reflect question interpretation more than substantive disagreement. Experts’ TA5 capability estimates fall sharply between the \$10 billion and \$100 billion thresholds (25% to 2%), whereas forecasters’ do not (65% to 58%) – this is consistent with the author’s view that experts interpreted capability as including achievement of the outcome, whereas forecasters did not.

Willingness

Given the escalatory risk and likelihood of retaliation, the most capable actors appear to be significantly constrained in their willingness to carry out a catastrophic cyberattack on the grid during peacetime. A \$100 billion grid cyberattack would almost certainly be treated as an act of war, inviting severe diplomatic and military responses.

The most capable state actors have conducted extensive pre-positioning in US critical infrastructure networks, but pre-positioning does not clearly indicate willingness to carry out catastrophic attacks in peacetime. The most prominent examples of such activities are the China-linked “Volt Typhoon” and Russia-linked “Dragonfly” campaigns ([CISA 2024](#); [Slowik 2021](#)). Pre-positioning is often interpreted as developing capabilities for, and signaling possible courses of action in the event of a crisis or conflict.¹¹⁸

A key uncertainty concerns the conditions under which adversaries would activate these pre-positioned capabilities. If catastrophic grid cyberattacks were reserved for war or major crises, then the annual probability of attempts by such actors may be roughly bounded by the risk of major conflict between the US and its adversaries, which forecasting platforms place on the order of 1% per year.¹¹⁹ However, this may underestimate the risk from such actors, to the extent they are willing to launch such attacks in scenarios that fall short of war or as part of “gray zone” activities. It is less clear the extent to which similar strategic constraints hold for second-tier state actors, such as Iran or North Korea, whose behavior appears less predictable.

In active conflict situations, where state actors are already willing to inflict large-scale infrastructure damage, kinetic weapons have proved considerably more effective than cyberattacks. As we have seen, Russia’s cyberattacks against Ukraine’s grid caused brief outages affecting a few hundred thousand people. In contrast, kinetic attacks during the current war have destroyed much of Ukraine’s generation and transmission capacity, causing months of rolling blackouts for tens of millions – roughly four orders of magnitude more disruption than the cyberattacks.¹²⁰ (See [Appendix 9](#) for a detailed comparison of cyber vs. kinetic attacks on the grid.) For capable state actors during peacetime, the comparison is less straightforward, since cyberattacks plausibly retain strategic advantages over kinetic, insofar as they are generally more deniable, easier to pre-position, and less clearly escalatory.

¹¹⁸ See e.g. the analysis of Chinese and Russian cyber activities in [ODNI \(2023\)](#).

¹¹⁹ As of 27 May 2026, the forecasting platform Metaculus estimates a 12% chance of US–China war before 2035 ([Metaculus n.d.a](#)), implying **1.5%/year**, assuming a constant hazard rate. The platform estimates a 10% chance of US–Russia war before 2050 ([Metaculus n.d.b](#)), implying **0.5%/year**, assuming a constant hazard rate.

¹²⁰ [Dixi Group \(2025\)](#) and [Energy Map \(2025\)](#) report 1,951 hours of rolling blackouts in 2024 affecting “all or the vast majority” of regions. During outage hours, a mean of 2 of 6 customer groups were simultaneously disconnected. Assuming all regions affected and 33% of ~17.5 million household electricity customers without power at any given time gives ~11 billion customer-hours of outage (1,951h × 33% × 17.5M).

By contrast, lower-capability actors appear far more willing to attack grid infrastructure, and some may be willing to cause catastrophic harm given the capability. This is perhaps unsurprising, since these actors are considerably more numerous. Our survey estimates there are on the order of 1 million TA1 actors, so one would expect a much greater range of motivations. This is corroborated by the frequency of physical attacks and threats against US grid infrastructure by individuals and extremist groups, which have caused dozens of power outages in recent years.¹²¹ Attempts to achieve relatively small-scale disruption, however, do not necessarily imply willingness to cause catastrophic levels of damages envisaged in this report’s main scenario.¹²²

For lower-capability actors, kinetic attacks appear significantly more accessible and effective than cyber. A 2022 substation attack in Moore County, North Carolina offers a useful comparison. One or two perpetrators shot at two substations, causing a localized blackout lasting several days.¹²³ Preparations for the attack appear to have taken no more than 16 days.¹²⁴ The resulting outage exceeded the combined effects of Russia’s 2015 and 2016 grid cyberattacks.¹²⁵ In other words, a kinetic attack achieved greater impacts than Russia’s grid cyberattacks, and appears to have required substantially less effort.¹²⁶ To be sure, cyberattacks have numerous advantages for lower-skilled threat actors, such as the ability to cause disruption without physical proximity and potentially lower risk of attribution.

	2022 Moore County attack	2015 and 2016 Ukraine grid attacks
Actor	• 1–2 people	• Russian state-linked hacking group • Est. 3–20 technical staff
Resources	• High-powered rifle and minivan • <0.1 person-years effort	• 2.5 years calendar time preparation • ~7 person-years effort
Effect	• Physical damage to two substations • 3 million customer-hours of outage	• No physical damage • 1 million customer-hours of outage

Table 6.4 | Cyber vs. kinetic grid attack comparison.

¹²¹ Total of 100–185 incidents per year; 35–60 causing outages per year. Statistics for 2021–23 via [Politico \(2022\)](#); [Politico \(2023\)](#); [Politico \(2024\)](#).

¹²² Such motivations are not unheard of – in one recent criminal case the perpetrator expressed willingness to cause “billions of dollars” in economic damages ([United States v. Clendaniel and Russell 2023, 4](#)).

¹²³ [The Pilot \(2023\)](#). Reporting suggests the attack was motivated by opposition to a local drag show – the attack coincided with the event, and the person of interest was linked to a local activist group who opposed it.

¹²⁴ The attack was on December 3, 2022. The drag show was announced on November 18, 2022 ([Sandhills Pride and Sunrise Theater 2022](#)), suggesting that, if drag show–motivated, preparation time was no more than 16 days.

¹²⁵ Note this table uses units of customers rather than people (one electricity customer roughly corresponds to one household). **Moore County outages:** See [spreadsheet](#) for calculation using outage information from [Wikipedia \(n.d.\)](#). **Russia’s 2015 and 2016 cyberattacks:** 2015: 0.225m customers for 3.5h; 2016: $\leq 0.225\text{m}$ customers for 1.25h; $0.225\text{m} \times 3.5\text{h} + 0.225\text{m} \times 1.25\text{h} \approx \sim 1\text{m}$ customer-hours of outage.

¹²⁶ Compared to the author’s central estimate of 7 person-years for the 2015 and 2016 Ukraine attacks combined. See [Table 5.1](#) and [Appendix 6](#) for details of staffing and preparation time estimates.

Capability

The historical record is consistent with the survey's finding of a stark capability differential between state actors and lower-skilled actors. All known cyberattacks resulting in blackouts via manipulation of grid control systems have been attributed to state-linked actors (see [Table 2.1](#)), as have the significant disruptive OT cyberattacks reviewed in [Section 5](#). One apparent exception – a 2022 cyberattack on an Iranian steel plant that caused a major fire and was claimed by the hacktivist group “Predatory Sparrow” – has been credibly linked to Israeli state cyber operations.¹²⁷

However, the track record provides only limited evidence about the frontier of state capabilities – it includes only those attacks that have been publicly documented and attributed, and reveals only those capabilities that states have chosen to deploy. This limited evidence may partly explain the divergence in survey respondents' estimates of TA4 and TA5 capabilities. Even so, a \$100 billion grid cyberattack would represent a significant leap beyond the documented track record.

Implications for AI Uplift

The important implication is that lower-tier actors appear much less capable than state actors, while also appearing more willing to cause grid disruption if they could. On the other hand, state actors may already possess much of the operational capacity required, but face significant strategic constraints limiting their willingness to do so.

This implies two distinct questions about AI uplift:

- For TA1/TA2 actors: Could AI substitute for most of the operational capacity required to execute a state-level OT cyber campaign?
- For TA4/TA5 actors: Could AI materially lower the barriers to attacks that are already feasible or near-feasible?

6.3 AI Uplift Pathways

We now turn to a central question of this report: What AI capabilities would meaningfully increase the risk of catastrophic grid cyberattacks? The analysis above points to at least two ways in which AI could increase this risk:

- By enabling TA1/TA2 to launch \$100 billion grid cyberattacks, which – given their high conditional willingness – would likely meaningfully increase risk.

¹²⁷ See [BBC \(2022\)](#) on the attack and initial attribution. See [Times of Israel \(2022\)](#) on link to state operations, and [NYT \(2021\)](#) on US officials attributing a *previous* attack by the group to Israel. Notably, the attack seems to have closely mirrored a 2014 incident at a German steel mill, suggesting attackers deliberately chose an approach already demonstrated in the wild, which appears consistent with the operational discipline of state actors.

- By lowering the barriers to such attacks by TA4/TA5 actors. The extent to which this would increase the risk of such attacks is less clear, given the strategic and contextual factors shaping their willingness.

This analysis focuses primarily on lower-skilled and state-level actors because they illustrate the two clearest uplift pathways. TA3 actors occupy an intermediate position. They have more operational capacity than individuals but appear much less capable than state-level teams. This report's author has found it harder to draw firm conclusions about this heterogeneous class of actors. Uplifting TA3 actors would plausibly require less than full TA1/TA2 uplift but more than narrow assistance to state-level actors.

6.3.1 Uplifting Lower-Skilled Actors

Lower-skilled actors are currently very far from possessing the capability to launch \$100 billion grid cyberattacks. As we have seen in Sections 4 and 5, such an attack plausibly requires the resources, capabilities, and organizational capacity associated with state actors – far beyond that of TA1/TA2.

The challenge facing lower-skilled actors is not simply that they lack a few exquisite technical capabilities, such as particularly powerful exploits or malware. There are numerous examples of independent researchers with limited resources (TA1/TA2) replicating or independently demonstrating impressive technical elements from state OT cyberattacks (see [Box 6.1](#)).

Box 6.1. Replicating OT attack capabilities at low cost

Examples of independent researchers replicating important aspects of OT cyberattacks at relatively low cost include:

- **2015 Ukraine attack:** [Larsen \(2020\)](#) argues that the skill level of the attack has been overblown and reports having done a more complex hack of similar serial-to-ethernet converters in 14h and the full attack chain in 3 weeks.¹²⁸
- **2016 Ukraine attack:** In May 2016, a 17-year-old competitor in a Russian capture-the-flag competition demonstrated a DoS of Siemens protection relays, as part of a simulated attack on a transmission substation. This technique was demonstrated 7 months later in the 2016 Ukraine attack.¹²⁹
- **Triton malware:** [Di Pinto et al. \(2018\)](#) replicated the malware and concluded the skill and resources were significant but “certainly not at the [nation-state] level”. For \$5–10k they were able to buy the relevant equipment, software and documentation from auction sites [\(Krotofil et al. 2018, 17–23\)](#).

¹²⁸ He claims “[t]hese are doable things; they don’t take a lot of effort ... people other than APT guys can do it,” including his “bong-smoking hacking buddies.”

¹²⁹ [\(PHDays 2016\)](#). See also [this presentation](#) on the event (translated via Google).

- **Grid cyberattacks:** [Keliris et al. \(2020\)](#) demonstrate having built an accurate grid topology model using open-source information, having bought specialist grid equipment on eBay for \$1,000 and ultimately having found 2 vulnerabilities which could be used to modify circuit breaker behavior through hardware testing.

To be sure, replicating known techniques is easier than originating them and demonstrating techniques in a laboratory setting is considerably less challenging than using them in real-world attacks. Nonetheless, these examples suggest that the most capable actors are distinguished less by access to uniquely sophisticated technical expertise than by the operational capacity to integrate a wide range of capabilities into complex campaigns.

Uplifting TA1/TA2 actors would therefore require AI systems that can substitute for much of this operational capacity, not merely assist with isolated technical tasks. In effect, this requires capabilities approaching whole-operation automation – AI that can substitute for most of the skilled work involved in executing complex, dynamic offensive OT cyber operations. This is an extremely high capability threshold.

6.3.2 Uplifting State-Level Actors

The picture is substantially different for state-level actors. TA4/TA5 actors already have substantial teams, accumulated infrastructure and access, and operational experience. These actors could receive meaningful uplift from AI capabilities short of the high threshold identified above, if they reduce the cost, time, staffing, or risk involved in executing large OT operations. This report's analysis of this pathway is less developed and more speculative than its analysis of TA1/TA2 uplift, given the limited visibility into the capabilities and decision-making of state actors.

This section's analysis of how state actors can be uplifted by AI draws heavily on the discussion in Sections 4 and 5. A \$100 billion grid cyberattack would require achieving coordinated disruptive effects across many targets, whether by physically damaging critical equipment at scale or precipitating and sustaining a widespread blackout. Prior OT cyberattacks have required sustained effort across reconnaissance, intrusion, developing and testing tailored OT payloads, stealth, and coordination. AI assistance could lower barriers by reducing the time, cost, or labor required for human operators to conduct these attacks.

Capabilities sufficient to automate specific parts of the attack chain – for example, lowering the human labor requirements for reconnaissance or initial intrusion – will likely precede whole-operation automation. Such partial automation could make OT operations cheaper, faster, more reliable, or more scalable.

The two-stage structure of OT cyberattacks discussed in [Section 5.1.2](#) offers a useful framework for delineating these capabilities.¹³⁰ In Stage 1, attackers gain access, conduct reconnaissance, establish persistence, and identify promising target environments. In Stage 2, attackers translate access into operational effects: interpreting OT systems, developing or adapting payloads, testing physical or process consequences, maintaining stealth, and coordinating disruptive effects.

Acceleration of Stage 1 capabilities may arrive earlier, insofar as these overlap substantially with broad offensive cyber capabilities. This may not translate directly into uplift, if attackers cannot reliably convert access into significant physical and operational effects. Stage 2 uplift bears more directly on the bottlenecks that distinguish grid cyberattacks from ordinary IT intrusions: target-specific OT reasoning, payload adaptation, effect validation, and coordinated execution under defensive response. [Table 6.5](#) sketches several clusters of capabilities, primarily in Stage 2, that appear particularly relevant.

Uplift channel	Potential effect
Stage 1 acceleration	Compressing reconnaissance, intrusion, persistence, and triage of target environments
Target-specific OT reasoning	Speeding up interpretation of target protocols, device behavior, control logic, etc.
Payload and exploit adaptation	Faster tailoring of tools to heterogeneous environments
Testing and validation of physical or process effects	Better simulation or reasoning about attack effects, reducing the need for realistic test environments
Stealth and adaptive execution	Improving evasion, telemetry manipulation, and adaptation under defensive response
Campaign orchestration	Enabling coordinated effects on more targets per unit effort

Table 6.5 | Plausible channels for uplifting state-level actors.

However, the relationship between capability uplift and overall risk is less direct for state-level actors than for TA1/TA2. In contrast to lower-skilled actors, TA4/TA5 actors may often be constrained as much by strategic logic as by capability. Launching catastrophic cyberattacks against the US grid is unlikely to serve their interests outside of relatively narrow circumstances, such as active conflict or geopolitical crises. AI systems that reduce the resources involved in launching such attacks therefore have a less clear effect on overall risk.

The more important question may be whether AI capabilities can change the strategic logic for state-level actors. For example, were AI to substantially reduce the risk of attribution,

¹³⁰ ([Assante and Lee 2015](#)).

increase the reliability of catastrophic effects, or enable coordinated effects across more targets, such attacks could become more attractive in a wider range of circumstances.

7. Implications

7.1 Risk Prioritization

These two uplift pathways have different implications for risk management. Fully uplifting TA1/TA2 actors would have the clearest and most direct effect on marginal risk, given their high conditional willingness and very low baseline capability. However, this pathway requires AI systems to substitute for most of the work involved in complex OT cyber operations: reconnaissance, intrusion, target-specific OT reasoning, developing and testing tailored payloads, maintaining stealth and persistence, and coordinating multi-target effects. This is a very high threshold.

This capability threshold is also extremely broad. AI systems capable of fully uplifting TA1/TA2 actors in this threat model would likely be capable of automating or substantially accelerating many other complex multi-stage activities, including other cyber operations, AI research and development, and scientific research in high-consequence domains. This limits the usefulness of the \$100 billion grid cyberattack threat model as a narrow near-term trigger for targeted safety mitigations.

This broad threshold contrasts with the narrower threshold identified in recent analysis of AI-enabled computer worms. [Halstead and Righetti \(2026\)](#) identify elite exploit development as a key bottleneck for AI-enabled computer worms capable of causing comparable economic damages. AI systems will likely reach this narrow capability threshold well before they can automate the sustained, coordinated work of state cyber teams. That threat model, and others plausibly bottlenecked by narrower capabilities, may therefore be better candidates for near-term safety mitigations.

Partial uplift of state actors could occur much earlier, since it could result from AI accelerating specific stages of an attack. State actors appear constrained at least as much by strategic considerations as by capability. If their willingness is limited primarily by the risk of attribution and escalation or, in active conflict, by the availability of more effective kinetic options, then AI systems that reduce cost or staffing requirements may not substantially increase attack probability. As argued in [Section 6.3.2](#), the most concerning forms of state-actor uplift may be capabilities that shift this strategic calculus.

It is also worth reflecting on the \$100 billion damage threshold at the center of this report. As discussed in Sections [4](#) and [5](#), \$10 billion grid cyberattacks appear qualitatively easier, since they may be achievable through a sufficiently large cascading blackout without the additional challenge of sustaining disruption over days or damaging equipment at scale. Enabling threat actors to achieve smaller-scale attacks will likely require lower levels of capability. Lower damage thresholds may therefore yield more decision-relevant capability thresholds.

7.2 Risk Monitoring

Monitoring whether AI systems are approaching these capabilities will be important both for decision makers concerned about catastrophic grid cyberattacks specifically and for measuring broader progress toward advanced offensive cyber capabilities.

Our survey of experts and forecasters provides some insight into the types of evidence that would be most informative. We asked respondents how they would update their baseline risk estimate for \$100 billion grid cyberattacks, conditional on two hypothetical pieces of evidence. While the hypothetical evidence concerns uplift to lower-skilled actors, the overall risk estimate concerns attacks by any class of actor:¹³¹

- **Uplift in OT-specific capture-the-flag competition:** A study conducted at the end of 2025 finds that, with AI assistance, individual TA1 actors complete as high a proportion of tasks in an OT-specific CTF-style competition as a team of ten experienced cybersecurity professionals (equivalent to TA3) without AI assistance, when both are given the same amount of time.¹³²
- **Warning shot attack:** A cyberattack on the US grid causes a blackout leading to at least \$100 million in economic damages and is subsequently confirmed by credible sources to have been perpetrated by a TA1- or TA2-level actor using AI.

The CTF result represents substantial uplift to lower-skilled actors, elevating individual novices to match the performance of professional teams on OT-specific challenges. Despite this, experts and forecasters only modestly updated their overall risk estimates from 0.1% to 0.2–0.25% on the basis of this evidence. On the other hand, real-world evidence that AI is enabling grid attacks by lower-skilled actors would cause respondents to update significantly. Conditional on such an attack occurring, experts and forecasters updated their risk estimate from 0.1% to 1.3–1.4% (See [Figure 7.1](#)).¹³³

¹³¹ We asked the same questions for \$10 billion grid cyberattacks. These additional results are collected in [Appendix 8](#).

¹³² The survey suggested the study would be similar to the Dragos annual CTF competition ([Dragos 2024](#)), which is focused on OT and industrial control systems.

¹³³ Respondents noted that CTF challenges fail to capture the coordination, persistence, and stealth required for real-world operations, whereas a successful attack – even at limited scale – would demonstrate that AI had meaningfully lowered technical barriers.

Our survey also provided some evidence on the expected timelines for reaching these capability milestones. We asked respondents to estimate the year in which these milestones would be achieved. Estimates for CTF uplift were tightly clustered (median 2028 for experts, 2030 for forecasters), while the warning shot milestone showed more divergence (median 2031 for experts, 2035 for forecasters; expert IQR 2029–2044, forecaster IQR 2032–2036).

Capability 3: TA1 actor uplifted by AI to match a TA3 actor in an OT-specific CTF-style competition.

Capability 4: a US grid cyberattack causing a $\geq$ \$100 million blackout, attributed to a TA1/TA2 actor with AI assistance.

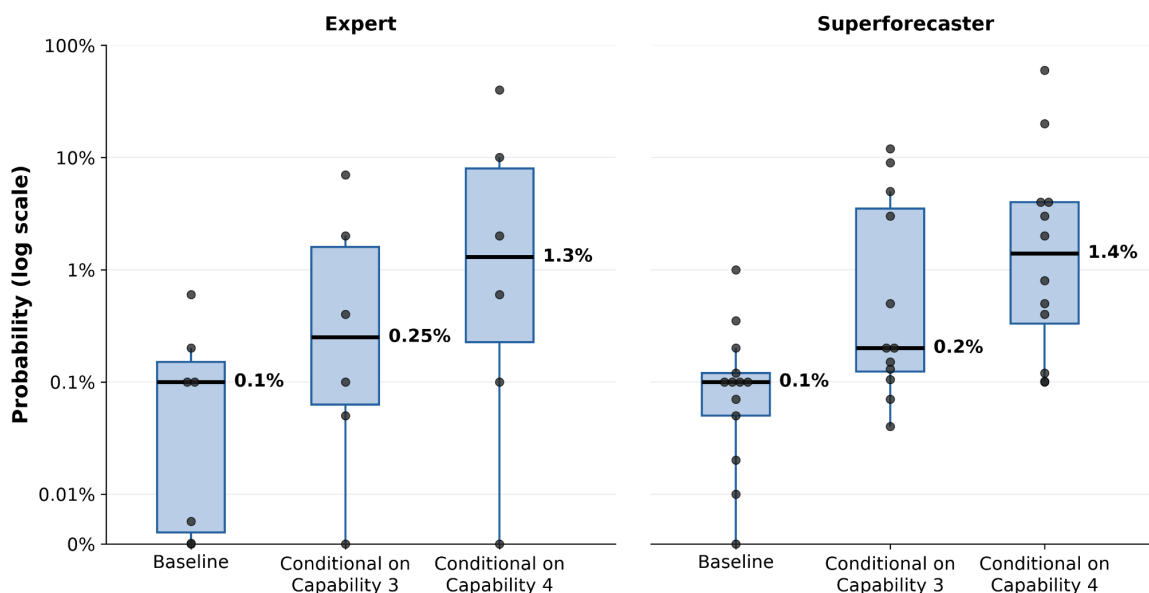


Figure 7.1 | Risk estimates conditional on hypothetical evidence. *Note: Estimated probability of a cyberattack against the US grid causing a blackout with at least \$100 billion in economic damages in 2026. Boxes span the interquartile range of estimates within each group; whiskers extend to the furthest estimate within 1.5× the interquartile range; black bars indicate medians. Source: [Ceppas de Castro et al. \(2026\)](#).*

Although the hypothetical evidence concerned TA1/TA2 uplift, respondents updated their overall, actor-agnostic risk estimate. This suggests that respondents treated real-world evidence of TA1/TA2 uplift as evidence about the broader threat model.¹³⁴

These results suggest that even sophisticated pre-deployment evaluations may provide weak signals of increased risk compared to in-the-wild monitoring. This is consistent with catastrophic grid cyberattacks being bottlenecked more by sustained operational integration than by isolated technical skill.

The results also suggest that real-world AI-enabled OT incidents, even if they fall significantly short of the \$100 billion threshold, provide stronger evidence about catastrophic risk than controlled evaluations. This is in part because they reflect the interaction between offensive capabilities and deployed defenses in practice. Indeed, there are reasons to expect smaller-scale events to precede catastrophic attacks. A \$100 billion grid cyberattack would

¹³⁴ This should not be interpreted as implying that respondents view lower-skilled actors as representing a large fraction of marginal risk. It could instead reflect, for example, the view that capabilities enabling TA1/TA2 uplift are also useful to higher-skilled actors, or that a warning-shot attack would indicate that the barriers to catastrophic grid cyberattacks are lower than previously assumed.

require coordinated effects across many targets, whereas less ambitious attacks would be considerably easier. Before observing TA1/TA2 actors achieving anything close to this catastrophic scenario, evidence of more limited successes seems likely: AI-assisted intrusions into OT-adjacent environments, smaller-scale grid disruptions, or physical damage to critical equipment. Were AI to begin lowering the barriers to OT cyberattacks for lower-skilled actors, one might also expect to see an uptick in failed or partially successful attempts, which would provide evidence that relevant barriers are being lowered.¹³⁵ A 2026 intrusion affecting a Mexican water utility might constitute an early, limited example of this sort of evidence. Analysis suggests a threat actor used AI assistance to identify OT-adjacent assets and potential routes from IT to OT ([Dragos 2026a](#); [Gambit 2026](#)).¹³⁶ However, the adversary does not appear to have successfully accessed OT environments or caused any OT disruption.

Other sources of evidence could inform whether AI systems are approaching the capabilities relevant to this threat model. These differ both in the strength of evidence they can provide and in whether they bear more strongly on the TA1/TA2 or TA4/TA5 uplift channels.

Recent work on autonomous cyber time horizons may be particularly relevant to assessing TA1/TA2 uplift. The UK AI Security Institute measures the length of offensive cyber tasks that AI systems can complete autonomously, benchmarked against the time taken by skilled humans.¹³⁷ As of May 2026, AISI estimates that the length of such tasks that models complete successfully at an 80% rate has doubled approximately every five months since late 2024. And recent models are beginning to saturate AISI's narrow cyber task suite, including some tasks that would take human experts several hours. This is suggestive of rapid progress in the kinds of autonomous cyber capabilities that would be necessary to increase the capabilities of TA1/TA2 actors.

However, this benchmark remains an imperfect proxy for the capabilities most relevant to grid cyberattacks. It uses a relatively narrow suite of self-contained tasks, which likely does not capture much of the difficulty of attacks against defended real-world targets. It also does not appear to capture many of the OT-specific bottlenecks emphasized above. This and similar benchmarks are therefore useful for tracking the trajectory of autonomous offensive cyber capabilities but provide weaker evidence about whether AI systems could automate multi-target OT operations of the scale envisaged in the central \$100 billion grid cyberattack scenario.

OT-specific evaluations could provide useful, but still limited, evidence for TA4/TA5 uplift. Such evaluations could test whether models reduce the effort required to test and validate

¹³⁵ Historical OT cyberattacks have often involved mistakes or unintended consequences. For example, [Slowik \(2022c\)](#) points out that none of Russia's grid cyberattacks appear to have worked as planned.

¹³⁶ Dragos's overall assessment is that "current AI models do not provide novel ICS or OT-specific capabilities, yet can make OT more visible to adversaries already operating inside IT environments" ([Dragos 2026a](#)).

¹³⁷ [UK AISI 2026](#).

physical effects. They could also test the effort required to understand unfamiliar and poorly documented OT environments. The strength of this evidence would depend on whether the capabilities tested mapped onto concrete bottlenecks facing such actors. As with other narrow capability evaluations, they would provide only partial evidence of the capabilities required to accelerate real-world attacks against defended targets in realistic environments. [Table 7.1](#) summarizes how the author would interpret these different lines of evidence.

Type of evidence	Primary relevance	What it would show	Key limitations
Generic cyber evaluations (e.g. CyBench)	Weak evidence for TA1/TA2 and TA4/TA5 uplift	Whether models possess narrow cyber skills	Not OT-specific; low operational realism
OT-specific evaluations	Some evidence for TA4/TA5 uplift; weaker evidence for TA1/TA2 uplift	Whether models possess skills specific to OT attacks. Degree of relevance depends on mapping to concrete bottlenecks	Low operational realism; limited evidence for end-to-end automation; may poorly measure human+AI collaboration
Autonomous cyber time horizons (e.g. UK AISI 2026)	Better evidence for long-horizon autonomous cyber capabilities	Extent to which models can autonomously complete increasingly long cyber tasks	Currently not OT-specific; limited operational realism; inherits limitations of underlying task set
Realistic OT uplift studies	Strong evidence for TA4/TA5 partial uplift	Extent to which models reduce time, staffing, or expertise required for capable actors to execute realistic OT operations	Expensive; long serial time; difficult to make realistic
Real-world AI-assisted OT incidents (e.g. Dragos 2026a)	Strongest evidence that AI is lowering real-world barriers	Whether AI is lowering barriers in operational settings	Rare and lagging; attribution and granular attack details may be unavailable

Table 7.1 | Interpreting different lines of evidence.

7.3 Mitigations

The case for safety mitigations depends on whether AI increases the risk of catastrophic grid cyberattacks enough to justify their costs. Absent stronger evidence that AI is eroding OT-specific bottlenecks, substantially shifting the strategic logic facing state actors, or enabling lower-skilled actors to conduct realistic OT operations, costly measures targeted specifically at the \$100 billion grid scenario appear hard to justify on this basis alone.

This conclusion is limited to the threat model analyzed here – a US grid cyberattack causing at least \$100 billion in economic damages. Considerations beyond the scope of this report, such as geopolitical consequences, non-economic harms from prolonged power outages, or correlations with other AI-enabled cyber and national-security risks, could strengthen the case for mitigations.

Moreover, lower-cost mitigations such as monitoring for AI-enabled OT attacks, refusal training, and output filtering are likely worthwhile, especially insofar as they address a broader range of cyber misuse risks. More costly mitigations – such as delaying model deployment or implementing stringent security measures to prevent theft of model weights – are unlikely to be justified on the basis of this threat model alone, absent much stronger evidence that AI is lowering relevant bottlenecks.

An important factor limiting the benefits of model-level mitigations is the proliferation of AI capabilities. At present, open-weight models lag frontier models by around four months ([Epoch 2026](#)). Given this dynamic, model-level mitigations may primarily buy time for defensive adaptation, rather than provide durable risk reduction.¹³⁸

Finally, many of the most robust mitigations for this threat model may fall outside the traditional scope of AI safety. Two key factors limiting damages from blackouts are the US grid's resilience to isolated failures – owing to decentralized generation capacity and redundancy in the transmission system – and the speed at which operators can restore power following large outages. Investments in grid resilience may therefore provide greater benefit than mitigations targeted narrowly at preventing intrusions. These could include increasing stockpiles of transformers and critical OT devices, expanding black-start capability – the ability to re-energize parts of the grid without external power – and improving operator readiness for manual power restoration. These investments may also increase resilience to kinetic attacks, severe weather, and equipment failures. They are, moreover, consistent with a strategy of deterrence-by-denial that reduces the expected benefits to adversaries of attacks by limiting achievable damages.

7.4 Summary of Implications

Taken together, these considerations suggest that the \$100 billion grid cyberattack scenario is unlikely to provide a strong standalone basis for near-term safety mitigations. The lower-skilled actor pathway would be highly consequential, but appears to require extremely broad and advanced AI capabilities. The state-actor pathway may become relevant much earlier, but has a less straightforward relationship to overall risk. Instead, this threat model is useful for understanding the extreme end of AI-enabled cyber misuse risk, and – in the near term – for surfacing early-warning signs for a wider class of high-consequence, AI-enabled OT cyberattacks. Such warning signs include evidence from real-world AI-assisted attacks, uplift studies in realistic OT operational environments, or evidence of AI beginning to lower key Stage 2 bottlenecks to OT attacks. The next section discusses the main limitations and uncertainties of this report's analysis.

¹³⁸ See [Toner \(2025\)](#).

8. Limitations and Uncertainties

Before concluding, it is important to highlight several limitations and uncertainties that could qualify this report's findings. These fall into three broad categories: the scope of the analysis, the evidence base and methodology, and the modeling assumptions.

8.1 Scope

The analysis is centered on the \$100 billion threshold for catastrophic damages used in OpenAI's Preparedness Framework, which provides a concrete benchmark but remains somewhat arbitrary. As discussed in [Section 1.3](#), a lower threshold of \$1 billion has since been incorporated into state-level legislation and at least one AI company's safety framework. Different stakeholders may have different views on what level of harm warrants ex ante mitigations, and the report's use of this threshold should not be taken as an endorsement of this or any other threshold. While the broad conclusion that catastrophic grid cyberattacks face substantial barriers likely holds across a range of thresholds, the capability requirements will be meaningfully reduced for lower thresholds, as illustrated by the brief analysis of \$10 billion attacks. This is especially important for this report's policy conclusion. The claim that this threat model is a weak standalone basis for costly near-term mitigations applies most directly to the \$100 billion threshold and may not hold for lower-damage scenarios.

Several other scoping choices constrain the generalizability of the conclusions. The report focuses exclusively on direct economic damages, omitting welfare costs, health impacts, and geopolitical consequences that could dominate total social cost. It examines a relatively narrow threat model, which likely represents only a small portion of the broader AI-enabled cyber risk landscape. And it focuses on the US grid, which has distinctive features – decentralization of generation and control and extensive redundancy – that limit the conclusions that can be drawn about grids in other countries or other critical infrastructure.

8.2 Evidence Base and Methodology

This report's analysis relies heavily on the track record of well-documented historical cyberattacks, which provides some of the most direct evidence available for what complex real-world operations require. This record has inherent limitations. It is selective, consisting almost entirely of state-linked operations that later became public, and therefore overrepresents operations that failed or were unusually visible. It is also lagging. Stuxnet remains the most technically impressive OT attack despite being planned almost two decades ago, and attacker and defender capabilities have likely evolved substantially since Russia's 2015 and 2016 grid attacks. And because states typically prioritize espionage over disruption, it sheds limited light on the upper bound of destructive potential. These limitations introduce

significant uncertainty into the assessments of baseline capabilities and of the difficulty of catastrophic attacks. Moreover, the most directly relevant case studies – Russia’s attacks against Ukraine’s grid – may transfer imperfectly to the US context, given Ukraine’s heightened operator readiness and experience with manual restoration.

A further limitation is that the most relevant state capabilities are likely classified. The public record may therefore understate what TA4/TA5 actors can already do, while also providing limited evidence about the conditions under which they would use such capabilities.

The evidence base also limits the granularity of this report’s capability assessments. It can ground rough estimates of operational timelines, but lacks the resolution to decompose attacks into specific tasks. The track record alone would not necessarily reveal narrow technical bottlenecks even if they existed. The conclusion about the breadth of capability requirements therefore draws on additional evidence: input from domain experts, the broader cybersecurity literature, and the observation that certain narrow OT capabilities are already accessible to relatively low-resource actors.

To further supplement the track record, we conducted a survey of domain experts and superforecasters, but the findings of the survey are limited by a small sample (8 experts, 13 superforecasters), uncertain transferability of forecaster calibration to questions lacking historical precedent, and substantial divergence in baseline beliefs, especially around TA4/TA5 capability.¹³⁹ Several surveyed experts had also provided comments on earlier drafts of this report, which may have anchored their judgments.

8.3 Modeling Assumptions

The analysis relies on a stylized distinction between attackers’ capability and conditional willingness. This is useful for organizing the two uplift pathways, but these properties likely interact in practice. AI could affect willingness as well as capability: by reducing perceived attribution risk, lowering the cost of failed attempts, increasing confidence in catastrophic effects, or making attacks appear less escalatory. These effects are difficult to model and are especially important for state actors, where strategic constraints may be at least as important as capability bottlenecks. As a result, the analysis of partial uplift for state-level actors is less developed than the analysis of lower-skilled actor uplift, and it remains highly uncertain how such uplift would change overall risk.

The estimates of lower-skilled actor willingness are fairly uncertain. Physical attacks and threats against grid infrastructure provide evidence of willingness to cause disruption, but do not directly establish a desire to cause catastrophic-scale harm or to pursue long and complex campaigns.

¹³⁹ Similar disagreement was found in [Singer et al.’s \(2023\)](#) survey of 18 grid security experts on four grid cyberattack threats.

Relatedly, the threat actor categorization used in the survey and risk model is necessarily stylized. Real-world threat actors span a continuum rather than falling into discrete classes (see [Section 6.1](#)).

As noted in [Section 6](#), this report's uplift analysis focuses exclusively on how AI could lower barriers for attackers, without modeling corresponding effects on defense. Many of the capabilities discussed in this report could improve defenders' ability to harden their systems, discover and patch vulnerabilities, monitor for intrusions, and respond to incidents. This may be particularly relevant for catastrophic grid cyberattacks, where the difficulty of sustaining disruption depends heavily on defenders' ability to detect intrusions and react swiftly to isolate compromised systems. The net effect of AI progress on risk to the grid will depend on the relative improvements in offensive and defensive capabilities, and the pace of adoption by the relevant actors.

[Section 3](#)'s damage estimates are extrapolated from a handful of major historical blackouts, none of which involved the scope and duration of outage envisaged in the \$100 billion scenario. The analysis assumes roughly linear scaling of damages with scope and duration. This may not hold for extreme outages. Damages could be sublinear in the first hours as backup systems absorb impact, then turn superlinear after a week as those systems fail and supply chains collapse.

The analysis assumes utilities can largely restore power independently, implying that sustained nationwide disruption would require interfering with dozens of largely independent restoration efforts. The precise level of independence will depend on the degree of overlap in control systems and communication infrastructure. Similarly, it is assumed that OT environments in the US grid remain heterogeneous. However, consolidation among IT/OT vendors and grid standardization may have reduced this heterogeneity.

Finally, the analysis is centered on two illustrative attack pathways, which do not exhaust the possible routes to large-scale grid disruption. Supply chain compromises are considered as a cross-cutting access mechanism ([Section 4.3](#)), but other potential vectors remain unexamined. These include demand-side attacks, hybrid cyber-kinetic operations, and combined attacks on the grid and other infrastructure networks such as telecommunications. Two alternative pathways appear particularly worthy of further investigation. First, bricking operationally critical OT devices at scale may offer a route to severe disruption with meaningfully lower barriers than physical destruction of generators or transformers, especially if failures overwhelm stockpiles of spare equipment. Second, grid modernization may introduce new correlated vulnerabilities through which relatively localized faults or compromises could cascade into large-scale disruption. Both pathways could reduce the number or diversity of targets an attacker would need to affect to cause large-scale outages.

9. Conclusion

This report finds that achieving \$100 billion in economic damages from a cyberattack on the US grid would require a blackout orders of magnitude larger than any cyber-induced blackout to date. Under the two attack pathways analyzed, this would require either physically destroying dozens or more critical grid assets or precipitating a large cascading blackout and preventing restoration for roughly a week.

A key finding is that sustaining a widespread blackout for roughly a week is a qualitatively different challenge from precipitating one. Triggering a cascading outage might be achievable via attacks on a handful of critical nodes. However, absent physical damage to equipment, power is typically restored within hours or a few days. Sustaining a large blackout for a week is a dramatically harder problem, requiring either extensive physical damage or prolonged interference with dozens of largely independent efforts to restore power.

The baseline risk of such an attack appears low in the near term. In the survey of domain experts and superforecasters, the median estimate of a US grid cyberattack causing at least \$100 billion in economic damages in 2026 was 0.1% in both groups, and 1% for a \$10 billion attack.

The AI uplift analysis centered on two distinct channels. The clearest route would be AI systems enabling lower-skilled but more willing actors to execute grid cyberattacks that currently appear to require state-level operational capacity. This would require AI systems capable of substituting for most of the work involved in complex OT attacks: reconnaissance, intrusion, target-specific OT reasoning, developing and testing tailored payloads, stealth, persistence, and coordination across many targets. This is a very high threshold.

A second route is partially uplifting state-level actors. This could arrive earlier, since partial automation of individual attack stages could provide meaningful uplift by reducing the cost, time, labor, or detection risk of operations these actors could plausibly already conduct. However, it is much less clear that uplifting these actors would translate into significant marginal risk. State-level actors appear significantly constrained in their willingness to carry out cyberattacks on the grid, due to the risk of attribution and escalation, and the availability of more effective kinetic options during active interstate conflict. AI capabilities that relieve these strategic constraints may be particularly significant from a risk perspective.

This uplift analysis holds defensive capabilities fixed, and increasing AI capabilities would plausibly also strengthen grid defense and recovery. The report therefore does not determine the overall effect on risk to the grid from more capable AI systems, which will depend on the balance of gains to attackers and defenders and the relative pace of adoption.

The \$100 billion grid cyberattack scenario is useful for understanding the more extreme end of AI-enabled OT cyber capability. However, these findings suggest it may not be well-suited for grounding costly near-term safety mitigations. By the time AI systems could fully uplift lower-skilled actors enough to significantly increase risk, they would likely pose serious risks across many other domains. More decision-relevant near-term triggers for mitigations may come from narrower cyber threat models, lower-damage grid scenarios, or evidence that AI is beginning to erode OT-specific bottlenecks.

For risk monitoring, we argue that real-world evidence of AI-assisted OT incidents may be more informative than narrow pre-deployment evaluations. Realistic OT uplift studies and evidence of AI assistance with Stage 2 tasks – such as interpreting OT systems, validating physical and process effects, maintaining stealth, or coordinating multi-target operations – could also be important warning signs.

These conclusions remain uncertain. The public record of OT cyberattacks and state capabilities remains limited, and there may be alternative attack pathways with meaningfully lower barriers. These uncertainties qualify this report's conclusions but do not overturn the central picture. The alternative pathways discussed in [Section 8.3](#) remain speculative and would likely face many of the same challenges of scale and coordination. For AI risk assessment, the catastrophic grid cyberattack threat model may be more useful for understanding the upper end of AI-enabled cyber capability in complex OT environments than for grounding near-term safety decisions.

About the Author



Matthew van der Merwe ✉

Research Fellow, GovAI

Matthew works on threat modeling, with a particular focus on CBRN and cyber misuse risk. Before joining GovAI he worked at the University of Oxford's Future of Humanity Institute. Previously, he worked as an equities analyst, researching cement companies. He holds a BA in Philosophy from the University of Cambridge.

Appendices

Appendix 1 | How the Grid Works

The power grid consists of the following basic components:

- **Generators** (e.g. power plants, solar farms) produce electricity.
- **Transmission networks** carry electricity at high voltages over long distances from generators to distribution networks, often via overhead lines on tall pylons.
- **Distribution networks** carry electricity from transmission networks to end customers, at lower voltages and over shorter distances, often on utility poles.¹⁴⁰
- **Substations** are junctions in the grid where voltage can be altered by running electricity through transformers, and where operators can control the flow of electricity by opening and closing circuit breakers.

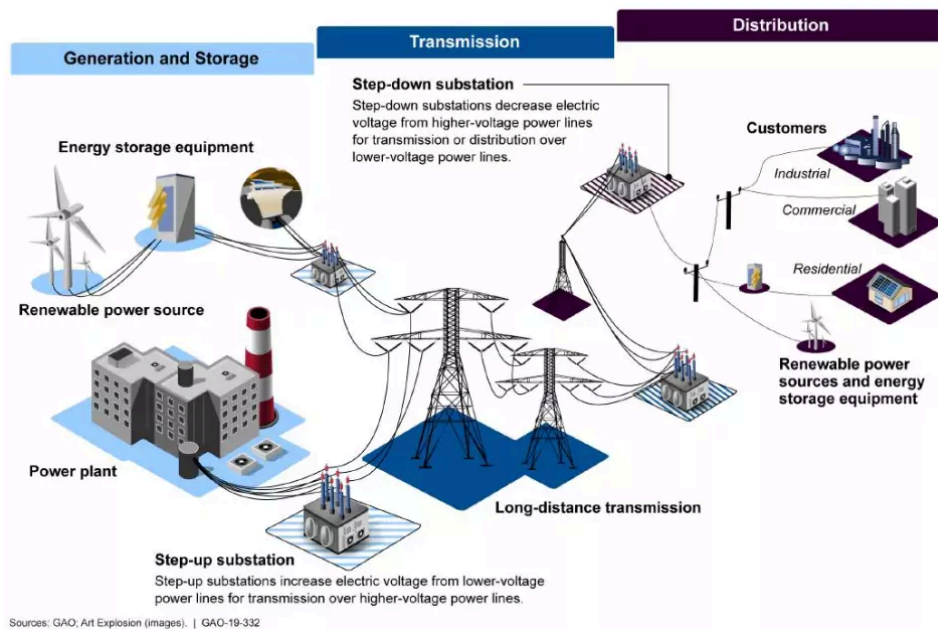


Figure A1 | The basic structure of the power grid. Source: [GAO \(2019\)](#)

IT vs. OT Systems

It is useful to distinguish between two types of computer systems used in grid operations, and other infrastructure and industrial processes:

¹⁴⁰ I find an analogy to road networks helpful – transmission lines are like the interstate highways where vehicles can travel fast over longer distances; distribution is the denser network of local roads where vehicles travel shorter distances at slower speeds to reach their final destinations.

- **IT systems:** The familiar systems that handle many business operations such as billing, email, and administration, and are typically connected to the internet.
- **OT systems:** The “broad range of programmable systems or devices that interact with the physical environment or manage devices that interact with the physical environment” ([Dragos 2024b, 3](#)). In grid operations these are the systems with which operators monitor and control equipment like generators, substation breakers, and transformers. They include:
 - **SCADA (supervisory control and data acquisition):** The basic monitoring and control infrastructure grid operators rely on to receive real-time data from the grid and to send commands to equipment.¹⁴¹
 - **Safety and protection systems:** Digital relays and other equipment that detect and automatically respond to unsafe conditions.

The important point here is that grid cyberattacks will likely involve compromising OT environments, since this is a precondition for interfering directly with grid behavior. OT environments present some additional challenges:

- They typically are, or should be, segmented from IT and internet-facing networks.¹⁴²
- They use more niche software and protocols requiring more specialized knowledge. OT devices may rely on relatively bespoke configurations for a given environment. These features pose obstacles to potential attackers, though these may be partially offset by increasing commonality in the IT/OT equipment used by grid operators.¹⁴³

How Is the Grid Controlled?

The grid is operated from several types of control centers with different scopes:

- Individual power plants have their own control rooms with fine-grained control over the on-site generators, e.g. increasing and decreasing output and disconnecting them from the grid.
- Utilities operate control centers that monitor their portions of the transmission and distribution networks with direct control over substation breakers, and more limited control over generation plants in their network.¹⁴⁴

¹⁴¹ Also relevant are energy management systems (EMS) – more advanced analytics tools built on top of SCADA that help operators to understand and manage the grid, e.g. to identify if lines are at risk of being overloaded and reroute electricity to prevent damage by opening and closing breakers.

¹⁴² See [Garton \(2019\)](#). This may create correlated vulnerabilities: whereby vulnerabilities in a widely deployed device or system could be reused across multiple operators. Nonetheless, there will likely still be target-specific differences in configuration, network architecture, and operational context.

¹⁴³ ([Lee 2019](#)). For example, the SCADA/EMS equipment market appears to be dominated by a few major suppliers (Siemens, GE Vernova, ABB).

¹⁴⁴ One expert reported that utility control centers’ Automatic Generation Control (AGC) systems can increase and decrease generation at power plants within set limits, but cannot normally switch units on and off, and that utility control centers typically have no control over nuclear plants.

- At the highest level, reliability operators run control centers that monitor and coordinate power flow at a regional level across multiple utilities' transmission networks, but they generally rely on utility control centers to execute the actual switching operations needed to maintain stability across regions.

Appendix 2 | Estimating Damages from Ukraine Cyber Blackouts

I take two approaches to estimating the damages from the 2015 grid attack.

- Method 1:
 - The outage lasted ~3.5 hours and affected 500k people ([INL 2020, 5](#)).
 - This represents ~0.04% of a calendar year,¹⁴⁵ and ~1.2% of Ukraine's 2015 population.¹⁴⁶
 - Ukraine's 2015 GDP was \$91 billion (current \$).¹⁴⁷
 - Assuming 100% economic output loss during the blackout in the affected areas (and no further impacts), this would represent a GDP impact of $[0.04\% \times 1.2\%] \times [\text{\$91 billion}] = \text{\$0.4m}$.
- Method 2:
 - I index on the 2003 Northeast Blackout, which – compared to the 2015 Ukraine attack – affected approximately 100x more people (50m vs. 0.5m). The time to fully restore power was ~12.3x longer (~43 hours vs. 3.5 hours).
 - Therefore the 2015 Ukraine blackout was approximately 1,200x smaller magnitude, in terms of scope x duration.
 - The economic damages per unit disruption will depend primarily on the relative sizes of the affected economies. In current \$, 2003 US GDP per capita was \$39.5k ([World Bank 2026b](#)), which is 18.8x higher than 2015 Ukraine GDP per capita of \$2.1k ([World Bank 2026a](#)).
 - So the 2015 Ukraine blackout was ~1,200x smaller magnitude and is assumed to have caused ~19x less damages per unit disruption – amounting to roughly 23,000x lower economic damages.
 - Economic damages from the 2003 Northeast Blackout were around \$13 billion.¹⁴⁸
 - This suggests that economic damages from the 2015 Ukraine blackout were ~\$0.6m (\$2025).

¹⁴⁵ A 3.5-hour outage represents approximately 0.04% of a year: $3.5 \div (365 \times 24) = 0.04\%$.

¹⁴⁶ See footnote to [Table 2.2](#) in main report. Approximately 0.5 million people were affected, or about 1.2% of Ukraine's population of approximately 41 million ([UA Gov 2023](#)).

¹⁴⁷ Ukraine's 2015 GDP was approximately [~\\$91 billion at current prices](#).

¹⁴⁸ See footnote to [Table 2.3](#).

- These are rough and uncertain estimates, but they do converge. Altogether the author's best guess is \$0.2m to \$2m (mean=\$0.6m).

2016 attack:

- Compared to the 2015 attack, the outage was ~2.8x shorter ([INL 2020, 5–6](#)).
- All sources suggest it was of smaller scope than 2015, affecting at most 225,000 customers (or 0.5 million people). Precise figures for the affected population could not be found.
- All else equal, this suggests the magnitude of the outage was >3x smaller.
- Notably, the 2016 blackout occurred around midnight on a weekend ([Motherboard 2016](#)) whereas 2015 was mid-afternoon on a weekday ([Zetter 2016](#)), which would likely have resulted in lower damages due to lower levels of economic activity.
 - e.g. [Leahy and Tol \(2011\)](#) estimate weekend night-time outages, in Northern Ireland, are ~5.5x less costly than weekday daytime outages.¹⁴⁹
 - Given the uncertainty, the author's best guess for this factor is 3–7x.
- Combining this with the magnitude difference, and rounding, implies economic damages from the 2016 Blackout are ~10–20x lower than 2015.
- Altogether, I estimate damages of \$10,000 to \$200,000 (mean=\$50,000) from the 2016 attack.

April 2022 attack:

- No outage resulted.

October 2022 attack:

- Unclear because scope or duration have not been reported.

Beyond the economic damages of the blackouts, a more comprehensive accounting of the social cost of these attacks could incorporate factors like:

- Remediation costs and longer-term disruption to the targeted utility companies from wiping systems and bricking OT devices.
- Spending on Ukrainian cyberdefense prompted by Russian cyberattacks.
- Impacts on Ukrainian national security, in the context of their conflict with Russia, which had escalated in 2014 following Russia's invasion and annexation of Crimea.

Appendix 3 | European Blackout Data

The damage estimate in [Section 3](#) relies exclusively on major US outages with available damage estimates. To understand whether these cases are representative of the relationship between outage scope and duration, [Stankovski et al. \(2023\)](#)'s dataset of European power

¹⁴⁹ See [Leahy & Tol \(2011\)](#) Table A2: system-wide VOLL for weekday day is 11 EUR, vs. 2 EUR for weekend night.

outages was reviewed.¹⁵⁰ [Figure A3](#) plots these outages by number of people affected, duration, and cause.

The data is consistent with the patterns described in [Section 3](#). Outages affecting large populations are typically short, while long-duration outages tend to be relatively localized. Outages caused by weather events are typically longer than those from other causes. This analysis does not directly inform Section 3's estimate but provides some further evidence that a blackout affecting tens or hundreds of millions of people for several days would sit far outside the observed distribution.

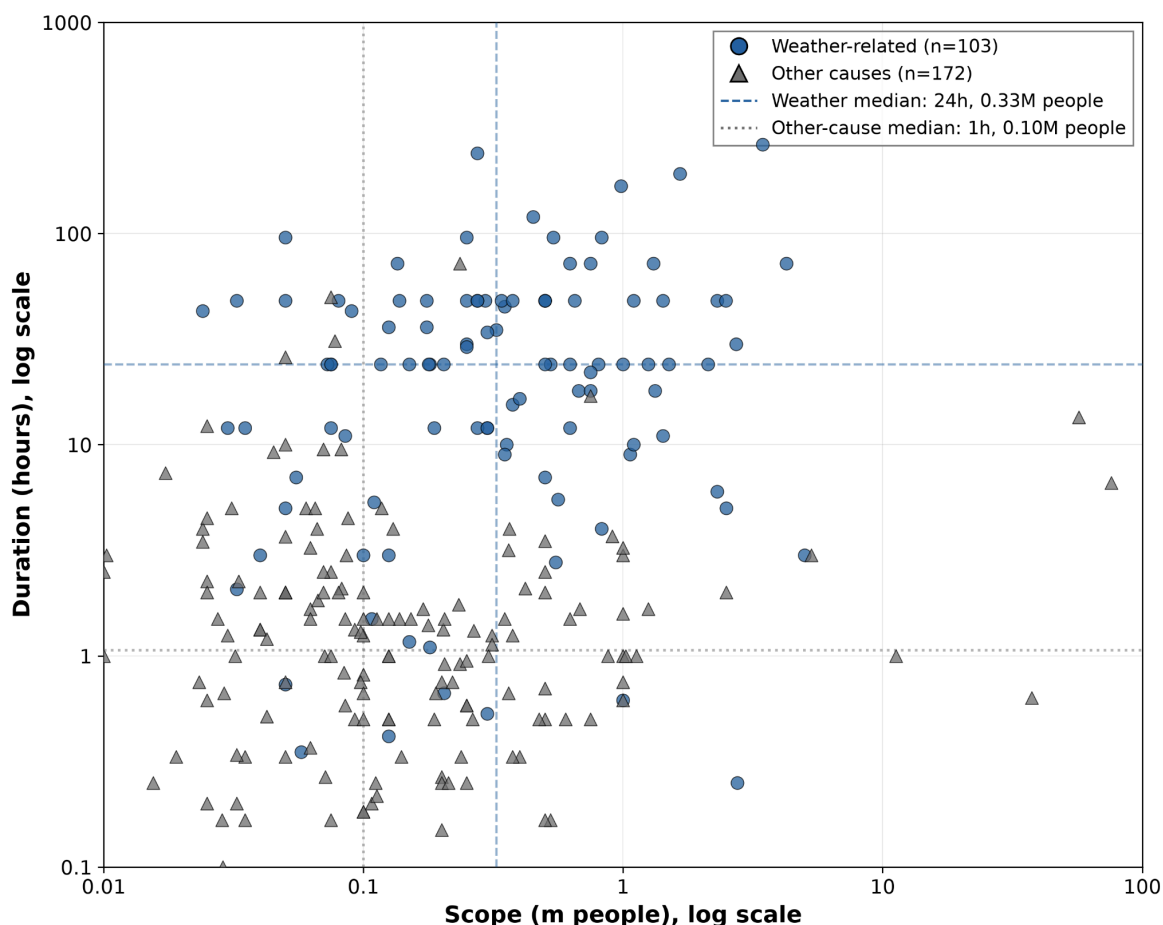


Figure A3 | European outages by scope and duration. Source: Data from [Stankovski et al. \(2023\)](#)

Appendix 4 | Methodology for VOLL-based Estimate

The value of lost load (VOLL) differs substantially between places, between different types of end-users, at different times of day, and for different durations of outage. For the back of the

¹⁵⁰ See [spreadsheet](#) for raw data.

envelope calculation, a recent estimate for system-wide VOLL for Texas for long-duration outages was used:

- [Gibbons and Sergici \(2024, 52\)](#) estimate a VOLL for 16 hour outages (the longest duration given) at \$13.6/kWh, with a 95% confidence interval of [\$10.6–18.7]

This was modeled using [Squiggle](#):

- The model fits their 95% confidence range to a lognormal distribution, with mean of ~\$14.5/kWh.
- Given this, the total power loss equivalent to \$100 billion in damages is estimated as 5.4–9.4 TWh (mean: 7.1 TWh).
- Total annual US electricity consumption is ~4,070 TWh ([EIA 2023](#)). So this power loss is equivalent to a blackout spanning the whole US lasting 0.5–0.8 outage-days (mean: 0.6),¹⁵¹ or an outage spanning 100 million people and lasting 1.6 to 2.9 outage-days.¹⁵²

Converting these “outage-days” figures into terms that are more easily comparable to historical outages requires additional assumptions about recovery schedules. I hope that readers who have made it this far will permit a brief digression into measuring outage duration.

A Note on Outage-Days

Power restoration after blackouts is a gradual process, rather than a period of total blackout followed by instantaneous recovery. The calculation above yielded an estimate in what I have called “outage-days”. One outage-day is equivalent to 100% of customers in the affected area being without power for 24 hours.

For example, consider an illustrative blackout, where power is restored linearly, with 50% restored after 1.5 days, 90% restored after 2.7 days, and 100% after 3 days (Figure A4.1).

¹⁵¹ Calculation: Dividing 5.4–9.4 TWh by average daily US electricity consumption (4,070 TWh ÷ 365 days) yields 0.48–0.84 outage-days.

¹⁵² Calculation: Dividing the above figure by (100m/US population=0.3) for a duration of 1.6 to 2.9 outage-days.

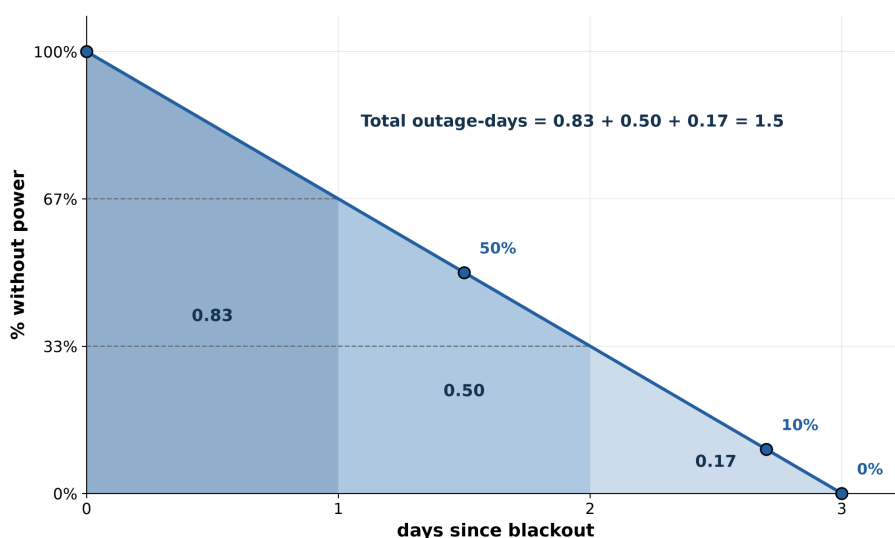


Figure A4.1 | Outage-day calculation example.

The outage-days correspond to the area under the curve in the above chart:

- In day 1, an average of 83% of customers are without power = 0.83 outage-days
- In day 2, an average of 50% of customers are without power = 0.5 outage-days
- In day 3, an average of 17% of customers are without power = 0.17 outage-days

So altogether this hypothetical blackout lasting 3 calendar days from start-to-finish, amounts to 1.5 outage-days.

This illustrates a simple rule. Assuming a linear recovery schedule, the total number of outage-days is half the time to restore 100% of power (or, equivalently, the time to restore 50% of power). Since this report has primarily been measuring blackout duration as the time to restore 90% of power, the more useful rule of thumb is that, assuming linear recovery, the time to restore 90% of power is 9/5 the number of outage-days (and 9/5 the time to restore 50% of power).

In reality, power restoration is typically slower than linear. (For example, see the restoration following Hurricane Sandy in Figure A4.2). And so, to convert the outage-day figure from our VOLL estimate into the time it takes to restore 90% of power, assumptions about the shape of the restoration curve are required.

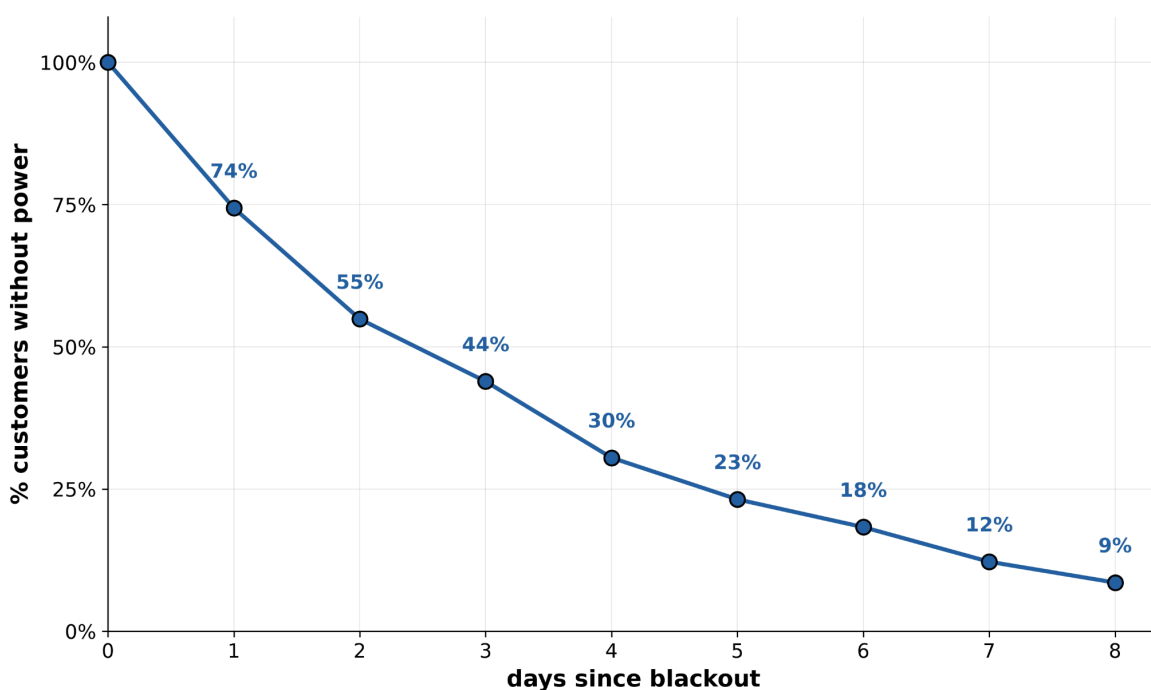


Figure A4.2 | Power restoration following Hurricane Sandy. Source: Reconstructed from [DOE \(2013, 11\)](#).

Finalizing the Estimate

To estimate the shape of the restoration curve, I ran a brief analysis of the recovery schedules of the historical blackouts mentioned in [Section 2](#) (excluding the 2021 Texas Power Crisis, since this was a rolling blackout):

- For the 2003 Northeast Blackout, one can infer from the table in [ICF \(2003, 2\)](#) that 50% load was restored after 14 hours, and ~89% load was restored after ~32 hours, a ratio of 2.3x.
- For Hurricane Sandy, one can infer from the chart in [DOE \(2013, 11\)](#) that ~50% customers were restored after ~2.5 days and ~91% after 8 days, a ratio of 3.2x.

These data points enable a rough assumption about the recovery schedule:

- The time taken to restore 90% of power will be roughly 2.3–3.2x the time to restore 50% (and therefore, 2.3–3.2x the total number of outage-days). Since this analysis uses only two data points, I treat this range as a 50% confidence interval fitted to a log-normal distribution (mean: 2.8; 90% CI: 1.8–4.1).

Above, it was estimated that \$100 billion in damages would require a blackout equivalent to 1.6 to 2.9 outage-days for 100 million people.

Using the above recovery schedule assumption, we can now work out how long this blackout would actually last in terms of the time taken to restore 90% of power. Multiplying the outage days [90%: 1.6 to 2.9] by the recovery parameter [90%: 1.8 to 4.1], suggests the time taken to restore 90% of power would be ~6 days [90%: 3.5 to 9.5].

In summary, the VOLL-based estimate finds that a blackout affecting 100 million people would have to last 3.5 to 9.5 days (mean: 6 days).

For this calculation, [Squiggle](#) was used to work with distributions rather than point estimates for the uncertain parameters. The model can be viewed [here](#).

Appendix 5 | Historical OT Cyberattacks

Attack/Actor	Calendar time	Resourcing	Other hard steps	Effects
Aurora test (2007) ¹⁵³ <i>Idaho National Laboratory</i>	Unclear	\$3–5m budget, ~30 staff	- Novel attack chain - Test environment	Destruction of a small diesel generator
Stuxnet (2009–10) ¹⁵⁴ <i>US/Israel</i>	≤4.5 years	Budget: \$0.3–2 billion (malware: \$10–20 million). 20–50 technical staff	- Multiple zero days - Human intelligence ¹⁵⁵ - Novel attack chains - Mock centrifuge cascade	Damage to several hundred centrifuges; ~3-month delay to enrichment program (Slayton 2017).
German steel mill (2014) ¹⁵⁶ <i>Russian APT</i>	Unclear (≥4 year campaign)	Unclear	- Multi-year reconnaissance - ICS know-how - Sophisticated intrusions	Physical damage to blast furnace (likely unintentional)
Ukraine grid (2015) ¹⁵⁷ <i>Russian APT</i>	~18 months	3–20 technical staff	- ICS know-how and testing - Multi-year reconnaissance ¹⁵⁸ - Coordinated impacts on 3 targets (E-ISAC 2016, 5)	Blackout: 3.5h outage for ~0.5m

¹⁵³ **Time:** Internal budget request ([Aurora FOIA 2015, 50](#)) dated February 13, 2007, Demo was March 4, 2007 ([Aurora FOIA 2015, 62](#)). **Budget:** Est. \$2.9m (or \$2.2m without contingency). Inflation-adjusted via [US Inflation Calculator](#) to \$3.4–4.5m ([Aurora FOIA 2015, 57](#)). **Staff:** 31 “test team participants” listed ([Aurora FOIA 2015, 108–110](#))

¹⁵⁴ See [Appendix 6](#).

¹⁵⁵ [Volkskrant \(2024\)](#) suggests a Dutch engineer was recruited by US intelligence to provide physical access to Natanz’s air-gapped networks, which he infected via a USB stick or a water pump he installed. He died in a motorcycle accident shortly after leaving Iran in January 2009.

¹⁵⁶ **Time:** [Slowik \(2021, 3\)](#) dates the start of the campaign to no later than 2010. **Effects:** [Slowik \(2021, 8\)](#) describes the event as “almost certainly unintended”.

¹⁵⁷ See [Appendix 6](#).

¹⁵⁸ ([E-ISAC 2016, 8](#)): “the strongest capability of the attackers was not in their choice of tools or in their expertise, but in their capability to perform long-term reconnaissance operations required to learn the environment and execute a highly synchronized, multistage, multisite attack.”

Ukraine grid (2016) ¹⁵⁹ <i>Russian APT</i>	~12 months	3–20 technical staff	(see 2015 attack)	Blackout: ~1.3h outage for <0.5m
Triton (2017) ¹⁶⁰ <i>Russian APT</i>	3–36 months	1–10 technical staff	- ICS know-how - Multi-year reconnaissance	2x unplanned shutdowns of petrochemical plant
Ukraine grid (2022a) ¹⁶¹ <i>Russian APT</i>	1.5 months	3–20 technical staff	(see 2015 attack)	Attack thwarted
Iranian steel plant (2022) ¹⁶² <i>Israel-linked actor</i>	Unclear	Unknown	Unclear	Serious malfunction and fire

Table A5 | Detail on significant historical OT cyberattacks.

Appendix 6 | Resource Requirements for Prior OT Cyberattacks

The person-years of skilled labor required for prior OT cyberattacks are estimated as follows:

Ukraine Grid Cyberattacks

Technical headcount:

- Eight individuals were indicted by the US Department of Justice in relation to Sandworm-linked activities. Three members of Unit 74455 (the GRU [Russian military intelligence] unit within which Sandworm operates) as part of a 2018 indictment ([DOJ 2018](#)) spanning two hacking groups, and five additional members as part of a 2020 indictment ([DOJ 2020](#)).
- [Molfar \(2023\)](#) name 11 employees of Unit 74455. [Bellingcat \(2020\)](#) found an additional 46 people linked to an address used by Unit 74455 and reported that “it is highly likely that the majority of these 49 people are directly associated with the GRU themselves”. Some of these will likely be non-technical staff, and some technical staff might not be involved in Sandworm activities.
- Altogether, I estimate the total headcount of the hacking group is 10–50, of whom 20% to 60% are in technical roles.

Time spent:

- [Maschmeyer \(2021\)](#) estimates 19 months calendar time preparation for the 2015 attack, and an additional 12 months for the 2016 attack.

¹⁵⁹ See [Appendix 6](#).
¹⁶⁰ See [Appendix 6](#).
¹⁶¹ **Time:** Attack scheduled for April 8, 2022. Initial compromise was 17 February, 2022 or earlier ([Lipovsky and Cherepanov 2021, minute 20:20](#)). Excluding preparation for 2015 and 2016 attacks.
¹⁶² [BBC \(2022\)](#) on the attack and initial attribution. See [Times of Israel \(2022\)](#) on link to state operations, and [NYT \(2021\)](#) on US officials attributing a previous attack by the group to Israel.

- I estimate that technical staff spent 10% to 60% of this calendar time on preparing for grid attacks.

Taking these uncertain estimates as inputs and using [Squiggle](#) to calculate the person-years of technical labor, the model estimates that:

- 2015 attack took ~4 person-years effort, with 90% CI [0.7–11]
- 2016 attack took ~2.6 person-years effort, with 90% CI [0.5–7]

Stuxnet

Headcount:

- [De Falco \(2012, 25\)](#) estimates ~45 technical and intelligence staff. [Gostev \(2011\)](#) estimates 22–30 technical staff.
- Combining and slightly expanding these ranges, the author's 90% CI is 20–50 technical staff.

Time spent:

- The attack was discovered in June 2010 ([Fildes 2010](#)).
- Calendar time preparation:
 - Upper bound of 4.5 years, since V0.5 of Stuxnet worm in development no later than November 2005 ([Ars Technica 2013](#)).
 - Lower bound of 0.5 years via an expert quoted in [Hesseldahl \(2010\)](#).
- I estimate that technical staff spent roughly 20–100% of this calendar time on the attack.

Using [Squiggle](#) to calculate the person-years technical labor with these uncertain estimates, the model finds that Stuxnet took ~30 person-years effort, with 90% CI [5–85]. This aligns with Microsoft's estimate of 27 man-years, quoted in [De Falco \(2012, 25\)](#).

Costs:

- **Malware:** The cost to develop the Stuxnet malware has been estimated on the order of \$10 million ([Halstead n.d.](#)). Costin Raiu estimates \$20 million ([SecurityWeek 2019](#)).
- **Total budget:** [Volkskrant \(2024\)](#) cites a figure of \$1–2 billion for the total operation, from the CIA director at the time (via Dutch intelligence). [Slayton \(2017\)](#) estimates ~\$300 million cost to the US and Israel. Most of this is intelligence (\$240 million). The more relevant total is the non-intelligence costs of ~\$50 million (exploits, cyberoperative labor, and infrastructure).
- **Physical testing:** A significant portion of the total budget would have likely been the cost of building a mock centrifuge cascade for testing the malware. Langner estimates: "well over 50 percent of Stuxnet's development cost went into efforts to hide the attack, with the bulk of that cost dedicated to the overpressure attack ... at the cost of

having to build a fully functional mockup IR-1 centrifuge cascade operating with real uranium hexafluoride” ([Langner 2013a](#)).

Triton

This estimate is particularly uncertain and heavily reliant on subjective judgment.

- The attack has been linked to a research institute in Russia, the State Research Center of the Russian Federation Central Scientific Research Institute of Chemistry and Mechanics (TsNIIKhM).¹⁶³
- In 2022, US DOJ [indicted one TsNIIKhM employee](#) in connection with the Triton campaign.
- This APT is also known as XENOTIME and Temp.Veles ([CSA p.7](#)).

Headcount:

- Estimates of this APT’s size could not be found, so the above estimate for Sandworm’s technical headcount (~3 to 20) is used as an imperfect proxy.
- This group is estimated to have roughly 20–80% as many technical staff as Sandworm. The author’s 90% confidence interval is [1 to 10], with a mean of 4.

Time:

- Calendar time preparation between 3 months and 3 years:
 - The first intrusion may have been in 2014 (e.g. see [Slowik \(2019c, 21\)](#). The latest possible start date is May 2017 ([Slowik 2022d, 6](#)).
- I estimate that XENOTIME technical staff spent 30% to 90% of this calendar time on Triton.

Using [Squiggle](#) to calculate the person-years of technical labor from these uncertain parameters, I estimate that Triton took ~2.5 person-years effort, with 90% CI [0.3–8].

Appendix 7 | Threat Actor Classification Methodology

Our survey and uplift analysis rely on a classification of threat actors, which is explained here in greater detail. The analysis uses operational capacity levels from [RAND \(2024, 9–10\)](#), which defines five levels of cyberattack operational capacity [OC] in terms of the resources and capabilities available to the operation, ranging from OC1 to OC5 operations.¹⁶⁴ Based on this, five threat actor categories, ranging from TA1 to TA5, are defined in terms of the highest OC operation they are able to carry out. States are treated as single or unitary threat actors, rather than classifying each state-backed team as an individual actor. In this respect, this report uses the term “threat actor” differently from the wider cybersecurity literature. For

¹⁶³ See [Slowik \(2022d\)](#) for detailed analysis.

¹⁶⁴ By definition, each category includes the capacities of all preceding ones. For example, the most competent nation-states, such as the US and China, are able to carry out OC5 operations, but also all operations below that level.

ease of analysis, it is assumed that each threat actor in a threat actor class has the same capability level. For example, China and the US are TA5 actors, so they are assumed to have the same capability level.

Appendix 8 | Additional Survey Results

This section reports additional results from the survey conducted with the Forecasting Research Institute. See [Ceppas de Castro et al. \(2026\)](#) for a comprehensive overview of the survey results, including some questions not reported here.

Estimates for \$10 Billion Damage Threshold

The survey conducted with FRI also asked respondents to estimate baseline capability and willingness for grid cyberattacks at a lower damage threshold of \$10 billion in economic damages.

Threat actor	Capability		Conditional willingness	
	Probability a given actor is capable of launching a \$10 billion grid cyberattack with six months of effort		Annual probability at least one actor spends six or more months actively attempting the attack, if capable	
	Experts	Forecasters	Experts	Forecasters
TA1	0%	0%	90%	67%
TA2	0%	0%	80%	60%
TA3	0.15%	2%	25%	50%
TA4	2.5%	20%	15%	5%
TA5	25%	65%	5%	5%

Table A8 | Results of the survey for the \$10 billion threshold.

Capability 3: TA1 actor uplifted by AI to match a TA3 actor in an OT-specific CTF-style competition.

Capability 4: a US grid cyberattack causing a $\geq$ \$100 million blackout, attributed to a TA1/TA2 actor with AI assistance.

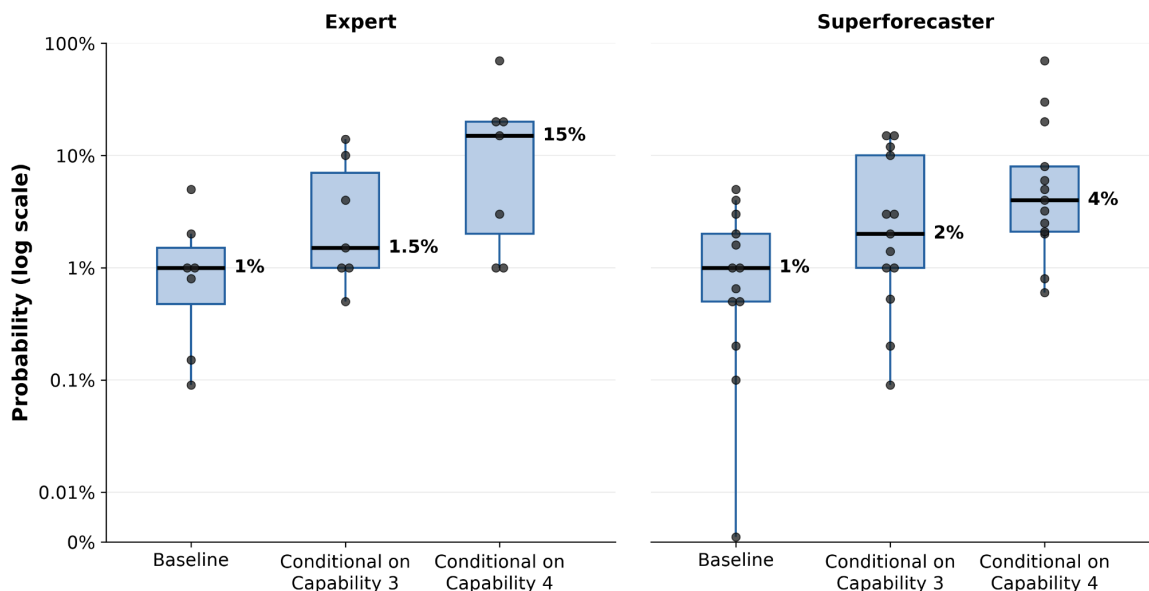


Figure A8.1 | Risk estimates conditional on hypothetical evidence. *Note: Estimated probability of a cyberattack against the US grid causing a blackout with at least \$10 billion in economic damages in 2026. Boxes span the interquartile range of estimates within each group; whiskers extend to the furthest estimate within 1.5× the interquartile range; black bars indicate medians. Source: [Ceppas de Castro et al. \(2026\)](#)*

Conditional Attribution to Threat Actor Classes

One survey question asked respondents to estimate the probability that a major grid cyberattack was caused by different classes of threat actor, conditional on such an attack having occurred in 2026. At both \$100 billion and \$10 billion thresholds, respondents placed a significant majority of probability mass on state-level actors (TA4 and TA5).

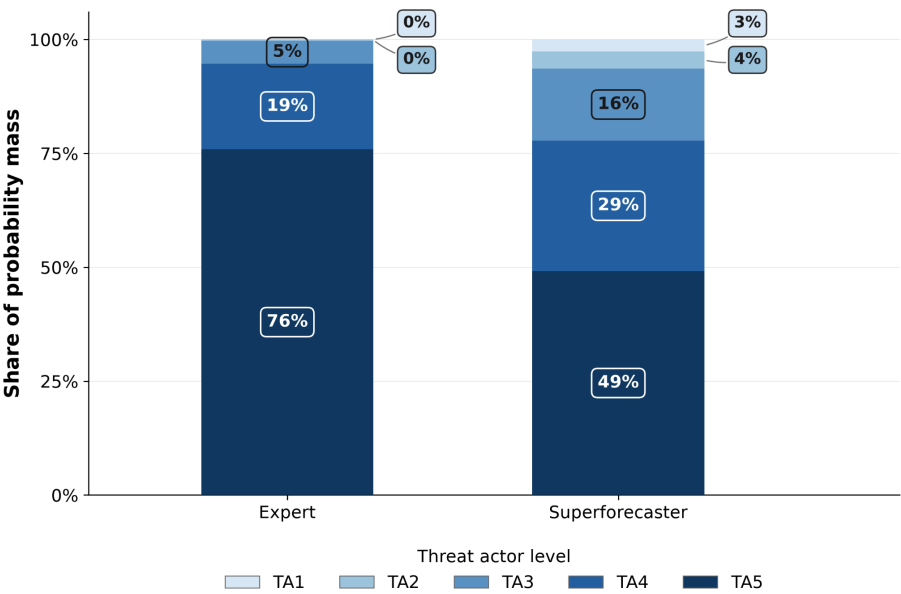


Figure A8.2 | Conditional attribution of a \$100 billion grid cyberattack. *Note: conditional on a cyberattack causing a blackout with at least \$100 billion in economic damages in 2026, respondents’ estimated probability that it was caused by an actor of each threat actor level.* Source: [Ceppas de Castro et al. \(2026\)](#)

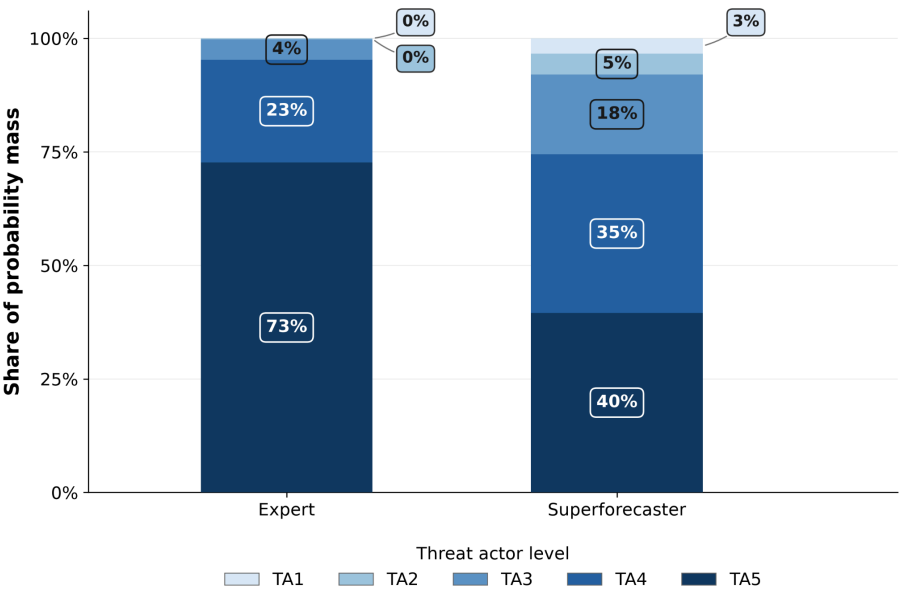


Figure A8.3 | Conditional attribution of a \$10 billion grid cyberattack. *Note: conditional on a cyberattack causing a blackout with at least \$10 billion in economic damages in 2026, respondents’ estimated probability that it was caused by an actor of each threat actor level.* Source: [Ceppas de Castro et al. \(2026\)](#)

Appendix 9 | Kinetic vs. Cyber Grid Attacks

A [2022 attack on a substation in Moore County, North Carolina](#) provides an interesting data point, which is more relevant for TA1–2 actors. On the evening of 3 December 2022, an attacker used a high-powered rifle to shoot at two substations 10 miles apart in Moore County, causing a localized power outage lasting several days ([The Pilot 2023](#)). The attack appears to have been motivated by objections to a local drag show that coincided with the time of the attack.¹⁶⁵

I estimate that the Moore County attack achieved greater impact than the Russian cyberattacks on the Ukrainian grid in terms of customer-hours of outage.¹⁶⁶ In terms of resource inputs, the Moore County attackers used a rifle and a minivan, and investigations have centered on 1–2 culprits ([The Pilot 2023](#)). Assuming the attack was indeed drag show–motivated, and therefore that preparations began only after the show’s announcement,¹⁶⁷ time investment can be bounded at <0.1 person-years.¹⁶⁸

This suggests the Moore County attacks achieved a greater effect than the 2015 and 2016 Ukraine grid cyberattacks, for perhaps 1/100th the person-years effort.

Comparison of Russia’s Kinetic and Cyberattacks against Ukraine’s Grid

Drawing on analysis by [Bateman \(2022\)](#), the Ukraine war can be used to compare the effects of kinetic vs. cyberattacks on the grid. As seen above, the effects of Russia’s cyberattacks on the actual delivery of power have been fairly modest – 3 localized blackouts lasting for a few hours.

Russia’s kinetic attacks have done dramatically more damage to Ukraine’s grid. Electricity demand appears to have fallen ~30% from 2021 to 2024.¹⁶⁹ Rolling blackouts appear very common, with 1,951 hours of scheduled outages in 2024 (~20% of the year).¹⁷⁰ As of late 2024, missile attacks had damaged or destroyed most of the country’s power plants and half of the transmission substations.

¹⁶⁵ The attack coincided with a controversial drag show in the area ([The Pilot 2023](#)), and the person of interest was linked to a local activist group opposed to the drag show. The drag show was announced on 18 November 2022 ([Sandhills Pride, and Sunrise Theater 2022](#)), suggesting that, if drag show–motivated, preparation time was no more than 16 days.

¹⁶⁶ **Moore County outages:** See [spreadsheet](#) for calculation using outage information from [Wikipedia \(n.d.\)](#).

Russia’s 2015 and 2016 cyberattacks: 2015: 0.225m customers for 3.5h; 2016: ≤0.225m customers for 1.25h; $0.225\text{m} \times 3.5\text{h} + 0.225\text{m} \times 1.25\text{h} \approx 1\text{m}$ customer-hours of outage.

Economic damages per customer-hour of outage are likely significantly higher for the Moore County attack, given the difference in US/Ukraine GDP per capita.

¹⁶⁷ Announced on 18 November 2022 ([Sandhills Pride, and Sunrise Theater 2022](#)).

¹⁶⁸ Using the figure of 16 days calendar preparation time, 1 to 2 people, and assuming 100% of calendar time spent on preparation, yields an estimated ~0.06 person-years effort [90%: 0.04, 0.09]. ([Squiggle](#))

¹⁶⁹ From [ACAPS \(2024, 2\)](#). Figure 1 – chart is labeled “electricity consumption” but it looks like either average or peak demand since the units are GW. Per [Chestney \(2024\)](#), 2024 winter peak demand was 18 GW.

¹⁷⁰ [Dixi Group \(2025\)](#) using data from [Energy Map \(2025\)](#). 1,951 hours is “total of scheduled stabilization outages for households introduced in all or the vast majority of Ukrainian regions (over 50%)”

	Impact of cyberattacks	Impact of kinetic attacks
Generation	None	As of December 2024, Russia has damaged or destroyed all of Ukraine’s thermal and hydropower plants (Harmash 2024).
Transmission	2016 attack opened a breaker in a single transmission substation.	“45% of controlled transmission substations were destroyed or damaged.” (IEC 2023, 12)
Distribution	2015 attack opened breakers in ~50 distribution substations, and bricked some substation OT equipment, causing lasting disruption. ¹⁷¹	“As of the beginning of January 2023, more than one thousand overhead lines (6-150 kV) and more than eight thousand transformers (6-150 kV) were damaged or disconnected” (IEC 2023, 13)
Overall economic damages	\$0.2–2 million ¹⁷²	“[T]otal damage to Ukraine’s energy sector exceeds \$56 billion, including \$16 billion in direct physical destruction and over \$40 billion in indirect financial losses” (Harmash 2024)
Overall outages	<10 hours of outages, each for <1 million people. ¹⁷³	~2,000 hours outage (in 2024), each for >19m people. ¹⁷⁴

Table A9 | Comparing impacts of Russia’s cyber and kinetic attacks on Ukraine. *Note: This table is adapted from one presented in [Bateman \(2022\)](#).*

Appendix 10 | Other Attack Pathways Not Considered in Depth

This report has not considered all possible routes to causing catastrophic outages via cyberattack. Other routes include:

- Demand-side attacks that cause a cascading blackout by manipulating power consumption, causing grid instability – e.g. turning every internet-connected AC unit to max in conditions of tight supply/demand balance.¹⁷⁵
 - These are just other routes to precipitating a cascading blackout, and do not themselves provide a route to sustaining a cascading blackout, which might limit the potential damages.
- Forced oscillations that trigger a disturbance in power flow that can propagate quickly through the grid and damage sensitive equipment (discussed in more depth in [Tomes 2024](#)).

¹⁷¹ [Slowik \(2022b\)](#): “anecdotal evidence indicates the system wiping and effective destruction of serial-to-ethernet converters produced damage taking years to effectively correct.”
¹⁷² See [Appendix 2](#).
¹⁷³ See [Appendix 2](#).
¹⁷⁴ [Dixi Group \(2025\)](#) using data from [Energy Map \(2025\)](#). 1,951 hours is “total of scheduled stabilization outages for households introduced in all or the vast majority of Ukrainian regions (over 50%)”. Ukraine’s 2024 population was 38m ([World Bank n.d.](#)) so we roughly estimate that each outage affects at least 19m people.
¹⁷⁵ See e.g. [Ars Technica \(2025\)](#), [Soltan et al. \(2018\)](#) for discussion.

- This is not a different pathway so much as an alternative route to causing physical damage to critical grid equipment (and there may be others), though importantly it may represent a pathway to damaging multiple pieces of equipment via a single effect.
- Mixed cyber and kinetic attacks, e.g. hardware supply chain attacks – placing explosive implants into grid equipment before installation to be triggered remotely to cause physical damage (e.g. Israel’s Hezbollah pager attack).¹⁷⁶ Other approaches could combine a physical attack on generators or transformers with a cyberattack on control centers.
 - This pathway was set aside in part because kinetic attacks appear considerably more effective at damaging grids than cyberattacks, which suggests hybrid attacks may be concerning primarily for their kinetic elements.

Appendix 11 | Grid Cyberattack Case Studies

This appendix provides extensive methodology and evidence supporting [Table 2.1](#) documenting historical grid cyberattacks.¹⁷⁷

Methodology

To create the comprehensive list of historical grid cyberattacks in [Table 2.1](#), we reviewed existing datasets and scholarly lists of electric grid-related cyberattacks and checked them against open sources from 1999 to 2024. We reviewed all entries in CSIS’s Significant Cyber Incidents ([2026](#)) and CFR’s Cyber Operations Tracker (2026), the subset of entries classified as “Energy” in ICS Strive’s Threat Report appendix ([Machtemes et al. 2025](#)) and web database ([n.d.](#)), and entries from academic and government lists of ICS cyber incidents, such as [Miller et al. \(2021\)](#) and [Canadian Centre for Cyber Security \(2021\)](#). We uncovered some additional incidents through a general search of open sources, including news articles and government filings.¹⁷⁸

[Table 2.1](#) included incidents that were:

- (a) Cyber incidents (e.g. conducted through software),
- (b) Perpetrated maliciously (e.g. for ransom, for sabotage),
- (c) By actors of any type (e.g. ransomware groups, hackers, state actors),
- (d) Against any part of the electrical grid (e.g. generators, substations, control centers),
- (e) Attested credibly to have affected the normal functioning of the electrical grid, including:
 - (i) Power outages, both:
 - (1) Directly, via cyber manipulation of operational technology (OT),

¹⁷⁶ On the pager attack see [Washington Post \(2024\)](#).

¹⁷⁷ I am grateful for Connor Aidan Stewart Hunter’s efforts in preparing [Table 2.1](#) and this section.

¹⁷⁸ The 2025 Poland attack was added after the initial construction of this dataset.

- (2) Indirectly, via cyber manipulation of information technology (IT) which the flow of power is dependent on,
 - (ii) Physical effects on grid assets,
 - (iii) Manipulating industrial control systems (ICS), even where there is no physical damage,
 - (iv) Disruption to operators' view and/or control of grid assets, including indirectly by rendering grid operator computers unusable,
 - (v) Delaying restart of generators,
 - (vi) Delaying restoration of power during an accidental or natural power outage,
 - (vii) Inducing the victim to voluntarily incur any of the above effects out of "an abundance of caution".

Incidents that were excluded were:

- (a) Not cyber incidents (e.g. physical attacks against the electrical grid),
- (b) Caused by non-malicious software bugs or human errors via software (we consider these separately as "accidents" elsewhere in the report),
- (c) Against the energy sector, but not the electrical grid (e.g. pipelines, wind turbine manufacturers, street lamps, EV charging stations, ICS providers)
- (d) Not credibly attested,
- (e) Resulted in effects or access merely to IT, such as exfiltration of data, wiping of data, infiltrating IT networks but not successfully proceeding to OT networks.

Classification

Outcomes

Incidents were classified into "outcomes" of the following types. Outcome categories are based on observed grid effect, not inferred attacker intent:

- (a) Outage via OT: An incident where a cyberattack causes a power outage by manipulating operational technology. This classification is modified from ICS Strive's classificatory scheme for OT incidents in general, which includes "Direct, on OT" for incidents where a cyberattack directly manipulates OT through OT software.
- (b) Incidental outage via IT: An incident where a cyberattack incidentally causes a power outage because power distribution is dependent on an IT system which was rendered inoperable. This classification is from ICS Strive's classificatory scheme which includes "Indirect, IT dependency" and this report's list fully overlaps with theirs for both "Energy" incidents of this type.
- (c) Process compromise: An incident where a cyberattack impacted grid control systems, including direct manipulation of safety monitoring or supervisory systems, installation of malware on the OT network (even absent a physical effect).

- (d) Loss of view/control: An incident where grid operators lost either view or digital control or both of grid assets due to a cyberattack, but did not suffer a process compromise. This includes loss of view/control caused by a cyberattack against upstream service providers (such as telecommunications satellites), and out of “an abundance of caution”.
- (e) Operational impact: An incident where a cyberattack substantially impacted the normal functioning of the grid through harassment of grid operators without causing a loss of view/control, a process compromise, or an incidental outage via IT.

Actors

Incidents are classified as attributed to “actor” types as follows:

- (a) State: An incident credibly attested to have been perpetrated by a state or state-sponsored actor, including the military, the intelligence community, or government-contracted entities.
- (b) Non-state: An incident credibly attested to an actor that is not a state nor state-sponsored, such as cybercrime groups.
- (c) Unknown: An incident with no attribution, or conflicting and insufficiently conclusive attributions.

Evidence

Ukraine	Oct 2022	State	Outage via OT	Outage in one city coinciding with missile strike (Mandiant 2023)
---------	----------	-------	---------------	---

Attribution: Sandworm ([Mandiant 2023](#)). The US Cybersecurity and Infrastructure Security Agency (CISA) ([CISA 2022a](#)) identifies Sandworm as a Russian state-sponsored cyber actor under the Russian military intelligence unit known as the GRU.

Evidence: The only public source on this cyberattack is from Mandiant. In their report, Mandiant attributes the cyberattack to Sandworm. “In late 2022, Mandiant responded to a disruptive cyber physical incident in which the Russia-linked threat actor Sandworm targeted a Ukrainian critical infrastructure organization.” ([Mandiant 2023](#))

Ukraine	Dec 2016	State	Outage via OT	1.25 hour blackout across northern Kyiv (INL 2020 ; Polityuk 2016)
---------	----------	-------	---------------	--

Attribution: Sandworm ([CISA 2022a](#); [Hultquist 2016](#))

Evidence: CISA and Mandiant both attribute the 2016 Ukraine grid cyberattack to Sandworm (part of GTsST): “GTsST actors conducted a cyberattack against Ukrainian energy distribution companies in December 2015 ... In 2016, GTsST actors conducted a cyber-intrusion campaign against a Ukrainian electrical transmission company and deployed CrashOverride malware (also known as Industroyer) specifically designed to attack power grids” ([CISA 2022a](#)). The SANS ICS blog confirms conclusions previously reached by iSIGHT regarding the nature of the Ukrainian attacks (specifically the role of destructive malware and phone disruption) and attribution to Sandworm Team ([Hultquist 2016](#)).

Ukraine	Dec 2015	State	Outage via OT	Blackout for 225k customers for 3.5h (INL 2020)
---------	----------	-------	---------------	---

Attribution: Sandworm ([CISA 2022a](#); [Hultquist 2016](#))

Evidence: See above entry Ukraine 2016.

Ghana	2022	Unknown	Incidental outage via IT	“5+ days of power outages” (Machtemes et al. 2025)
-------	------	---------	--------------------------	--

Attribution: Unknown ([Machtemes et al. 2025](#))

Evidence: News reporting on this incident does not identify a specific threat actor. This OT cyberattack is cataloged by ICS Strive ([Machtemes et al. 2025](#)) as perpetrated by an “unknown” threat actor.

Justification: Threat actor is classified as unknown by a major OT cyberattack dataset.

South Africa	2019	Unknown	Incidental outage via IT	250k customers lost power, delayed restoration (Machtemes et al. 2025)
--------------	------	---------	--------------------------	--

Attribution: Unknown ([Machtemes et al. 2025](#))

Evidence: News reporting on this incident does not identify a specific threat actor. This OT cyberattack is catalogued by ICS Strive ([Machtemes et al. 2025](#)) as perpetrated by an “unknown” threat actor.

Justification: Threat actor is classified as unknown by a major OT cyberattack dataset.

Poland	2025	State	Process compromise	OT devices bricked; no outage (Midnight Blue 2026 ; MITRE 2025)
--------	------	-------	--------------------	---

Attribution: Russian state; specific group contested – Sandworm/ELECTRUM (GRU) or Dragonfly/Berserk Bear (FSB) ([MITRE 2025](#); [FCDO 2026](#)).

Evidence: The UK and EU issued an official attribution to Russia’s FSB (Centre 16) ([FCDO 2026](#)), consistent with some prior reporting ([CERT Polska 2026](#)). Other prior reporting attributed the destructive wiper activity to the GRU-linked ELECTRUM (also tracked as Sandworm) with medium confidence ([Dragos 2026b](#)). The actor is classified as “State”, since both candidate attributions are to Russian state-sponsored groups.

Ukraine	Apr 2022	State	Process compromise	Malware in OT network; foiled attack (ESET 2022 ; O’Neill 2022)
---------	----------	-------	--------------------	---

Attribution: Sandworm ([CERT-UA 2022](#); [ESET 2022](#); [O’Neill 2022](#); [Greenberg 2022](#); [Vicens 2022](#))

Evidence: Ukraine’s state-run Computer Emergency Response Team (CERT) attributed the largely foiled cyberattack to Russia ([CERT-UA 2022](#)). The cybersecurity firm ESET confirmed that “Sandworm attackers made an attempt to deploy Industroyer2 against high-voltage electrical substations in Ukraine.” ([ESET 2022](#))

Notes: An initial Ukrainian government report claimed that 9 substations were briefly offline (without causing a power outage), affirmed by multiple news organizations ([O’Neill 2022](#); [Greenberg 2022](#)). However, this finding was later retracted by the Ukrainian government as “preliminary” and a “mistake” ([O’Neill 2022](#); [Vicens 2022](#)), replaced with the ambiguous claim that the government was “aware about only one substation (with up to nine outgoing transmission lines) among which only one line may have witnessed minute idle period, potentially caused by fault tolerance mechanism” ([Vicens 2022](#)). This claim is ambiguous because it is not clear if the statement means that one substation witnessed an idle period due to a fault tolerance mechanism which was triggered because of the cyberattack, or that this was a natural non-cyberattack explanation for a true event which was confabulated in the earlier report as larger (9 substations instead of 1) and due to a cyberattack. Given the uncertainty of the 9 substation claim, as well as the 1 substation claim, this entry reflects merely the presence of malware in the ICS network, credibly and unambiguously claimed by ESET.

US	2014	Unknown	Process compromise	Utility’s control system software accessed (ICS-CERT 2014)
----	------	---------	--------------------	--

Attribution: Unknown

Evidence: A “sophisticated threat actor” compromised an internet-accessible control system network for a public utility, protected by a simple password that was brute-forced ([ICS-CERT 2014](#)).

US	2013	State	Process compromise	Dam’s SCADA accessed during maintenance (DOJ 2016)
----	------	-------	--------------------	--

Attribution: ITSecTeam, contracted by Islamic Revolutionary Guard Corps ([DOJ 2016](#))

Evidence: US prosecutors accused Hamid Firoozi of obtaining remote access to a computer controlling the SCADA systems of a New York state dam ([DOJ 2016, 14](#)). Hamid Firoozi is identified as one of a group of “experienced computer hackers” working for a private security company (non-state actor) based in Iran, which performed work for the Islamic Revolutionary Guard Corps (IRGC; a state actor).

Justification: This event is at the edge of the distinction between state and non-state actors. The perpetrator of this cyberattack was Hamid Firoozi, an experienced computer hacker working for a private company. Nonetheless, these activities were reportedly sponsored by the Iranian state.

US	2012	Unknown	Process compromise	Malware delayed plant restart for 3 weeks (ICS-CERT 2012)
----	------	---------	--------------------	---

Attribution: Unknown

Evidence: ICS-CERT, which is the only source for this incident, does not attribute it. The malware is identified as a variant of the Mariposa virus ([ICS-CERT 2012](#)).

Brazil	2011	Unknown	Process compromise	Conficker worm downed power plant ICS (Branquinho 2011)
--------	------	---------	--------------------	---

Attribution: Unknown

Evidence: The Canadian Centre for Cyber Security ([2021](#)), a government organization, cites Branquinho and lists this incident as “unattributed”. No attribution is made by Branquinho ([2011](#)). The infecting virus was Conficker, which has been non-conclusively attributed to Ukrainian cybercriminals.

US	2003	Unknown	Process compromise	5h nuclear plant safety system loss by Slammer worm (Poulsen 2003)
----	------	---------	--------------------	--

Attribution: Unknown

Evidence: There is no attribution for the Slammer worm. Litchfield thinks there were at least two authors ([Litchfield 2003](#)) but does not make a full attribution.

UK	1999	Non-state	Process compromise	Guard trips access control/electronic door lockdown (Miller et al. 2021)
----	------	-----------	--------------------	--

Attribution: Security guard ([Miller et al. 2021, 4](#))

Evidence: “A security guard set off a high-level alarm at Bradwell Nuclear Power Plant. The guard reportedly hacked into one of the computers to alter and delete sensitive information. This ‘caused a shutdown of the station’s access control system’ which automatically locked the facilities [sic] electronic doors” ([Miller et al. 2021, 4](#)).

Notes: The apparent intent of this insider threat was the deletion of sensitive information, not causing an effect on operational technology. Nonetheless, their actions shut down the access control system and triggered physical effects on a power plant.

Denmark	2022	Unknown	Loss of view/control	Operator lost visibility into 3 remote assets (SektorCERT 2023)
---------	------	---------	----------------------	---

Attribution: Unknown

Evidence: SektorCERT published a play-by-play sequence of events ([SektorCERT 2023](#)) indicating Sandworm as a possible perpetrator but cautioning that they could not make a confident attribution. The cybersecurity firm Forescout subsequently published an analysis indicating that the attack was perpetrated by multiple threat actors, none of which could be confidently attributed to Sandworm ([Forescout 2024](#)).

Germany	2022	State	Loss of view/control	Loss of control/view of 5,800 wind turbines (Willuhn 2022)
---------	------	-------	----------------------	--

Attribution: Russia ([Blinken 2022](#); [Willuhn 2022](#))

Evidence: A cyberattack on the satellite internet provider Viasat at the outset of Russia’s 2022 invasion of Ukraine disrupted German wind turbines ([Willuhn 2022](#)). The US State Department

and other American security agencies later attributed the attack to Russia and confirmed that the attack had spillover effects on European countries ([Blinken 2022](#); [CISA 2022b](#)).

US	2019	Unknown	Loss of view/control	Brief loss of view/control for remote assets (NERC 2019 ; Sobczak 2019)
----	------	---------	----------------------	---

Attribution: Unknown

Evidence: A “cyber event” was reported to the US Department of Energy by a utility. No attribution is provided across news reporting ([Sobczak 2019](#)), expert commentary ([Alrich 2019](#)), declassified documents ([GWU 2019](#)), or official public NERC documentation.

Notes: Different sources provide different wordings as to the precise nature of the operational technology effect of the cyberattack. [Sobczak \(2019\)](#) reports a denial-of-service attack likely resulted in a “temporary loss of visibility to certain parts of the utility’s supervisory control and data acquisition (SCADA) system”. [NERC \(2019\)](#) reports the incident “resulted in brief communications outages (i.e., less than five minutes) between field devices at sites and between the sites and the control center.”

US	2018	Unknown	Loss of view/control	Control center offline for 12–24h out of caution (Alrich 2020)
----	------	---------	----------------------	--

Attribution: Unknown

Evidence: The only detailed public information about this event is from the cybersecurity expert Tom Alrich, who learned that an electrical grid control center went offline as a precautionary cleanup measure to remove what he infers was ransomware that may or may not have infected the control center’s computers ([Alrich 2020](#)).

Germany	2022	Unknown	Loss of view/control	Cautionary 1–2 day loss of view of wind turbines (DW 2022)
---------	------	---------	----------------------	--

Attribution: Unknown

Evidence: The wind turbine maintenance and repair company, Deutsche Windtechnik, lost digital view of wind turbines due to an unknown threat actor. [DW \(2022\)](#) describes the incident as “a targeted professional cyber attack”. According to [Petkauskas \(2022\)](#), Deutsche

Windtechnik deactivated data monitoring for wind turbines that were switched off, due to security concerns.

Montenegro	2022	Unknown	Loss of view/control	Switched to “manual handling” out of caution (Reuters 2022b)
------------	------	---------	----------------------	--

Attribution: Unknown

Evidence: Multiple attributions have been made for the threat actor that carried out a cyberattack against government organizations in Montenegro in August 2022, with the most recent attribution being to the ransomware group called Cuba Ransomware ([S4P 2023](#); [Reuters 2022a](#); [Byrne 2024](#)).

Norway	2019	Non-state	Loss of view/control	Took “months” to regain full operational capability (Jeffries et al. 2022)
--------	------	-----------	----------------------	--

Attribution: FIN6 ([Jeffries et al. 2022](#))

Evidence: MITRE attributes the incident targeting Norsk Hydro to the cybercriminal group known as FIN6 ([Jeffries et al. 2022](#)).

Ukraine	2017	State	Loss of view/control	Forced to “paper standards” for 10 days (Borys 2017)
---------	------	-------	----------------------	--

Attribution: Russian military/government

Evidence: The US, UK, and Australia, confidently attribute NotPetya to the Russian military and government ([CISA 2018](#); [UK Foreign Office 2018](#); [Taylor 2018](#)).

Notes: NotPetya not only impacted IT functions in Ukraine’s energy sector, but also directly impacted grid operators, with BBC reporting indicating that a “dispatcher” saw black screens ([Borys 2017](#)).

US	2003	Unknown	Loss of view/control	Control center SCADA downed by Slammer worm (Owens 2009)
----	------	---------	----------------------	--

Attribution: Unknown

Evidence: The North American Electric Reliability Council reported ([see original from GWU 2015](#)) the Slammer worm downed one utility’s SCADA network, moving onto the control center’s local area network ([Owens 2009](#)).

US	2003	Unknown	Loss of view/control	Slammer worm downed SCADA traffic via telecom (Poulsen 2003)
----	------	---------	----------------------	--

Attribution: Unknown

Evidence: The North American Electric Reliability Council reported ([see original from GWU 2015](#)) the Slammer worm incidentally downed a power company’s SCADA traffic by infecting a telecommunications company which the SCADA traffic was dependent on ([Poulsen 2003](#)).

UK (BVI)	2024	Unknown	Operational impact	IT attack slowed hurricane blackout recovery (Machtemes et al. 2025)
----------	------	---------	--------------------	--

Attribution: Unknown ([Machtemes et al. 2025](#))

Evidence: ICS Strive, in its 2025 Threat Report, indicates the threat actor as “Unknown”.

Notes: This incident is an edge case, since the ransomware infected IT systems and indirectly hampered restoration of power that was out due to an incidental storm ([Kaspersky 2025](#)).

References

ACAPS. 2024. *UKRAINE Energy infrastructure attacks: outlook and impact during 2024–2025 cold season*. Thematic Report.

https://www.acaps.org/fileadmin/Data_Product/Main_media/20240913_ACAPS_Ukraine_Analysis_hub_Attacks_on_the_energy_infrastructure_in_Ukraine.pdf. Archived March 15, 2026, at the [Wayback Machine](#).

Acharya, Samrat, et al. 2020. “Cybersecurity of Smart Electric Vehicle Charging: A Power Grid Perspective.” *IEEE Access* 8: 214434–214453. <https://doi.org/10.1109/access.2020.3041074>.

Alrich, Tom. 2019. “It’s Official: The Event Reported in March Was a Real Cyber Attack.” <http://tomalrichblog.blogspot.com/2019/09/its-official-event-reported-in-march.html>. Archived July 6, 2026, at the [Wayback Machine](#).

———. 2020. “When Will a Ransomware Attack Impact the Bulk Electric System? 2018.” <http://tomalrichblog.blogspot.com/2020/10/when-will-ransomware-attack-impact-bulk.html>. Archived February 19, 2026, at the [Wayback Machine](#).

———. 2024. “Joe, I suggest the reason two of those items weren’t discussed was they’re imaginary threats.” Comment on Joseph Weiss, “The 2024 RSA Cybersecurity Conference: What Wasn’t Addressed Can Hurt You,” *Energy Central*, May 31, 2024.

<https://www.energycentral.com/intelligent-utility/post/2024-rsa-cybersecurity-conference-what-wasn-t-addressed-can-hurt-you-HKtKlJvCipbBUw0?highlight=6lldZPL1m3B2JUC>. Archived June 24, 2026, at the [Wayback Machine](#).

Ankit, Aman, et al. 2022. “U.S. Resilience to large-scale power outages in 2002–2019.” *Journal of Safety Science and Resilience* 3 (2): 128–135. <https://doi.org/10.1016/j.jnlssr.2022.02.002>.

Anthropic. 2023. *Anthropic's Responsible Scaling Policy, Version 1.0*. Anthropic. <https://www-cdn.anthropic.com/1adf000c8f675958c2ee23805d91aaade1cd4613/responsible-scaling-policy.pdf>. Archived June 16, 2026, at the [Wayback Machine](#).

———. 2025a. *Claude 3.7 Sonnet System Card*. Anthropic. <https://assets.anthropic.com/m/785e231869ea8b3b/original/claude-3-7-sonnet-system-card.pdf>. Archived April 29, 2026, at the [Wayback Machine](#).

———. 2025b. *Responsible Scaling Policy, Version 2.2*. Anthropic. <https://www-cdn.anthropic.com/872c653b2d0501d6ab44cf87f43e1dc4853e4d37.pdf>. Archived June 22, 2026, at the [Wayback Machine](#).

Ars Technica. 2013. “Revealed: Stuxnet “Beta’s” Devious Alternate Attack on Iran Nuke Program.” *Ars Technica*.

<https://arstechnica.com/information-technology/2013/02/new-version-of-stuxnet-sheds-light-on-iran-targeting-cyberweapon/>. Archived March 13, 2026, at the [Wayback Machine](#).

———. 2025. “Researchers Say New Attack Could Take Down the European Power Grid.” *Ars Technica*.

<https://arstechnica.com/security/2025/01/could-hackers-use-new-attack-to-take-down-european-power-grid/>. Archived February 18, 2026, at the [Wayback Machine](#).

- Assante, Michael J., and Robert M. Lee. 2015. *The Industrial Control System Cyber Kill Chain*. SANS Institute.
https://icscsi.org/library/Documents/White_Papers/SANS%20-%20ICS%20Cyber%20Kill%20Chain.pdf. Archived June 24, 2026, at the [Wayback Machine](#).
- Baker, George, et al. 2021. *Large Transformer Criticality, Threats, and Opportunities*. Center for Security Policy.
<https://centerforsecuritypolicy.org/wp-content/uploads/2022/06/LARGE-TRANSFORMER-THREATS-OPPORTUNITIESJCIP-PUBLISHED-VERSION.pdf>. Archived October 29, 2025, at the [Wayback Machine](#).
- Bartley, William H. 2003. *Analysis of Transformer Failures*. IMIA WGP 33 (03). The Hartford Steam Boiler Inspection and Insurance Co.
<https://www.imia.com/wp-content/uploads/2023/07/wgp3303.pdf>. Archived June 24, 2026, at the [Wayback Machine](#).
- Bateman, Jon. 2022. *Russia's Wartime Cyber Operations in Ukraine: Military Impacts, Influences, and Implications*.
<https://www.jstor.org/stable/resrep45856.5?mag=ukraine-russia-and-the-west-a-background-reading-list&seq=1>.
- BBC. 2011. "Cyprus: Navy chief killed by base munitions blast." BBC.
<https://www.bbc.com/news/world-europe-14115103>. Archived June 24, 2026, at the [Wayback Machine](#).
- . 2022. "Did hackers start this steel factory fire in Iran?" BBC.
<https://www.bbc.com/news/av/technology-62099474>. Archived June 24, 2026, at the [Wayback Machine](#).
- . 2025. "Sri Lanka: Minister blames monkey for nationwide power cut." BBC.
<https://www.bbc.co.uk/news/articles/c8d92n28pqjo>. Archived August 16, 2025, at the [Wayback Machine](#).
- Bellingcat. 2020. "Russian Vehicle Registration Leak Reveals Additional GRU Hackers." Bellingcat.
<https://www.bellingcat.com/news/uk-and-europe/2020/10/22/russian-vehicle-registration-leak-reveals-additional-gru-hackers/>. Archived May 29, 2026, at the [Wayback Machine](#).
- Bhatt, Manish, et al. 2024. "CyberSecEval 2: A Wide-Ranging Cybersecurity Evaluation Suite for Large Language Models." *arXiv preprint*. <https://doi.org/10.48550/arxiv.2404.13161>.
- Blinken, Antony J. 2022. "Attribution of Russia's Malicious Cyber Activity Against Ukraine: Press Statement." U.S. Department of State.
<https://2021-2025.state.gov/attribution-of-russias-malicious-cyber-activity-against-ukraine/>. Archived May 6, 2026, at the [Wayback Machine](#).
- Blount, Joseph. 2021. *Testimony of Joseph Blount, President and Chief Executive Officer, Colonial Pipeline Company*. U.S. House of Representatives, Committee on Homeland Security.
<https://www.congress.gov/117/meeting/house/112689/witnesses/HHRG-117-HM00-Wstate-BlountJ-20210609.pdf>. Archived December 7, 2025, at the [Wayback Machine](#).

- Borys, Christian. 2017. "Ukraine braces for further cyber-attacks." BBC News. <https://www.bbc.co.uk/news/technology-40706093>. Archived May 18, 2026, at the [Wayback Machine](#).
- Branquinho, Marcelo. 2011. "Recent Malware Infections on Control System Networks in Brazil." ACS Conference. <https://www.slideshare.net/slideshow/apresentao-tnica-infeces-por-malware-no-brasil/10824740>. Archived December 7, 2025, at the [Wayback Machine](#).
- Brighton, Henry, and Gerd Gigerenzer. 2015. "The bias bias." *Journal of Business Research* 68 (8): 1772–1784. <https://doi.org/10.1016/j.jbusres.2015.01.061>.
- Broad, William J., John Markoff, and David E. Sanger. 2011. "Israeli Test on Worm Called Crucial in Iran Nuclear Delay." *New York Times*. <https://www.nytimes.com/2011/01/16/world/middleeast/16stuxnet.html>. Archived February 21, 2026, at the [Wayback Machine](#).
- Brookings. 2021. *What Investors and the SEC Can Learn from the Texas Power Crises*. Brookings Institution. <https://www.brookings.edu/wp-content/uploads/2021/06/TX-Report-Final.pdf>. Archived March 14, 2026, at the [Wayback Machine](#).
- Buchanan, Ben. 2017. *The Legend of Sophistication in Cyber Operations*. <https://www.belfercenter.org/publication/legend-sophistication-cyber-operations>. Archived October 14, 2025, at the [Wayback Machine](#).
- Byrne, Malcolm. 2024. "Chronology of Cyber Aspects of the War in Ukraine, 2022--Present." <https://nsarchive.gwu.edu/sites/default/files/2024-07/Ukraine%20cyber%20chronology%20-%20as%20of%20July%2015.%202024.pdf>. Archived August 5, 2025, at the [Wayback Machine](#).
- Cabinet Office. 2025. *National Risk Register: 2025 Edition*. Cabinet Office. https://assets.publishing.service.gov.uk/media/67b5f85732b2aab18314bbe4/National_Risk_Register_2025.pdf. Archived June 18, 2026, at the [Wayback Machine](#).
- California State Legislature. 2025. SB 53: Artificial Intelligence Models: Large Developers. California Legislative Information. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202520260SB53. Archived June 23, 2026, at the [Wayback Machine](#).
- Canadian Centre for Cyber Security. 2021. *Cyber Threat Bulletin: the Cyber Threat to Operational Technology*. Government of Canada. https://www.cyber.gc.ca/sites/default/files/cyber/2021-12/Cyber-Threat-to-Operational-Technology-white_e.pdf. Archived March 15, 2026, at the [Wayback Machine](#).
- Ceppas de Castro, Rebecca, et al. 2026. *Forecasting AI Cyber Risks and Capabilities: Results of a 2025 Pilot Study*. FRI Report #7. Forecasting Research Institute. <https://forecastingresearch.org/pdf/ai-cyber-risks-capabilities.pdf>.
- CERC. 2012. *Report on the Grid Disturbance on 30th July 2012 and 31st July 2012*. Petition No. 167/Suo-Motu/2012. Central Electricity Regulatory Commission. https://www.cercind.gov.in/2012/orders/Final_Report_Grid_Disturbance.pdf. Archived June 14, 2025, at the [Wayback Machine](#).

CERT Polska. 2026. *Energy Sector Incident Report – 29 December*. NASK – National Research Institute.

https://cert.pl/uploads/docs/CERT_Polska_Energy_Sector_Incident_Report_2025.pdf.

Archived May 20, 2026, at the [Wayback Machine](#).

CERT-UA. 2022. Кібератака зруну Sandworm (UAC-0082) на об'єкти енергетики України з використанням шкідливих програм INDUSTROYER2 та CADDYWIPER (CERT-UA#4435).

<https://cert.gov.ua/article/39518>. Archived June 16, 2026, at the [Wayback Machine](#).

Chestney, Nina. 2024. "Ukraine Faces Winter Power Shortfall of One-Third of Peak Demand, IEA Says."

<https://www.reuters.com/world/europe/ukraine-faces-6-gw-power-supply-shortfall-this-winter-iea-says-2024-09-19/>. Archived September 19, 2024, at [archive.today](#).

Chondrogiannis, S., et al. 2017. *Power grid recovery after natural hazard impact*. European Commission, Joint Research Centre. <https://doi.org/10.2760/87402>.

CISA. 2024. *PRC State-Sponsored Actors Compromise and Maintain Persistent Access to U.S. Critical Infrastructure*. AA24-038A. Cybersecurity and Infrastructure Security Agency.

<https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-038a>. Archived June 22, 2026, at the [Wayback Machine](#).

Committee on Enhancing the Resilience of the Nation's Electric Power Transmission and Distribution System, et al. 2017. *Enhancing the Resilience of the Nation's Electricity System*. National Academies Press. <https://doi.org/10.17226/24836>.

Conway, Tim, Robert M. Lee, and Jeff Shearer. 2020. *Analysis of the Recent Report of Supply Chain Attacks on US Electric Infrastructure by Chinese Actors (ICS Defense Use Case No. 7)*. SANS ICS.

https://ics.sans.org/media/SANS_ICS_DUC_7_supply_chain_attacks_on_US_electric_infrastructure.pdf#page=8. Archived January 22, 2021, at the [Wayback Machine](#).

CRS. 2014. *Physical Security of the U.S. Power Grid: High-Voltage Transformer Substations*.

R43604. Congressional Research Service. <https://sgp.fas.org/crs/homesec/R43604.pdf>.

Archived March 8, 2026, at the [Wayback Machine](#).

———. 2024. *Attacks on Ukraine's Electric Grid: Insights for U.S. Infrastructure Security and Resilience*.

<https://www.congress.gov/crs-product/R48067>. Archived February 18, 2026, at the [Wayback Machine](#).

CSIS. 2026. "Significant Cyber Incidents." Center for Strategic and International Studies, Strategic Technologies Program.

<https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents>.

Archived July 18, 2026, at the [Wayback Machine](#).

Cybersecurity Infrastructure Security Agency. 2018. "Petya Ransomware."

<https://www.cisa.gov/news-events/alerts/2017/07/01/petya-ransomware>. Archived April 12,

2026, at the [Wayback Machine](#).

———. 2022a. “Russian State-Sponsored and Criminal Cyber Threats to Critical Infrastructure.” <https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-110a>. Archived June 11, 2026, at the [Wayback Machine](#).

———. 2022b. “U.S. Government Attributes Cyberattacks on SATCOM Networks to Russian State-Sponsored Malicious Cyber Actors.” U.S. Department of Homeland Security. <https://www.cisa.gov/news-events/alerts/2022/05/10/us-government-attributes-cyberattacks-satcom-networks-russian-state-sponsored-malicious-cyber-actors>. Archived May 4, 2026, at the [Wayback Machine](#).

Cys Centrum. 2016. “Киберугроза BlackEnergy2/3. История атак на критическую ИТ инфраструктуру Украины.” CyS Centrum. https://cys-centrum.com/ru/news/black_energy_2_3. Archived March 11, 2026, at the [Wayback Machine](#).

Dallas Fed. 2021. “Cost of Texas’ 2021 deep freeze justifies weatherization.” <https://www.dallasfed.org/research/economics/2021/0415>. Archived April 23, 2026, at the [Wayback Machine](#).

Daniel, Frank Jack. 2012. “Power cut hits millions, among world’s worst outages.” Reuters. <http://in.reuters.com/article/india-blackout-delhi-northern-power-idINDEE86U05C20120731>. Archived June 14, 2017, at the [Wayback Machine](#).

De Falco, Marco. 2012. *Stuxnet Facts Report: A technical and strategical analysis*. CCDCOE. https://ccdcoe.org/uploads/2018/10/Falco2012_StuxnetFactsReport.pdf. Archived June 8, 2026, at the [Wayback Machine](#).

Department of Homeland Security. 2014. FOIA Documents: Control Systems Security Aurora Update Brief. US Department of Homeland Security. https://muckrock.s3.amazonaws.com/foia_files/14f00304-Documents.pdf. Archived October 11, 2015, at the [Wayback Machine](#).

———. 2024a. *DHS Publishes Guidelines and Report to Secure Critical Infrastructure and Weapons of Mass Destruction from AI-Related Threats*. Department of Homeland Security. <https://www.dhs.gov/archive/news/2024/04/29/dhs-publishes-guidelines-and-report-secure-critical-infrastructure-and-weapons-mass>. Archived June 24, 2026, at the [Wayback Machine](#).

———. 2024b. *Mitigating Artificial Intelligence (AI) Risk: Safety and Security Guidelines for Critical Infrastructure Owners and Operators*. Department of Homeland Security. https://www.dhs.gov/sites/default/files/2024-04/24_0426_dhs_ai-ci-safety-security-guidelines-508c.pdf. Archived June 22, 2026, at the [Wayback Machine](#).

Deutsche Windtechnik. 2022. “Cyber attack on Deutsche Windtechnik.” <https://www.deutsche-windtechnik.com/nl/news/news/detail/cyber-attack-on-deutsche-windtechnik/>. Archived March 14, 2026, at the [Wayback Machine](#).

Di Pinto, Alessandro, Younes Dragoni, and Andrea Carcano. 2018. *TRITON: The First ICS Cyber Attack on Safety Instrument Systems*. Nozomi Networks (Black Hat USA 2018). https://icscsi.org/library/Documents/Cyber_Events/Nozomi%20-%20TRITON%20-%20The%20First%20SIS%20Cyberattack.pdf. Archived April 26, 2025, at the [Wayback Machine](#).

Dixi Group. 2025. “Electricity Outages Lasted Almost 2 Thousand Hours in 2024.” Dixi Group. <https://dixigroup.org/en/electricity-outages-lasting-2-thousand-hours-for-ukrainian-households-in-2024/>. Archived June 24, 2026, at the [Wayback Machine](#).

DOE. 2004. *Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*. U.S.-Canada Power System Outage Task Force. <https://www.energy.gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf>. Archived June 10, 2026, at the [Wayback Machine](#).

———. 2009. *Comparing the Impacts of the 2005 and 2008 Hurricanes on U.S. Energy Infrastructure*. OE/ISER Report. U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability. <https://www.oe.netl.doe.gov/docs/HurricaneComp0508r2.pdf#page=12>. Archived November 28, 2024, at the [Wayback Machine](#).

———. 2013. *Economic Benefits of Increasing Electric Grid Resilience to Weather Outages*. Executive Office of the President (President’s Council of Economic Advisers) and U.S. Department of Energy. https://www.energy.gov/sites/prod/files/2013/08/f2/Grid%20Resiliency%20Report_FINAL.pdf. Archived June 14, 2026, at the [Wayback Machine](#).

———. 2014a. *Considerations for a Power Transformer Emergency Spare Strategy for the Electric Utility Industry*. Electric Power Research Institute (for U.S. DHS Science and Technology Directorate). <https://www.dhs.gov/sites/default/files/publications/RecX%20-%20Emergency%20Spare%20Transformer%20Strategy-508.pdf>. Archived June 22, 2026, at the [Wayback Machine](#).

———. 2014b. *Large Power Transformers and the U.S. Electric Grid (April 2014 Update)*. U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability. <https://www.energy.gov/sites/prod/files/2014/04/f15/LPTStudyUpdate-040914.pdf>. Archived June 9, 2026, at the [Wayback Machine](#).

———. 2015. *Review of Controls for Protecting Nonpublic Information at the Federal Energy Regulatory Commission*. Inspection Report DOE/IG-0933. U.S. Department of Energy, Office of Inspector General. <https://www.energy.gov/sites/default/files/2015/02/f19/DOE-IG-0933.pdf>. Archived February 27, 2025, at the [Wayback Machine](#).

———. 2016a. *Eastern Renewable Generation Integration Study*. <https://gmlc.doe.gov/news-and-events/eastern-renewable-generation-integration-study>. Archived April 13, 2026, at the [Wayback Machine](#).

———. 2016b. *Electric Grid Security and Resilience: Establishing a Baseline for Adversarial Threats*. U.S. Department of Energy. <https://www.energy.gov/sites/prod/files/2017/01/f34/Electric%20Grid%20Security%20and%20Resilience--Establishing%20a%20Baseline%20for%20Adversarial%20Threats.pdf>. Archived February 16, 2026, at the [Wayback Machine](#).

———. n.d.. “Electric Emergency Incident and Disturbance Report (Form DOE-417).” U.S. Department of Energy, Office of Cybersecurity, Energy Security, and Emergency Response. <https://doe417.pnnl.gov/>. Archived February 3, 2026, at the [Wayback Machine](#).

DOJ. 2021. *United States v. Evgeny Viktorovich Gladkikh: Indictment*. U.S. Department of Justice. <https://drive.google.com/file/d/1bJRTkFg17FScaYdOTpbPks1HyNfsw1uO/view>.

Dragos. 2017. *CRASHOVERRIDE: Threat to the Electric Grid Operations*. Dragos. <https://nsarchive.gwu.edu/sites/default/files/documents/3869008/Dragos-CRASHOVERRIDE-Analyzing-the-Threat-to.pdf>. Archived February 22, 2026, at the [Wayback Machine](#).

———. 2019. “The Evolution of Cyber Attacks on Electric Operations.” <https://www.dragos.com/blog/the-evolution-of-cyber-attacks-on-electric-operations/>. Archived March 14, 2026, at the [Wayback Machine](#).

———. 2024a. “The 4th Annual Dragos Capture the Flag (CTF) Results Are In!” <https://www.dragos.com/blog/the-4th-annual-dragos-capture-the-flag-ctf-results-are-in>. Archived March 8, 2026, at the [Wayback Machine](#).

———. 2024b. *OT QuickStart Guide for IT Professionals*. Dragos. https://hub.dragos.com/hubfs/116-Datasheets/Dragos_OT_QuickStart_Guide_for_IT_Professionals.pdf#page=3. Archived July 1, 2026, at the [Wayback Machine](#).

———. 2026a. “AI in the Breach: How an Adversary Leveraged AI to Target a Water Utility’s OT.” <https://www.dragos.com/blog/ai-assisted-ics-attack-water-utility>. Archived June 11, 2026, at the [Wayback Machine](#).

———. 2026b. *ELECTRUM: Cyber Attack on Poland’s Electric System 2025*. Dragos. <https://5943619.hs-sites.com/hubfs/Reports/dragos-2025-poland-attack-report.pdf>. Archived June 1, 2026, at the [Wayback Machine](#).

E-ISAC. 2016. *Analysis of the Cyber Attack on the Ukrainian Power Grid (Defense Use Case)*. E-ISAC and SANS ICS. https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf. Archived April 1, 2018, at the [Wayback Machine](#).

EconomyNext. 2025. “Sri Lanka Hit by Nation-Wide Cascading Power Failure.” *EconomyNext*. <https://economynext.com/sri-lanka-hit-by-nation-wide-cascading-power-failure-204393/>. Archived June 24, 2026, at the [Wayback Machine](#).

EEL. 2023. *Spare Equipment and Grid Resilience*. Edison Electric Institute. <https://www.eei.org/-/media/Project/EEL/Documents/Issues-and-Policy/Reliability-and-Emergency-Response/Spare-Equipment-and-Grid-Resilience-Programs.pdf>. Archived February 15, 2026, at the [Wayback Machine](#).

EIA. 2012. *Electric Power Annual 2012*. U.S. Energy Information Administration. <http://large.stanford.edu/courses/2016/ph241/orozco2/docs/eia-epa-2012.pdf>. Archived June 4, 2024, at the [Wayback Machine](#).

———. 2020. *About 25% of U.S. Power Plants Can Start up within an Hour*. U.S. Energy Information Administration. <https://www.eia.gov/todayinenergy/detail.php?id=45956>. Archived May 5, 2026, at the [Wayback Machine](#).

———. 2024. Table 4.1. Count of Electric Power Industry Power Plants, by Sector, by Predominant Energy Sources within Plant, 2014 through 2024. U.S. Energy Information Administration. https://www.eia.gov/electricity/annual/html/epa_04_01.html. Archived June 22, 2026, at the [Wayback Machine](#).

Elcon. 2004. *The Economic Impacts of the August 2003 Blackout*. Electricity Consumers Resource Council (ELCON). <https://www.nrc.gov/docs/ml1113/ml111300584.pdf>. Archived October 7, 2025, at the [Wayback Machine](#).

Electric Power Research Institute. 2019. *High-Altitude Electromagnetic Pulse and the Bulk Power System: Potential Impacts and Mitigation Strategies*. 3002014979. Electric Power Research Institute. <https://www.roxtec.com/globalassets/03.-files/campaign-pages/emc/2019-epri-report.pdf>. Archived January 18, 2026, at the [Wayback Machine](#).

EMP Commission. 2008. *Report of the Commission to Assess the Threat to the United States from Electromagnetic Pulse (EMP) Attack: Critical National Infrastructures*. Electromagnetic Pulse (EMP) Commission. <https://apps.dtic.mil/sti/pdfs/ADA484672.pdf>. Archived October 23, 2025, at the [Wayback Machine](#).

Energy Map. 2025. "Information on Electricity Consumption Limitation Measures." Energy Map. <https://map.ua-energy.org/en/resources/0f8f9882-1fb2-47c6-81dc-31fbad914f16/>. Archived May 24, 2025, at the [Wayback Machine](#).

Epoch. 2026. "Open models lag state-of-the-art closed models by 4 months." Epoch. <https://epoch.ai/data-insights/open-closed-eci-gap>. Archived June 24, 2026, at the [Wayback Machine](#).

ESET. 2017. *Win32/Industroyer: A New Threat for Industrial Control Systems*. ESET (Anton Cherepanov). https://web-assets.esetstatic.com/wls/2017/06/Win32_Industroyer.pdf. Archived June 10, 2026, at the [Wayback Machine](#).

———. 2022. "Industroyer2: Industroyer reloaded." <https://www.welivesecurity.com/2022/04/12/industroyer2-industroyer-reloaded/>. Archived May 15, 2026, at the [Wayback Machine](#).

ESET Research. 2022. *ESET Threat Report T1 2022*. https://www.eset.com/fileadmin/ESET/INT/B2B_Resource_centrum/Reports/T1-2022_Threat-Report.pdf. Archived October 13, 2025, at the [Wayback Machine](#).

Fassihi, Farnaz, and Ronen Bergman. 2021. "Israel and Iran Broaden Cyberwar to Attack Civilian Targets." *New York Times*. <https://www.nytimes.com/2021/11/27/world/middleeast/iran-israel-cyber-hack.html>. Archived January 20, 2026, at the [Wayback Machine](#).

FCDO. 2026. "UK and EU strike Russian cyber networks with new sanctions." Press release, Foreign, Commonwealth & Development Office. <https://www.gov.uk/government/news/uk-and-eu-strike-russian-cyber-networks-with-new-sanctions>. Archived July 26, 2026, at the [Wayback Machine](#).

FERC et al. 2021. *The February 2021 Cold Weather Outages in Texas and the South Central United States* (FERC-NERC-Regional Entity Staff Report). Federal Energy Regulatory Commission and North American Electric Reliability Corporation.

https://www.nerc.com/pa/rrm/ea/Documents/February_2021_Cold_Weather_Report.pdf.
Archived August 30, 2025, at the [Wayback Machine](#).

Fildes, Jonathan. 2010. “Stuxnet worm ‘targeted high-value Iranian assets’” BBC News.
<https://www.bbc.co.uk/news/technology-11388018>. Archived June 19, 2026, at the [Wayback Machine](#).

Forescout Research, and Vedere Labs. 2024. *Clearing the Fog of War: A Critical Analysis of Recent Energy Sector Attacks in Denmark and Ukraine*. Forescout.
<https://www.forescout.com/resources/clearing-the-fog-of-war/>. Archived December 13, 2024, at the [Wayback Machine](#).

Forster, Malcolm, and Elliott Sober. 1994. “How to Tell When Simpler, More Unified, or Less Ad Hoc Theories will Provide More Accurate Predictions.” *The British Journal for the Philosophy of Science* 45 (1): 1–35. <https://doi.org/10.1093/bjps/45.1.1>.

FRED. n.d.. “Population.” <https://fred.stlouisfed.org/series/B230RC0A052NBEA>. Archived March 6, 2026, at the [Wayback Machine](#).

Freeman, Sarah G., et al. 2024. *Attack Surface of Wind Energy Technologies in the United States*. INL/RPT-24-76133. Idaho National Laboratory.
<https://inl.gov/content/uploads/2024/02/INL-Wind-Threat-Assessment-v5.0.pdf>. Archived November 6, 2025, at the [Wayback Machine](#).

Frontier Model Forum. 2025. “Frontier AI Biosafety Thresholds.” Frontier Model Forum.
<https://www.frontiermodelforum.org/issue-briefs/frontier-ai-biosafety-thresholds/>. Archived May 9, 2026, at the [Wayback Machine](#).

Fugate, Craig. 2013. *Statement of Craig Fugate, Administrator, FEMA, before the Senate Committee on Homeland Security and Governmental Affairs (Hurricane Sandy recovery)*. Federal Emergency Management Agency.
https://www.fema.gov/sites/default/files/2020-07/craig-fugate_recovery-hurricane-sandy_testimony_11-6-2013.pdf. Archived March 27, 2026, at the [Wayback Machine](#).

GAO. 2023. *Electricity Grid: DOE Could Better Support Industry Efforts to Ensure Adequate Transformer Reserves*. GAO-23-106180. U.S. Government Accountability Office.
<https://www.gao.gov/assets/d23106180.pdf>. Archived February 8, 2026, at the [Wayback Machine](#).

Garton, David. 2019. *Purdue Model Framework for Industrial Control Systems & Cybersecurity Segmentation*. National Petroleum Council.
https://www.energy.gov/sites/default/files/2022-10/Infra_Topic_Paper_4-14_FINAL.pdf.
Archived June 23, 2026, at the [Wayback Machine](#).

- Gennari et al. 2024. *Considerations for Evaluating Large Language Models for Cybersecurity Tasks*. CMU Software Engineering Institute.
<https://sei.cmu.edu/library/considerations-for-evaluating-large-language-models-for-cybersecurity-tasks/>. Archived March 6, 2026, at the [Wayback Machine](#).
- Glover, J. Duncan, Mulukutla S. Sarma, and Thomas J. Overbye. 2017. "Chapter 3: Power Transformers." In *Power System Analysis and Design*. Cengage Learning.
http://web.ecs.baylor.edu:80/faculty/lee/ELC4340/Lecture%20note/Chapter3_GSO5.pdf. Archived March 28, 2018, at the [Wayback Machine](#).
- Google DeepMind. 2025a. *Frontier Safety Framework, Version 2.0*. Google DeepMind.
<https://storage.googleapis.com/deepmind-media/DeepMind.com/Blog/updating-the-frontier-safety-framework/Frontier%20Safety%20Framework%202.0.pdf>. Archived June 16, 2026, at the [Wayback Machine](#).
- . 2025b. *Frontier Safety Framework, Version 3.0*. Google DeepMind.
https://storage.googleapis.com/deepmind-media/DeepMind.com/Blog/strengthening-our-frontier-safety-framework/frontier-safety-framework_3.pdf. Archived June 16, 2026, at the [Wayback Machine](#).
- Gorman, Will. 2022. "The quest to quantify the value of lost load: A critical review of the economics of power outages." *The Electricity Journal* 35 (8): 107187.
<https://doi.org/10.1016/j.tej.2022.107187>.
- Gostev, Alexander. 2011. "The Mystery of Duqu: Part One." Securelist, Kaspersky Lab.
<https://securelist.com/the-mystery-of-duqu-part-one/31177/>. Archived July 3, 2026, at the [Wayback Machine](#).
- Green, Kesten C., and J. Scott Armstrong. 2015. "Simple Versus Complex Forecasting: The Evidence." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2643534>.
- Greenberg, Andy. 2022. "Russia's Sandworm Hackers Attempted a Third Blackout in Ukraine." WIRED. <https://www.wired.com/story/sandworm-russia-ukraine-blackout-gru/>. Archived January 9, 2023, at the [Wayback Machine](#).
- Halstead, John, and Luca Righetti. 2026. *Assessing the Risk of AI-Enabled Computer Worms*. GovAI.
https://govai.b-cdn.net/Report_Assessing_the_Risk_of_AI_Enabled_Computer_Worms.pdf. Archived July 13, 2026, at the [Wayback Machine](#).
- Halstead, John, and Matthew van der Merwe. 2024. "Cyber case studies." Google Docs document.
https://docs.google.com/document/d/1u1vnb20BIbBjr0Y2vEXhY84I99SwJu60m5be-I43_AM/edit.
- Harmash, Olena. 2024. "Ukrainians find new energy sources to beat blackouts as winter arrives." Reuters.
<https://www.reuters.com/world/europe/ukrainians-find-new-energy-sources-beat-blackouts-winter-arrives-2024-12-03/>. Archived November 16, 2025, at the [Wayback Machine](#).

- Hemsley, Kevin E., and Ronald E. Fisher. 2018. *History of Industrial Control System Cyber Incidents*. INL/CON-18-44411. Idaho National Laboratory.
<https://www.osti.gov/servlets/purl/1505628>. Archived October 29, 2025, at the [Wayback Machine](#).
- Hesseldahl, Arik. 2010. "Computer Worm May Be Targeting Iranian Nuclear Sites." Bloomberg.
<https://www.bloomberg.com/news/articles/2010-09-24/stuxnet-computer-worm-may-be-ai-med-at-iran-nuclear-sites-researcher-says>. Archived May 14, 2015, at the [Wayback Machine](#).
- Hines, Paul, Jay Apt, and Sarosh Talukdar. 2009. *Large Blackouts in North America: Historical Trends and Policy Implications*. Working Paper CEIC 09-01. Carnegie Mellon Electricity Industry Center. https://phines.w3.uvm.edu/publications/2009/hines_2009_blackouts.pdf. Archived March 26, 2026, at the [Wayback Machine](#).
- Hoole, P. R. P., et al. 2017. "Power Transformer Fire and Explosion: Causes and Control." *International Journal of Control Theory and Applications* 10 (16): 211-219.
https://serialsjournals.com/abstract/33400_ch_21_f_-_ijcta_paper2.pdf. Archived November 12, 2023, at the [Wayback Machine](#).
- Hultquist, John. 2016. "Sandworm Team and the Ukrainian Power Authority Attacks." Google Cloud Blog.
<https://cloud.google.com/blog/topics/threat-intelligence/ukraine-and-sandworm-team>. Archived March 11, 2026, at the [Wayback Machine](#).
- ICF. 2003. *The Economic Cost of the Blackout: An Issue Paper on the Northeastern Blackout, August 14, 2003*. ICF Consulting. <https://www.nrc.gov/docs/ML1113/ML111361105.pdf>. Archived May 15, 2025, at the [Wayback Machine](#).
- ICS-CERT. 2012. "ICS-CERT Monitor: October/November/December 2012."
https://www.cisa.gov/sites/default/files/Monitors/ICS-CERT_Monitor_Oct-Dec2012.pdf. Archived February 14, 2026, at the [Wayback Machine](#).
- . 2014. "ICS-CERT Monitor: January--April 2014."
https://www.cisa.gov/sites/default/files/Monitors/ICS-CERT_Monitor_Jan-April2014.pdf. Archived January 5, 2026, at the [Wayback Machine](#).
- ICS Investigation Expert Panel. 2025. *Grid Incident in Spain and Portugal on 28 April 2025: Factual Report*. ICS Investigation Expert Panel.
https://eepublicdownloads.blob.core.windows.net/public-cdn-container/clean-documents/Publications/2025/iberian-blackout/entso-e_incident_report_ES-PT_April_2025_06.pdf. Archived February 14, 2026, at the [Wayback Machine](#).
- ICS STRIVE. n.d. "ICS STRIVE Incident Database." ICS STRIVE (ISSSource and Waterfall Security).
<https://icsstrive.com/>. Archived 5 July, 2026, at the [Wayback Machine](#).
- Idaho National Laboratories. 2025. *TAIGR: Testing the limits of AI on the power grid*.
<https://inl.gov/feature-story/taigr-testing-the-limits-of-ai-on-the-power-grid/>. Archived February 2, 2026, at the [Wayback Machine](#).

IEC. 2023. *Cooperation for Restoring the Ukrainian Energy Infrastructure: Ukrainian Energy Sector Evaluation and Damage Assessment - IX*. International Energy Charter.
https://www.energycharter.org/fileadmin/DocumentsMedia/Occasional/2023_04_27_UA_sectoral_evaluation_and_damage_assessment_Version_IX.pdf. Archived September 22, 2023, at the [Wayback Machine](#).

Indian Express. 2021. "Mumbai 2020 outage due to 'cascade tripping,' not sabotage: Report." *Indian Express*.
<https://indianexpress.com/article/cities/mumbai/mumbai-2020-outage-due-to-cascade-tripping-not-sabotage-report-7594295/>. Archived June 24, 2026, at the [Wayback Machine](#).

INL. 2020. *CCE Case Study: Baltavia Substation Power Outage (Consequence-driven Cyber-informed Engineering)*. Idaho National Laboratory, Cybercore Integration Center.
https://inldigitallibrary.inl.gov/sites/sti/sti/Sort_26135.pdf. Archived February 22, 2025, at the [Wayback Machine](#).

Jeffries, Blaine, et al. 2022. *Cyber Risk to Mission Case Study: Norsk Hydro*. MITRE Corporation.
<https://apps.dtic.mil/sti/trecms/pdf/AD1183007.pdf>. Archived May 3, 2025, at the [Wayback Machine](#).

Kapoor et al. 2024. "On the Societal Impact of Open Foundation Models."
<https://crfm.stanford.edu/open-fms/>. Archived May 19, 2026, at the [Wayback Machine](#).

Karnofsky, Holden. 2024. "If-Then Commitments for AI Risk Reduction." Karnofsky.
<https://carnegieendowment.org/research/2024/09/if-then-commitments-for-ai-risk-reduction?lang=en>. Archived April 25, 2026, at the [Wayback Machine](#).

Kaspersky ICS CERT. 2025. "Q3 2024 -- A Brief Overview of the Main Incidents in Industrial Cybersecurity."
<https://ics-cert.kaspersky.com/publications/q3-2024-a-brief-overview-of-the-main-incidents-in-industrial-cybersecurity/>. Archived February 9, 2026, at the [Wayback Machine](#).

Keliris, Anastasis. 2020. "Are Cyber-Attacks on the Power Grid Limited to Nation-State Actors?" Black Hat YouTube. <https://www.youtube.com/watch?v=9kgzDZjl748>. Archived June 1, 2025, at the [Wayback Machine](#).

Koelemij, Sinclair. 2020. "Are Power Transformers hackable?" Koelemij.
<https://otcybersecurity.blog/2020/06/10/are-power-transformers-hackable/>. Archived November 29, 2023, at the [Wayback Machine](#).

Koessler, Leonie, Jonas Schuett, and Markus Anderljung. 2024. "Risk thresholds for frontier AI." *arXiv preprint*. <https://doi.org/10.48550/arxiv.2406.14713>.

Krotofil, Marina, Andrea Carcano, and Younes Dragoni. 2018. *TRITON: How It Disrupted Safety Systems and Changed the Threat Landscape of Industrial Control Systems, Forever*. Black Hat USA 2018.
<https://i.blackhat.com/us-18/Wed-August-8/us-18-Carcano-TRITON-How-It-Disrupted-Safety-Systems-And-Changed-The-Threat-Landscape-Of-Industrial-Control-Systems-Forever.pdf>. Archived July 16, 2019, at the [Wayback Machine](#).

- Kučinskas, Simas, et al. 2025. *Assessing Near-Term Accuracy in the Existential Risk Persuasion Tournament*. Forecasting Research Institute.
<https://static1.squarespace.com/static/635693acf15a3e2a14a56a4a/t/68b6ce72b3435a79858344b7/1756810866830/near-term-xpt-accuracy.pdf>. Archived July 9, 2026, at the [Wayback Machine](#).
- Kunz, M., et al. 2013. "Investigation of superstorm Sandy 2012 in a multi-disciplinary approach." *Natural Hazards and Earth System Sciences* 13 (10): 2579–2598.
<https://doi.org/10.5194/nhess-13-2579-2013>.
- Lakshminarayana, Subhash, et al. 2025. "Cybersecurity Threats to Power Grid Operations From the Demand-Side Response Ecosystem: A Comprehensive Overview." *IEEE Industrial Electronics Magazine* 19 (2): 62–72. <https://doi.org/10.1109/mie.2024.3524752>.
- Langner, Ralph. 2013a. "Stuxnet's Secret Twin." *Foreign Policy*.
<https://foreignpolicy.com/2013/11/19/stuxnets-secret-twin/>. Archived May 25, 2026, at the [Wayback Machine](#).
- . 2013b. *To Kill a Centrifuge: A Technical Analysis of What Stuxnet's Creators Tried to Achieve*. The Langner Group.
<https://www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf>. Archived June 8, 2025, at the [Wayback Machine](#).
- Larsen, Jason. 2020. "14 Hours and an Electric Grid." Larsen.
<https://youtu.be/UQceYzbhOkA?feature=shared&t=1330>. Archived February 17, 2026, at the [Wayback Machine](#).
- Leahy, Eimear, and Richard S.J. Tol. 2011. "An estimate of the value of lost load for Ireland." *Energy Policy* 39 (3): 1514–1520. <https://doi.org/10.1016/j.enpol.2010.12.025>.
- Lee, Robert M. 2019. "Homogeneous Infrastructure and Scalable Attacks."
<https://www.robertmlee.org/homogeneous-infrastructure-and-scalable-attacks/1000/>. Archived July 1, 2026, at the [Wayback Machine](#).
- Lee, Robert M., Michael J. Assante, and Tim Conway. 2014. *German Steel Mill Cyber Attack (ICS CP/PE Case Study)*. SANS ICS.
https://ics.sans.org/media/ICS-CPPE-case-Study-2-German-Steelworks_Facility.pdf. Archived October 1, 2020, at the [Wayback Machine](#).
- Lee, Susan, Michael Moskowitz, and Jane Pinelis. 2018. *Quantifying Improbability: An Analysis of the Lloyd's of London Business Blackout Cyber Attack Scenario*. National Security Report. Johns Hopkins University Applied Physics Laboratory.
<https://www.jhuapl.edu/sites/default/files/2022-12/QuantifyingImprobability.pdf>. Archived August 13, 2025, at the [Wayback Machine](#).
- Lipovský, Robert, and Anton Cherepanov. 2021. "Industroyer2: Sandworm's Cyberwarfare Targets Ukraine's Power Grid Again." Black Hat.
https://www.youtube.com/watch?v=xC9iM5wVedQ&t=1220s&ab_channel=BlackHat. Archived July 1, 2026, at the [Wayback Machine](#).

Litchfield, David. 2003. "The Inside Story of SQL Slammer." Threatpost. <https://threatpost.com/inside-story-sql-slammer-102010/74589/>. Archived March 15, 2025, at the [Wayback Machine](#).

Lloyd's and CRS. 2015a. *Business Blackout: Appendix 2 --- Scenario Design and Impact Modelling Methodologies*. University of Cambridge Centre for Risk Studies and Lloyd's. <https://assets.lloyds.com/assets/pdf-business-blackout-appendix-2/1/pdf-business-blackout-appendix-2.pdf>. Archived March 1, 2026, at the [Wayback Machine](#).

———. 2015b. *Business Blackout: The Insurance Implications of a Cyber Attack on the US Power Grid*. Emerging Risk Report 2015. Lloyd's and Cambridge Centre for Risk Studies. <https://www.jbs.cam.ac.uk/wp-content/uploads/2020/08/crs-lloyds-business-blackout-scenario.pdf>. Archived March 1, 2026, at the [Wayback Machine](#).

Machtemes, Rees, et al. 2025. *OT Cyber Threat Report: Navigating the Future of OT Security*. Waterfall Security Solutions and ICS STRIVE. <https://waterfall-security.com/wp-content/uploads/2025/03/2025-OT-Cyber-Security-Threat-Report.pdf>. Archived February 11, 2026, at the [Wayback Machine](#).

Makridakis, Spyros, and Michèle Hibon. 2000. "The M3-Competition: results, conclusions and implications." *International Journal of Forecasting* 16 (4): 451–476. [https://doi.org/10.1016/S0169-2070\(00\)00057-1](https://doi.org/10.1016/S0169-2070(00)00057-1).

Mandiant. 2017. "Attackers Deploy New ICS Attack Framework "TRITON" and Cause Operational Disruption to Critical Infrastructure." Google Cloud Blog. <https://cloud.google.com/blog/topics/threat-intelligence/attackers-deploy-new-ics-attack-framework-triton/>. Archived June 24, 2026, at the [Wayback Machine](#).

———. 2023. "Sandworm Disrupts Power in Ukraine Using a Novel Attack Against Operational Technology." Google Cloud Blog. <https://cloud.google.com/blog/topics/threat-intelligence/sandworm-disrupts-power-ukraine-operational-technology/>. Archived June 24, 2026, at the [Wayback Machine](#).

———. 2025. "Protecting the Core: Securing Protection Relays in Modern Substations." Google Cloud Blog. <https://cloud.google.com/blog/topics/threat-intelligence/securing-protection-relays-modern-substations>. Archived May 7, 2026, at the [Wayback Machine](#).

Maschmeyer, Lennart. 2021. "The Subversive Trilemma: Why Cyber Operations Fall Short of Expectations." *International Security* 46 (2): 51–90. https://doi.org/10.1162/isec_a_00418.

Mellers, Barbara, et al. 2015. "Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions." *Perspectives on Psychological Science* 10 (3): 267–281. <https://doi.org/10.1177/1745691615577794>.

Mellers, Barbara, et al. 2014. "Psychological Strategies for Winning a Geopolitical Forecasting Tournament." *Psychological Science* 25 (5): 1106–1115. <https://doi.org/10.1177/0956797614524255>.

Melman, Yossi, and Dan Raviv. 2025. “Israel Secretly Recruited Iranian Dissidents to Attack Their Country From Within.” *ProPublica*.

<https://www.propublica.org/article/israel-iran-war-mossad-iranian-recruits>. Archived May 30, 2026, at the [Wayback Machine](#).

Meta. 2025. “Frontier AI Framework, Version 1.1.”

<https://ai.meta.com/static-resource/meta-frontier-ai-framework/>. Archived February 27, 2026, at the [Wayback Machine](#).

Metaculus. n.d.a. “US-China war before 2035?”

<https://www.metaculus.com/questions/8362/us-china-war-by-2035/>. Archived January 30, 2026, at the [Wayback Machine](#).

———. n.d.b. “US-Russia War before 2050.”

<https://www.metaculus.com/questions/7452/us-russia-war-before-2050/>. Archived February 12, 2026, at the [Wayback Machine](#).

METR. 2024. “Common Elements of Frontier AI Safety Policies.”

https://metr.org/assets/common_elements_of_frontier_ai_safety_policies.pdf.

Midnight Blue. 2026. “On the risk of destructive bricking attacks against OT devices (part 1).”

<https://www.midnightblue.nl/blog/have-you-tried-turning-it-off-and-on-again-part-1>. Archived June 2, 2026, at the [Wayback Machine](#).

Miller, Thomas, et al. 2021. “Looking back to look forward: Lessons learnt from cyber-attacks on Industrial Control Systems.” *International Journal of Critical Infrastructure Protection* 35: 100464.

<https://doi.org/10.1016/j.ijcip.2021.100464>.

MITRE. 2025. “2025 Poland Wiper Attacks.” MITRE ATT&CK.

<https://attack.mitre.org/campaigns/C0063/>. Archived June 3, 2026, at the [Wayback Machine](#).

Modderkolk, Huib. 2024. “Sabotage in Iran: een missie in duisternis.” *de Volkskrant*.

<https://www.volkskrant.nl/kijkverder/v/2024/sabotage-in-iran-een-missie-in-duisternis-v989743/>. Archived May 18, 2026, at the [Wayback Machine](#).

Molfar Intelligence Institute. 2023. “Investigation into Russian Military Units Engaged in Psychological Operations (PSYOP) and Hacking Attacks.” Molfar Intelligence Institute.

<https://www.molfar.institute/en/russian-psyop-and-hacking-attacks/>. Archived September 12, 2024, at the [Wayback Machine](#).

Morehouse, Catherine. 2023. “Extremists Keep Trying to Trigger Mass Blackouts — and That’s Not Even the Scariest Part.” *Politico*.

<https://www.politico.com/news/2023/09/10/power-grid-attacks-00114563>. Archived September 27, 2025, at the [Wayback Machine](#).

Morgan, Millett Granger, and Max Henrion. 2012. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press.

<https://doi.org/10.1017/cbo9780511840609>.

Motherboard. 2016. “The Ukrainian Power Grid Was Hacked Again.” *Vice*.

<https://www.vice.com/en/article/ukrainian-power-station-hacking-december-2016-report/>.

- Nagy, John, and Jonathan Bym. 2023. "Moore County Warrants Shed Light on 2022 Power Grid Attacks." *The Pilot*.
https://www.thepilot.com/news/moore-county-warrants-shed-light-on-2022-power-grid-attacks/article_80d86202-9b93-11ee-bf23-bf21b492a7e4.html. Archived August 1, 2025, at the [Wayback Machine](#).
- National Academies of Sciences, Engineering, and Medicine. 2008. *Severe Space Weather Events: Understanding Societal and Economic Impacts*. National Academies Press.
<https://www.nationalacademies.org/read/12507/chapter/9>. Archived June 24, 2026, at the [Wayback Machine](#).
- . 2012. "Restoration of the Electric Power System After an Attack." In *Terrorism and the Electric Power Delivery System*. Chap. 7. National Academies Press.
<https://doi.org/10.17226/12050>.
- National Security Archive. 2019. "Attackers Trigger Firewall Reboot to Briefly Interrupt Power Company Network."
<https://nsarchive.gwu.edu/news/cyber-vault/2019-10-31/attackers-trigger-firewall-reboot-briefly-interrupt-power-company-network>. Archived November 16, 2025, at the [Wayback Machine](#).
- NERC. 2004. *Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn?* North American Electric Reliability Council.
https://www.nerc.com/pa/rrm/ea/August%2014%202003%20Blackout%20Investigation%20DL/NERC_Final_Blackout_Report_07_13_04.pdf. Archived November 21, 2023, at the [Wayback Machine](#).
- . 2025. "NERC Compliance Registry Matrix." Excel spreadsheet. North American Electric Reliability Corporation.
https://www.nerc.com/pa/comp/Registration%20and%20Certification%20DL/NERC_Compliance_Registry_Matrix_Excel.xlsx. Archived September 1, 2025, at the [Wayback Machine](#).
- . 2021. *The February 2021 Cold Weather Outages in Texas and the South Central United States* (FERC-NERC Joint Staff Report). North American Electric Reliability Corporation.
<https://www.nerc.com/globalassets/standards/projects/2021-07/ferc-presentation-phase-2.pdf>. Archived June 24, 2026, at the [Wayback Machine](#).
- NIST. 2025. *Managing Misuse Risk for Dual-Use Foundation Models (Second Public Draft)*. NIST AI 800-1 2pd. U.S. AI Safety Institute, NIST.
<https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-1.ipd2.pdf>. Archived January 15, 2025, at the [Wayback Machine](#).
- North American Electric Reliability Corporation. 2019. *Risks Posed by Firewall Firmware Vulnerabilities*.
https://www.nerc.com/globalassets/programs/event-analysis/lessons-learned/20190901_risks_posed_by_firewall_firmware_vulnerabilities.pdf. Archived May 7, 2025, at the [Wayback Machine](#).

NY DFS. 2021. *Report on the SolarWinds Cyber Espionage Attack and Institutions' Response*. New York State Department of Financial Services.

https://www.dfs.ny.gov/system/files/documents/2021/04/solarwinds_report_2021.pdf.

Archived June 24, 2026, at the [Wayback Machine](#).

O'Brien, James G., et al. 2022. *Electric Grid Blackstart: Trends, Challenges, and Opportunities*. PNNL-32773. Pacific Northwest National Laboratory.

https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-32773.pdf.

Archived March 26, 2026, at the [Wayback Machine](#).

O'Neill, Patrick Howell. 2022. "Russian Hackers Tried to Bring Down Ukraine's Power Grid to Help the Invasion." MIT Technology Review.

<https://www.technologyreview.com/2022/04/12/1049586/russian-hackers-tried-to-bring-down-ukraines-power-grid-to-help-the-invasion/>. Archived May 11, 2026, at the [Wayback Machine](#).

ODNI. 2023. *Annual Threat Assessment of the U.S. Intelligence Community*. Office of the Director of National Intelligence.

<https://www.dni.gov/files/ODNI/documents/assessments/ATA-2023-Unclassified-Report.pdf>.

Archived June 14, 2026, at the [Wayback Machine](#).

OpenAI. 2023. *Preparedness Framework (Beta)*. OpenAI.

<https://cdn.openai.com/openai-preparedness-framework-beta.pdf>. Archived June 16, 2026, at the [Wayback Machine](#).

———. 2025a. *OpenAI o3 and o4-mini System Card*. OpenAI.

<https://cdn.openai.com/pdf/2221c875-02dc-4789-800b-e7758f3722c1/o3-and-o4-mini-system-card.pdf>. Archived June 16, 2026, at the [Wayback Machine](#).

———. 2025b. *Preparedness Framework, Version 2*. OpenAI.

<https://cdn.openai.com/pdf/18a02b5d-6b67-4cec-ab64-68cdfbddebcdf/preparedness-framework-v2.pdf>. Archived June 20, 2026, at the [Wayback Machine](#).

Owens, William A., Kenneth W. Dam, and Herbert S. Lin, eds. 2009. *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities*. National Academies Press.

<https://doi.org/10.17226/12651>.

Petkauskas, Vilius. 2022. "Deutsche Windtechnik hit with a cyberattack, a third on Germany's wind energy sector." Cybernews.

<https://cybernews.com/news/deutsche-windtechnik-hit-with-a-cyberattack-a-third-on-germanys-wind-energy-sector/>. Archived October 5, 2025, at the [Wayback Machine](#).

PHDays. 2016. "PHDays — Positive Hack Days. Teenager Hacks Electrical Substation at PHDays."

<https://2016.phdays.com/press/news/62397/>. Archived September 18, 2019, at the [Wayback Machine](#).

PJM. 2023. *PJM Manual 36: System Restoration, Revision 32*. PJM, System Operations Division.

<https://www.pjm.com/-/media/DotCom/documents/manuals/archive/m36/m36v32-system-restoration-06-15-2023.pdf>. Archived July 3, 2026, at the [Wayback Machine](#).

- Politico. 2022. “Physical attacks on power grid surge to new peak.” Politico. <https://subscriber.politicopro.com/article/2022/12/physical-attacks-on-u-s-electrical-grid-surge-to-new-peak-00075216>. Archived June 9, 2026, at the [Wayback Machine](#).
- . 2024. “DOE data show grid security breaches reached all-time high in 2023.” Politico. <https://subscriber.politicopro.com/article/2024/04/grid-security-breaches-reached-all-time-high-in-2023-doe-data-00150223>. Archived December 3, 2024, at the [Wayback Machine](#).
- Polityuk, Pavel. 2016. “Ukraine Investigates Suspected Cyber Attack on Kiev Power Grid.” <https://www.reuters.com/article/us-ukraine-crisis-cyber-attacks-idUSKBN1491ZF/>. Archived June 11, 2017, at the [Wayback Machine](#).
- Poulsen, Kevin. 2003. “Slammer Worm Crashed Ohio Nuke Plant Net.” <https://www.theregister.com/security/2003/08/20/slammer-worm-crashed-ohio-nuke-plant-net/1487486>. Archived February 13, 2026, at the [Wayback Machine](#).
- RAND. 2024. *Securing AI Model Weights: Preventing Theft and Misuse of Frontier Models*. https://www.rand.org/pubs/research_reports/RRA2849-1.html. Archived June 18, 2026, at the [Wayback Machine](#).
- Reuters. 2018. “Power Restoration After Major U.S. Hurricanes.” Reuters. <https://www.reuters.com/graphics/STORM-MARIA-POWER/010050ZK27Z/>. Archived April 13, 2025, at the [Wayback Machine](#).
- . 2022a. “Montenegro Blames Criminal Gang for Cyber Attacks on Government.” <https://www.reuters.com/world/europe/montenegro-blames-criminal-gang-cyber-attacks-government-2022-08-31/>. Archived March 17, 2025, at the [Wayback Machine](#).
- . 2022b. “Montenegro’s State Infrastructure Hit by Cyber Attack -- Officials.” <https://www.reuters.com/world/europe/montenegros-state-infrastructure-hit-by-cyber-attack-officials-2022-08-26/>. Archived November 20, 2023, at the [Wayback Machine](#).
- Rieger, Craig G., et al. 2022. *Distributed Renewables Cyber Resilience*. INL/MIS-22-66611. Idaho National Laboratory. https://inldigitallibrary.inl.gov/sites/sti/sti/Sort_59626.pdf. Archived September 21, 2025, at the [Wayback Machine](#).
- Rodriguez, Mikel, et al. 2025. “A Framework for Evaluating Emerging Cyberattack Capabilities of AI.” *arXiv preprint*. <https://doi.org/10.48550/arxiv.2503.11917>.
- Roodman, David. 2015. *The Risk of Geomagnetic Storms to the Grid: A Preliminary Review*. GiveWell. <https://davidroodman.com/david/The%20risk%20of%20geomagnetic%20storms%205%20dr.pdf>. Archived June 2, 2026, at the [Wayback Machine](#).
- S4P. 2023. “US Offers up to \$10 Million for Info on Cyber Attacks in Montenegro.” Support4Partnership. <https://support4partnership.org/en/news/us-offers-up-to-10-million-for-info-on-cyber-attacks-in-montenegro>. Archived July 1, 2026, at the [Wayback Machine](#).

Sandhills Pride, and Sunrise Theater. 2022. “Downtown Divas @ Sunrise Theater.” Facebook event.

<https://web.facebook.com/events/sunrise-theater/downtown-divas-sunrise-theater/818165132773921/>. Archived June 29, 2026, at the [Wayback Machine](#).

Sanger, David E., and Emily Schmall. 2021. “China Appears to Warn India: Push Too Hard and the Lights Could Go Out.” *New York Times*.

<https://www.nytimes.com/2021/02/28/us/politics/china-india-hacking-electricity.html>. Archived January 21, 2026, at the [Wayback Machine](#).

Schröder, Thomas, and Wilhelm Kuckshinrichs. 2015. “Value of Lost Load: An Efficient Economic Indicator for Power Supply Security? A Literature Review.” *Frontiers in Energy Research* 3.

<https://doi.org/10.3389/fenrg.2015.00055>.

SecurityWeek. 2019. “Dutch Engineer Used Water Pump to Get Billion-Dollar Stuxnet Malware into Iranian Nuclear Facility: Report.” SecurityWeek.

<https://www.securityweek.com/dutch-engineer-used-water-pump-to-get-billion-dollar-stuxnet-malware-into-iranian-nuclear-facility-report/>. Archived May 18, 2026, at the [Wayback Machine](#).

SektorCERT. 2023. *The Attack Against Danish Critical Infrastructure*.

<https://sektorcert.dk/wp-content/uploads/2023/11/SektorCERT-The-attack-against-Danish-critical-infrastructure-TLP-CLEAR.pdf>. Archived April 24, 2026, at the [Wayback Machine](#).

Sela, Eyal. 2026. *The AI-Assisted Breach of Mexico's Government Infrastructure*. Gambit Security.

https://cdn.prod.website-files.com/69944dd945f20ca4a27a7c47/69d8bb5aea59e31efb3b8a7f_Tech_Report_ai_breach_mex_gov.pdf. Archived June 24, 2026, at the [Wayback Machine](#).

Silverman, S. M., and E. W. Cliver. 2001. “Low-latitude auroras: the magnetic storm of 14–15 May 1921.” *Journal of Atmospheric and Solar-Terrestrial Physics* 63 (5): 523–535.

[https://doi.org/10.1016/s1364-6826\(00\)00174-7](https://doi.org/10.1016/s1364-6826(00)00174-7).

Singer, Brian, et al. 2023. “Shedding Light on Inconsistencies in Grid Cybersecurity: Disconnects and Recommendations.” . <https://doi.org/10.1109/sp46215.2023.10179343>.

Slayton, Rebecca. 2017. “What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment.” *International Security* 41 (3): 72–109. https://doi.org/10.1162/isec_a_00267.

Slowik, Joe. 2018. *Anatomy of an Attack: Detecting and Defeating CRASHOVERRIDE*. Dragos (Virus Bulletin 2018). <https://pylos.co/wp-content/uploads/2022/10/VB2018-Slowik.pdf>. Archived December 17, 2022, at the [Wayback Machine](#).

———. 2019a. *CRASHOVERRIDE: Reassessing the 2016 Ukraine Electric Power Event as a Protection-Focused Attack*. Dragos.

<https://pylos.co/wp-content/uploads/2021/02/crashoverride.pdf>. Archived May 3, 2026, at the [Wayback Machine](#).

———. 2019b. *Evolution of ICS Attacks and the Prospects for Future Disruptive Events*. Dragos.

<https://pylos.co/wp-content/uploads/2021/02/Evolution-of-ICS-Attacks-and-the-Prospects-for-Future-Disruptive-Events-Joseph-Slowik-1.pdf>. Archived December 17, 2022, at the [Wayback Machine](#).

- . 2019c. *Stuxnet to CRASHOVERRIDE to TRISIS: Evaluating the History and Future of Integrity-Based Attacks on Industrial Environments*. Dragos.
<https://pylos.co/wp-content/uploads/2021/02/Past-and-Future-of-Integrity-Based-ICS-Attacks.pdf>. Archived March 4, 2026, at the [Wayback Machine](#).
- . 2020. “SolarWinds Incident: Deeper Analysis of the Breach.”
<https://www.domaintools.com/blog/continuous-eruption-further-analysis-of-the-solarwinds-supply-incident>. Archived May 8, 2026, at the [Wayback Machine](#).
- . 2021. *The Baffling Berserk Bear: A Decade's Activity Targeting Critical Infrastructure*. Gigamon (Virus Bulletin 2021). <https://vblocalhost.com/uploads/VB2021-Slowik.pdf>. Archived June 4, 2026, at the [Wayback Machine](#).
- . 2022a. *Exorcising the Ghost in the Machine: Debunking Myths around Supply Chain Intrusions*. Gigamon.
<https://www.gigamon.com/content/dam/resource-library/english/white-paper/wp-ghost-in-the-machine-supply-chain.pdf>. Archived February 18, 2026, at the [Wayback Machine](#).
- . 2022b. “Industroyer2 in Perspective.” Slowik.
<https://pylos.co/2022/04/23/industroyer2-in-perspective/>. Archived April 17, 2026, at the [Wayback Machine](#).
- . 2022c. “Thrice Is Nice: Ukraine In Review.” Slowik.
<https://www.slideshare.net/slideshow/thrice-is-nice-ukraine-in-review/252544998>. Archived March 5, 2026, at the [Wayback Machine](#).
- . 2022d. *Zeroing in on Xenotime: Analysis of the Entities Responsible for the Triton Event*. Virus Bulletin Conference.
<https://pylos.co/wp-content/uploads/2022/10/Day1-1400-Green-Zeroing-in-on-XENOTIME-analysis-of-the-entities-responsible-for-the-Triton-event.pdf>. Archived December 17, 2022, at the [Wayback Machine](#).
- Smeets, Max. 2022. *No Shortcuts*. Hurst Publishers.
<https://www.hurstpublishers.com/book/no-shortcuts/>. Archived March 5, 2026, at the [Wayback Machine](#).
- Sobczak, Blake. 2019. “Experts assess damage after first cyberattack on U.S. grid.” E&E News by POLITICO.
<https://www.eenews.net/articles/experts-assess-damage-after-first-cyberattack-on-u-s-grid/>. Archived June 26, 2022, at the [Wayback Machine](#).
- Soltan, Saleh, Prateek Mittal, and H. Vincent Poor. 2018. “BlackIoT: IoT Botnet of High Wattage Devices Can Disrupt the Power Grid.” In *27th USENIX Security Symposium (USENIX Security 18)*, 15–32. Baltimore, MD, USENIX Association.
<https://www.usenix.org/conference/usenixsecurity18/presentation/soltan>. Archived February 20, 2026, at the [Wayback Machine](#).
- Stankovski, Andrej, et al. 2023. “Power blackouts in Europe: Analyses, key insights, and recommendations from empirical evidence.” *Joule* 7 (11): 2468–2484.
<https://doi.org/10.1016/j.joule.2023.09.005>.

- Stock, James H., and Mark W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." https://www.princeton.edu/~mwatson/papers/Stock_Watson_JASA_2002.pdf. Archived December 5, 2024, at the [Wayback Machine](#).
- Stouffer, Keith, et al. 2023. *Guide to Operational Technology (OT) Security*. NIST SP 800-82r3. National Institute of Standards and Technology. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-82r3.pdf>. Archived June 9, 2026, at the [Wayback Machine](#).
- Tan, Rui, et al. 2017. "Modeling and Mitigating Impact of False Data Injection Attacks on Automatic Generation Control." *IEEE Transactions on Information Forensics and Security* 12 (7): 1609–1624. <https://doi.org/10.1109/tifs.2017.2676721>.
- Task Force. 2023. *Ukrainian Energy Sector Evaluation and Damage Assessment -- X: As of May 24, 2023*. https://www.energycharter.org/fileadmin/DocumentsMedia/Occasional/2023_05_24_UA_sectoral_evaluation_and_damage_assessment_Version_X_final.pdf. Archived February 19, 2026, at the [Wayback Machine](#).
- Taylor, Angus. 2018. "Australian Government Attribution of the 'NotPetya' Cyber Incident to Russia." <https://www.dfat.gov.au/sites/default/files/australia-attributes-notpetya-malware-to-russia.pdf>. Archived October 25, 2025, at the [Wayback Machine](#).
- Tetlock, Philip E., and Dan Gardner. 2015. *Superforecasting: The Art and Science of Prediction*. Crown. <https://www.penguinrandomhouse.com/books/227815/superforecasting-by-philip-e-tetlock-and-dan-gardner/>.
- The Brattle Group. 2024. *Review of Value of Lost Load in the ERCOT Market: Value of Lost Load Study Final Report*. PUCT Project No. 55837. The Brattle Group. <https://www.brattle.com/wp-content/uploads/2024/09/Value-of-Lost-Load-Study-for-the-ERCOT-Region.pdf>. Archived February 20, 2026, at the [Wayback Machine](#).
- Thunder Said Energy. 2024. "Reserve Margins: By ISO and over Time?" Thunder Said Energy. <https://thundersaidenergy.com/downloads/reserve-margins-by-iso-and-over-time/>. Archived June 24, 2026, at the [Wayback Machine](#).
- Times of Israel. 2022. "Gantz orders probe after TV reports hint IDF behind Iran steel plant cyberattack." *Times of Israel*. <https://www.timesofisrael.com/gantz-orders-probe-after-tv-reports-hint-idf-behind-iran-steel-plant-cyberattack/>. Archived June 24, 2026, at the [Wayback Machine](#).
- Tomes, George. 2024. "Electric Grid Cyberattack: An AI-Informed Threat Model." LessWrong. <https://www.lesswrong.com/posts/zc5uhndCoxEKZvXoQ/electric-grid-cyberattack-an-ai-informed-threat-model>. Archived April 15, 2026, at the [Wayback Machine](#).

- Toner, Helen. 2025. “Nonproliferation is the wrong approach to AI misuse.” <https://helentoner.substack.com/p/nonproliferation-is-the-wrong-approach>. Archived June 24, 2026, at the [Wayback Machine](#).
- Tsvetanov, Tsvetan, and Srishti Slaria. 2021. “The effect of the Colonial Pipeline shutdown on gasoline prices.” *Economics Letters* 209: 110122. <https://doi.org/10.1016/j.econlet.2021.110122>.
- U.S. Department of Homeland Security, Office of Cyber, and Infrastructure Analysis. 2015. *Nuclear Reactors, Materials, and Waste Sector Cyberdependencies*. <https://nsarchive.gwu.edu/sites/default/files/documents/2841137/Document-09.pdf>. Archived June 14, 2024, at the [Wayback Machine](#).
- U.S. Department of Justice. 2018. *United States v. Netyksho et al.* Indictment, No. 1:18-cr-00215, U.S. District Court for the District of Columbia. <https://cdn.cnn.com/cnn/2018/images/07/13/gru.indictment.pdf>. Archived June 25, 2026, at the [Wayback Machine](#).
- U.S. Department of Justice, Office of Public Affairs. 2020. “Six Russian GRU Officers Charged in Connection with Worldwide Deployment of Destructive Malware and Other Disruptive Actions in Cyberspace.” <https://www.justice.gov/archives/opa/pr/six-russian-gru-officers-charged-connection-worldwide-deployment-destructive-malware-and>. Archived February 4, 2025, at the [Wayback Machine](#).
- U.S. Energy Information Administration. 2023. “Use of Electricity.” U.S. Energy Information Administration. <https://www.eia.gov/energyexplained/electricity/use-of-electricity.php>. Archived June 30, 2026, at the [Wayback Machine](#).
- UA Gov. 2023. *Demographic Situation in 2021*. State Statistics Service of Ukraine. <https://stat.gov.ua/en/publications/demographic-situation-2021>. Archived February 27, 2026, at the [Wayback Machine](#).
- UK AISI. 2024. *Advanced AI Evaluations at AISI: May Update*. AI Security Institute. <https://www.aisi.gov.uk/blog/advanced-ai-evaluations-may-update>. Archived May 31, 2026, at the [Wayback Machine](#).
- . 2026. *How Fast Is Autonomous AI Cyber Capability Advancing?* AI Security Institute. <https://www.aisi.gov.uk/blog/how-fast-is-autonomous-ai-cyber-capability-advancing>. Archived June 16, 2026, at the [Wayback Machine](#).
- UK Foreign Office. 2018. “Foreign Office Minister Condemns Russia for NotPetya Attacks.” GOV.UK. <https://www.gov.uk/government/news/foreign-office-minister-condemns-russia-for-notpetya-a-attacks>. Archived February 16, 2018, at the [Wayback Machine](#).
- UK MoD. 2022. *Cyber Primer (3rd Edition)*. UK Ministry of Defence. https://assets.publishing.service.gov.uk/media/63623df5d3bf7f04e12196d0/Cyber_Primer_Edition_3.pdf. Archived August 8, 2025, at the [Wayback Machine](#).

United States Government Accountability Office. 2019. *Critical Infrastructure Protection: Actions Needed to Address Significant Cybersecurity Risks Facing the Electric Grid*.

<https://www.gao.gov/assets/gao-19-332.pdf>. Archived March 17, 2026, at the [Wayback Machine](#).

United States v. Clendaniel and Russell. 2023. *Affidavit in Support of Criminal Complaint*, No. 23-mj-00401-MJM. U.S. District Court for the District of Maryland.

<https://storage.courtlistener.com/recap/gov.uscourts.mdd.530116/gov.uscourts.mdd.530116.2.0.pdf>.

United States v. Fathi et al.: Sealed Indictment. 2016.

<https://www.justice.gov/archives/opa/file/834996/dl?inline=>. Archived March 18, 2026, at the [Wayback Machine](#).

Vicens, AJ. 2022. “Russian hackers thwarted in attempt to take out electrical grid, Ukrainians say.” CyberScoop.

<https://cyberscoop.com/ukrainian-electrical-grid-industroyer2-russia-sandworm/>. Archived July 5, 2026, at the [Wayback Machine](#).

Washington Post. 2024. “Mossad’s Pager Operation: Inside Israel’s Penetration of Hezbollah.” The Washington Post.

<https://www.washingtonpost.com/world/2024/10/05/israel-mossad-hezbollah-pagers-nasralah/>. Archived April 30, 2025, at the [Wayback Machine](#).

Wikipedia. n.d.. “Moore County Substation Attack.”

https://en.wikipedia.org/wiki/Moore_County_substation_attack. Archived June 8, 2026, at the [Wayback Machine](#).

Willuhn, Marian. 2022. “Satellite cyber attack paralyzes 11GW of German wind turbines - pv magazine Global.” pv magazine Global.

<https://www.pv-magazine.com/2022/03/01/satellite-cyber-attack-paralyzes-11gw-of-german-wind-turbines/>. Archived March 1, 2022, at the [Wayback Machine](#).

Wing, Ian Sue, et al. 2025. “A Method to estimate the economy-wide consequences of widespread, long duration electric power interruptions.” *Nature Communications* 16 (1): 3335.

<https://doi.org/10.1038/s41467-025-58537-4>.

World Bank. 2025. “World Bank Open Data.” World Bank.

<https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>. Archived June 24, 2026, at the [Wayback Machine](#).

———. 2026a. “GDP per capita (current US\$) - Ukraine.” World Bank Open Data.

<https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=UA>. Archived April 21, 2026, at the [Wayback Machine](#).

———. 2026b. “GDP per capita (current US\$) - United States.” World Bank Open Data.

<https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=US>. Archived July 5, 2026, at the [Wayback Machine](#).

———. n.d.. “Population, total - Ukraine.” World Bank Open Data.

<https://data.worldbank.org/indicator/SP.POP.TOTL?locations=UA>. Archived May 26, 2026, at the [Wayback Machine](#).

xAI. 2025. xAI Risk Management Framework. xAI.
<https://data.x.ai/2025-08-20-xai-risk-management-framework.pdf>. Archived July 12, 2026 at the [Wayback Machine](#).

Zachariadis, Theodoros, and Andreas Poullikkas. 2012. “The costs of power outages: A case study from Cyprus.” *Energy Policy* 51: 630–641. <https://doi.org/10.1016/j.enpol.2012.09.015>.

Zeller, Mark. 2011. *Common Questions and Answers Addressing the Aurora Vulnerability*. Schweitzer Engineering Laboratories.
https://cdn.selinc.com/assets/Literature/Publications/Technical%20Papers/6467_CommonQuestions_MZ_20101209_Web.pdf. Archived June 24, 2026, at the [Wayback Machine](#).

Zetter, Kim. 2016. “Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid.” WIRED.
<https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/>. Archived May 31, 2026, at the [Wayback Machine](#).



GovAI