Speaker Notes: "The New Nuclear: How the AI Race Will Rewrite the Rules of Global Conflict"
Presented by Daria Bahrami | Head of Policy, Dreadnode

[Slide One]

We have a lot of ground to cover in the next 25 minutes, so I'm going to offer this disclaimer up front: this is a spicy talk and if we had infinite time, I could couch a lot of what I'm about to say in more precise, descriptive language. But this session is the only thing standing between you and the reception, so we're going to enjoy the jalapeños and ghost peppers and get through this together.

[Slide Two]

Think back to 2023. Anyone here remember or experience the cultural phenomenon known as Barbenheimer?

[Slide Three]

2023 gave us this amazing cultural moment. Two completely different films, released on the same day, and the vibes could not be more different. Yet both somehow tackled existentialism.

Oppenheimer showed us the weight of nuclear power in a very intensely crafted drama film. Barbie tackled the complexities of gender identity and the human experience — a reimagination of 'what we're made for', with a lot more pink.

This contrast matters. Nuclear proliferation gave us a ceiling for kinetic warfare. Mutually Assured Destruction works because everyone understands that crossing certain lines means total annihilation. We respect that boundary. And yet, the digital domain lacks the same norms of behavior and standards we've deemed necessary to de-escalate and manage traditional. This is the difference between kinetic and cyber warfare: we can see the damage inflicted by the former with the naked eye, and we can't necessarily do the same for the latter.

[Slide Four]

In cyber warfare, we often treat strategy with a light touch. Yet cyber operations are threatening to dismantle the foundations upon which we've developed laws of armed conflict. Because cyber strategy is often absent from political dialogue, we're not seeing the level of institutional support required for this threat landscape to be managed properly.

And AI just made this exponentially more urgent.

We have a demonstrated history of creating guardrails on warfare. There is an incentive to avoid crossing lines we can't come back from. Mutually assured destruction is one of them, but let's dig a little deeper.

Look at the pattern.

Speaker Notes: "The New Nuclear: How the AI Race Will Rewrite the Rules of Global Conflict"
Presented by Daria Bahrami | Head of Policy, Dreadnode

[Slide Five]

Disastrous moments in history have compelled leaders to uniformly decide that there are certain crimes against humanity that we cannot afford to tolerate.

- The Lieber Code emerged after Civil War brutality, and marked the first time we said "maybe we shouldn't kill prisoners."
- The Geneva Conventions surfaced after 40,000 soldiers died without medical care at Solferino.
- We banned chemical weapons after witnessing mustard gas in World War I.
- And of course, the Holocaust prompted comprehensive civilian protections.

[Slide Six]

Every framework emerged from horror, and four core principles emerged: military necessity, humanity, distinction, and proportionality.

- **Principle of Military Necessity** - only use force needed for legitimate objectives, i.e. to defeat the enemy quickly and effectively
- **Principle of Humanity** - don't inflict unnecessary suffering, injury, or destruction on combatants and non-combatants
- **Principle of Distinction** - separate combatants from civilians, and direct attacks only at military targets, rather than civilian lives or property.
- **Principle of Proportionality** – don't pursue attacks expected to cause civilian harm that exceeds the military advantage

These gave us the language of warfare - clear rules for offense and defense. But these core principles are designed for kinetic warfare, where tools and targets are relatively static and observable. On the other hand, cyber operations break nearly ALL of these frameworks. And we haven't built new ones to replace them.

[Slide Seven]

Let me show you why traditional frameworks don't work.

- **Distinction is dead, or at least extremely difficult.** When you attack a power grid, who's the target? The military base that needs electricity? The hospital? The water treatment plant? All of the above? The line between civilian and military infrastructure doesn't exist in interconnected systems.

- **Proportionality is meaningless** when cascading effects are unpredictable. You can't calculate proportionality when you don't know if your attack will brick one system or cascade through fifty.
- **And attribution is deniable.** In kinetic warfare, you know who fired the missile. In cyber, attribution can take months or years - if you get it at all. We've experienced this recently.

[Slide Eight]

Take Stuxnet. 2010. Widely believed to be US and Israeli work - neither has confirmed it.

It physically destroyed hundreds of Iranian centrifuges and set back their nuclear program by years. This functionally served as a bombing raid conducted with code that achieved offensive cyber objectives for defensive reasons (i.e. nuclear nonproliferation). And the attacker never had to own it.

By traditional warfare standards, this was proportional, necessary, and legitimate. But it's over a decade later. No official attribution. No public framework. No declared rules of engagement.

This model worked when cyber operations were rare, sophisticated, and nation-state level.

But now? Criminal gangs can achieve nation-state effects and civilians are suffering.

[Slide Nine]

I need to be clear about what's at stake here. There's no obvious peacetime anymore. Attacks on civilian infrastructure happen constantly: sometimes by nation-states, sometimes by criminals, and sometimes we can't tell the difference. And **civilians bear the immediate cost.** Let me show you what that looks like in practice.

- **Düsseldorf, Germany, 2020.** Ransomware shut down a hospital. A woman with an aortic aneurysm died during emergency transfer to another facility. First documented death linked to a cyberattack.
- **Change Healthcare, US, 2024.** Ransomware affecting a third of US medical records. Cancer patients couldn't get chemotherapy authorization because of a billing system hack.

Even insurance companies know code kills. They sell cyber liability policies covering **Contingent Bodily Injury** from hacked medical devices, infrastructure collapse, and critical system failures.

The market has priced in that code kills. Our policy frameworks haven't.

[Slide Ten]

August 2025. St. Paul, Minnesota. Ransomware attack by a group called Interlock disrupted city systems, forced a mass password reset for employees, and led to the public release of 43 gigabytes of data after the city refused to pay ransom.

The Governor deployed the Minnesota National Guard to assist in recovery efforts.

Think about that. A financially-motivated criminal gang forced a military response.

Now here's the question: **What happens next?**

CISA, the FBI, and HHS had issued warnings about Interlock days before. But CISA can't compel action in this scope—they can only advise. When the attack happened, Minnesota mobilized as many resources as they could tap into. But the recovery time still took about seven weeks.

The St. Paul example is probably one of the optimal response scenarios in that local, state, and even federal government recognized the urgency of the situation and mobilized resources as quickly as possible.

[Slide Eleven]

August 2025. St. Paul, Minnesota. Ransomware attack by a group called Interlock disrupted city systems, forced a mass password reset for employees, and led to the public release of 43 gigabytes of data after the city refused to pay ransom.

The Governor deployed the Minnesota National Guard to assist in recovery efforts.

Think about that. A financially-motivated criminal gang forced a military response.

Now here's the question: **What happens next?**

CISA, the FBI, and HHS had issued warnings about Interlock days before. But CISA can't compel action in this scope—they can only advise. When the attack happened, Minnesota mobilized as many resources as they could tap into.

In fact, the St. Paul example is probably one of the optimal response scenarios in that local, state, and even federal government recognized the urgency of the situation and mobilized resources as quickly as possible.

But the recovery time still took about seven weeks.

[Slide Twelve]

Seven weeks to recovery is a demonstration of how we got off easy, knowing what emerging capabilities are making the rounds. The lag time here was not due the sophistication of the cyberattack, but due to the lack of resourcing necessary to respond. So what happens when the threats move faster, demonstrate more sophistication, and can establish persistence in the targeted network?

[Slide Thirteen]

We saw demonstrations of automated vulnerability and exploitation when threat actors weaponized AI-powered penetration testing modeled by Hexstrike-AI. Hexstrike's research highlighted a vulnerability exploitation attack chain that reduced the time from zero-day disclosure to successful exploit from days to **under 10 minutes.** Within hours of release, threat actors used it to exploit complex Citrix NetScaler vulnerabilities among others. Our patch cycle is usually measured in weeks. Exploitation is now measured in minutes.

We've also seen instances of polymorphic malware evasion. BlackMamba - a proof-of-concept from HYAS Labs - demonstrated AI-generated polymorphic malware. It reaches out to OpenAI's API at runtime to dynamically generate keylogger code. Every execution creates entirely new code with identical functionality. Tested against leading EDR solutions: Zero detections. Zero alerts. Say goodbye to traditional defenses that rely on signatures and patterns. AI-powered malware rewrites itself nearly every time.

[Slide Fourteen]

Seven weeks to recovery is a technology optimization problem. How much time are we losing by subjecting security operations analysts to manual log analysis, data correlation, tier 1 responses, or vulnerability management processes. We can get time back by developing and refining AI-powered cyber capabilities. I'm talking about automated vulnerability discovery and remediation, automated log analysis—manual efforts that can free up security engineer time to focus on the more interesting breadcrumbs.

[**Slide Fifteen]**

AIxCC demonstrated the potential for AVDR and incentivized companies across the country to focus on that use-case. Since AIxCC launched, Vulnhuntr released in October 2024: an AI tool that autonomously discovers zero-day vulnerabilities. In just a few hours, it found over a dozen remotely exploitable zero-days in popular open-source projects. Including full remote code execution.

Google's Big Sleep AI system discovered a zero-day vulnerability in SQLite, blocking hackers from exploiting this previously undisclosed vulnerability before threat actors could exploit it.

We are seeing an opportunity to build AI-powered solutions that can exponentially speed up threat detection and subsequent patching processes. Stack that against an OODA loop process and let me know what's faster. I'll wait.

We need to automate both defensive responses AND offensive responses within pre-approved rules of engagement.

[Slide Sixteen]

Seven weeks to recovery is a resourcing problem. We can reduce the inherent delays from "emergency deployments" by having teams pre-positioned. Yes, this term has mostly been referenced in the context of how threat actors are pre-positioning adversarial capabilities in our critical infrastructure. But we can flip this narrative in our favor as well.

We have an opportunity to authorize existing federally resourced cyber capabilities teams to provide continuous support at the state and local levels, to start. State and local entities are prime targets for adversarial attacks, and we can begin to set them up for success by establishing formal coordination between CISA and the National Guard under Title 32. I touch on these concepts in one of my recent blog posts: https://dreadnode.io/blog/from-compute-to-congress-to-address-cisas-authority-gap-reauthorize-cisa-2015-and-slcgp

[Slide Seventeen]

Your security operations center is drowning. Manual log analysis, chasing alerts, vulnerability management, unsustainable hours.

And now they're facing attackers who discover vulnerabilities in hours, exploit in minutes, and evade detection with self-modifying code.

This is not sustainable.

We need to **automate the tedious parts:** Log analysis, initial vulnerability assessment, pattern recognition, first-tier triage.

This isn't about replacing analysts. It's about giving them something interesting to do. Let machines handle repetitive work. Let humans focus on threat hunting, complex threat intelligence analysis, and tier 2 escalations.

Because **AI enables offense just as much as defense.** Autonomous reconnaissance at scale. Adaptive exploitation. Real-time target analysis.

[Slide Eighteen]

Seven weeks to recovery and the surrounding communications since sent an underwhelming message that we "tolerate" cyberattacks on critical infrastructure. Interlock just attacked critical infrastructure. They're still operating. Their infrastructure is still live. They will continue to regroup and deploy attacks as long as they can.

And so far, the message we've sent (publicly) is this: we're going to move on and quietly handle this behind closed doors. This isn't about real-time recovery efforts or what the FBI is up to behind closed doors. This is about public perception because this is officially a public concern. Cyber has physical ramifications, it does impose civilian harm. And that is undeniably a public interest issue that will affect votes and who's in power and what legislation passes.

[Slide Nineteen]

Here's the bottom line: **We're running human-speed processes in a machine-speed world.**

We have an opportunity to raise the stakes and make cyber-attacks on critical infrastructure more costly. We can make it less enticing, riskier, and that much harder to successfully bully businesses into handing over money or resources or data.

Interlock attacked a major US city. We deployed military assets in response. We need a clear strategy for balanced defensive-offensive cyber operations when we are attacked at scale.

Let me be explicit about what I'm proposing, because in the policy community, offensive cyber is currently translating to "hack back" and "cyber letters of marque" and this is all very loaded.

Absent the time and the necessary nuance to unpack these "tools," I'm NOT here to talk about:

- Vigilante cyber operations, uncontrolled or escalatory retaliation, or even cyber letters of marque.

But even in kinetic warfare, we have gained more public support for operations that demonstrate a clear political objective in response to a perceived threat. I'm talking about **deliberate cyber operations with the same defensive or political objectives, conducted under clear legal authority.** We have used the same rationale in kinetic warfare because we understood early on that we don't tolerate acts of terror or violence against our own.

And we can do so digitally in the near-term by leveraging existing institutional powers.

- **Elevate Response Posture (Title 32):** Immediately leverage existing Title 32 authorities for the National Guard to close CISA's authority gap and provide federally-supported resources for rapid, CISA-coordinated multi-state cyber responses.
- **Enforce Secure-by-Design AI Integrations:** Mandate inherent resilience by creating optimized, use-case specific AI-enabled cybersecurity evaluations and procurement benchmarks. Coupled with adversarial stress testing, these frameworks can push vendors toward building 'Secure-by-Design' technology from the start.
- **Automate Compliance (Policy-as-Code):** Convert slow, administrative compliance requirements into "policy-as-code" using strategic, machine-readable automations. This makes governance scalable, verifiable, and ensures real-time compliance across complex systems. You can read more about this in Dreadnode's latest RFI response: https://dreadnode.io/blog/dreadnode-response-to-the-2025-regulatory-reform-for-artificial-intelligence
- **Close the Detection Gap (Advanced Reconnaissance):** Embed advanced reconnaissance and detection capabilities to meet the tempo of threats. This requires constant observation and adaptive governance protocols to stay ahead of rapidly evolving AI-powered threats and Agentic AI systems.

We don't tolerate offensive kinetic attacks outside a defensive framework or narrative. We shouldn't treat cyber differently.

Stuxnet had the right objective - prevent nuclear proliferation - but no public framework.

St. Paul response should have a framework - defend critical infrastructure through both defensive hardening AND offensive action against the threat actor.

This is not controversial for kinetic warfare. It shouldn't be controversial for cyber.

[Slide Twenty]

Private AI labs—OpenAI, Anthropic, Google, Meta, just to name a few—are developing and investing in machine-speed solutions that can help map out the threat landscape faster and more efficiently than we've been able to thus far.

And as we all know, most of these AI models are commercially available, open-weigh, or even open-source. So what's the civilian-combatant distinction in this context? We have to bake that in and get over our discomfort of having military proximity to commercial offerings.

This is where institutional support in favor of secure-by-design AI systems that can support cyber resiliency across critical use cases.

- **We need fine-tuned, classified AI systems for government cyber operations.** Models evaluated against specific use cases - offensive cyber, defensive analysis, threat intelligence - with procurement standards optimized for national security.
    - NIST evaluation standards
    - AI testbeds through DOE National Labs at a minimum
    - OMB secure-by-design federal procurement standards for AI systems
- **What are these use cases? Incentives for AI-powered cyber operations** to automate whatever degree of manual effort can free up security engineering time for more complex, engaging problem-solving.
    - Automate: Log analysis, vulnerability assessment, threat intelligence, first-tier response.
    - Free up analysts for: Threat hunting, strategic analysis, offensive operations, framework development.

The clock is ticking. We can't wait for the next St. Paul or an attack on military infrastructure where we lack clear offensive authority.

Integrated Defensive-Offensive Capabilities

The architecture for many of the foundational policies we need are already in place, which Dreadnode Policy partially unpacks in this blog post:

- The Cyber Information Sharing Act of 2015 enables private entities to disclose cyberattack details to government stakeholders with amplified liability protections. Also known as CISA 2015, this act is up for a clean reauthorization in Congress as we speak—this time as the PACT Act.
- The State and Local Cybersecurity Grant Program promotes funding and support for state and local cyber capability development. While this expired on September 30, 2025, the U.S. Government has the power to renew this funding authority.
- And CIRCIA gives CISA mandatory reporting authority. If you're critical infrastructure and you're breached, we need to know immediately and we need to be able to act just as quickly. CISA is slated to turn into final rule in May 2026.

- Read more on Dreadnode's response to AI Regulatory Reform to learn how strategic compliance automations can optimize CIRCIA and deduplicate parallel defense-related incident reporting requirements.

To that end, there has to be a return on that information investment, which we can more easily realize by leveraging Title 32 to formalize CISA-National Guard coordination. When critical infrastructure is attacked, treat it as national security from day one.

Together, these create **integrated defensive capability at speed.** We need a national cyber strategy, to include legally covered defensive and offensive capabilities that would have optimized the St. Paul case study. This national strategy needs to account for:

- An attribution framework,
- A designation authority,
- An operational authority, and
- Clear rules of engagement.

[Slide Twenty-One]

We spent decades learning lessons about kinetic warfare the hard way. We've already learned these lessons and understood the need for frameworks to manage and navigate warfare. We saw mustard gas and placed heavy restrictions on chemical weapons. We saw nuclear weapons and built MAD doctrine.

We know what the cyber ceiling is. It's when we can't recover. When we lose access to critical systems and can't reclaim sovereign digital space, that primes that space for geopolitical negotiating power. Cyber deserves the same treatment: **institutional frameworks** that legitimize both defense and offense, with clear authorities, clear attribution, clear rules of engagement.

Here's what I want you to do:

If you're in the private sector: I urge you to consider how the foundation of information sharing, led by CISA and other voluntary cooperation model agencies, could support your organization's security. These threat advisories shouldn't be regarded as optional, and we need operational support to back up the solutions here. Advocate for integrating your defensive operations with government cyber units under Title 32 coordination. Continue to call out what's not working, but be part of the solution.

Speaker Notes: "The New Nuclear: How the AI Race Will Rewrite the Rules of Global Conflict"
Presented by Daria Bahrami | Head of Policy, Dreadnode

If you're in government: We need to drive conversations that will outline cyber operational capabilities at scale. The civilian sector has already been militarized in many ways. Just because the threat actors are "unaffiliated or unattributed" doesn't change the impact. Push for operationalizing nationally distributed cyber capabilities (i.e. by way of Title 32 + CIRCIA) for cyber resilience and response, AND clear ROE for offensive cyber operations. We need the whole strategy.

If you're a red teamer: Your tradecraft is a national security asset. It needs institutional backing - clear authorities for offensive operations within defensive frameworks. Don't forget that this is your community too, and continue to advocate for your teams.

If you develop AI: Consider (also) building use-case specific, classified systems optimized for government use cases with procurement standards that can prioritize national security needs across critical infrastructure.

Everyone: Start treating cyber attacks on critical infrastructure as what they are: acts of violence that demand an institutional response and unified, national strategy

Let's stop giving cyber warfare a light touch. Let's give it the institutional treatment it deserves - not just building the capabilities, but building the institutions to manage them.

Because the alternative is waiting for the next horror to force us to build frameworks reactively.

Let's choose to build them now, on our own terms.

[Slide Twenty-Two]

As always—don't hesitate to reach out with any questions or constructive feedback.