Briefing Summary

CISO Roundtable

AI and Cybersecurity: A CISO's Guide to Emerging Threat Detection and Response Applications

May 2023

Special thanks to our Research Partner:
ThreatWarrior

TRUSTED NETWORKS.
ADVANCING CYBER STRATEGIES.

ADVANCED CYBER SECURITY CENTER

Artificial Intelligence has taken the technology world by storm. Its impact on cybersecurity - both good and bad - is just beginning. Our most recent roundtable leveraged key thought leaders including **Pete Slade, Co-Founder and CTO of ThreatWarrior, Dr. Andres Molina-Markham, MITRE, and Colin Zick and Chris Hart from Foley Hoag**.

90 minutes only gave us time to skim the surface. This briefing summary is packed with insights and questions that are sure to be part of public and private discourse for months to come. To that end, this roundtable is a mere precursor to a deeper dive in our upcoming November 8th in-person event. Buckle your seatbelts. This one is a whopper!

There is reason to be bullish about AI in Cybersecurity!

- **Unsupervised machine learning will change the game of cybersecurity.** There is a nearly infinite amount of data that can be consumed for security analysis. And yet, give an analyst a data stream of every network connection and they won't be able to see anything. Add every firewall block, Windows event, DNS query and they'll see even less. In contrast, the more you feed into an ML algorithm, the better it models data. Legacy solutions can catch many problems by looking for known threat signatures. First generation ML did a little better using universal models and classifying against libraries of known threats. But it's still fuzzy matching on historical data. What do you do about things that these approaches miss? This is where machine learning comes in as unsupervised anomaly detection. It is a serious force multiplier.

- **AI will quiet today's noisy anomaly detection.** The state of the art for mitigating noise is to monitor multiple data feeds. When anomalous activity spans various feeds, stronger signals can be generated. Small changes in behavior - such as allowing external connections through a firewall to an unusual IP, running a process with an unusual name, downloading an unusual amount of data from a server (even one you are allowed to access), or leaving your computer running at times - are all situations that may trigger an independent alert. Alone, they are pure noise. Tied together, they paint a strong signal that an incident is occurring. This brings alert triage to an entirely new level, and gives the good guys a fighting chance at finally reducing dwell time - as you are escalating and contextualizing known threats that would otherwise be missed. Again, this is a force multiplier.

Contact us

(617) 584-0581

jdinneen@acscenter.org

- **The use cases are nearly endless.** Examples include: misconfigurations that allow dubious DNS server comms; external IPs contacting an internal IP over ICMP; unauthorized vulnerability scans. Viewed on their own these would often be considered benign, again if noticed at all. AI shines a light on the "every-dayness" that clever attackers rely upon to do their bidding.
- **Security jobs aren't going away.** Today's AI won't replace humans in cybersecurity. It rather allows humans to do more. Think "augmented intelligence." No SOC team has enough resources to do its job as well as it would like. Humans still have the extra knowledge necessary to interpret the world, e.g., people, politics, bureaucracy, infrastructure, and cause and effect. AI tools don't replace this. AI tools do, however, let your workers see further, act faster, and defend more of your threat surface.

At the same time, we need to be eyes wide open.

- **AI is the latest "glom on" term.** While generative AI and large language models (LLMs) are all the rage at present, some security vendors have been working diligently on integrating and building products that consume AI for years. But buyer beware! Every vendor is going to tell you they have AI and ML. Definitions are very loose.
- **Not all will benefit.** The organizations we see struggling the most with AI are the ones that haven't invested properly in their cybersecurity workforce. The better the team, the closer they are to operations. The closer to operations, the more they will effectively utilize AI.
- **The bad guys have it too.** AI can, and will, accelerate learning and software development for attackers just as with defenders. Spear phishing and pretexting attacks will become much harder to detect. Phishing and spam will become much harder to catch as messages will be more natural and organic.

Contact us

(617) 584-0581

jdinneen@acscenter.org

- **And it won't just be "traditional" security attacks.** Misinformation will become harder to refute - and spread faster. Supply chain attacks - where a vendor sends out a trained model, but it gets intercepted and replaced with a malicious stand-in - will happen. Man-in-the-middle attacks for large language model services, like open AI, where API requests will be intercepted will happen. Published models that are reverse engineered will lead to all kinds of havoc. Prompt injection attacks that trick LLMs into doing arbitrary unauthorized work for users will occur.
- **Expensive processing power will not be a malevolent barrier to entry.** Today, training an LLM is exorbitantly expensive and requires massive GPUs to run. So its development is largely in the hands of elite organizations with access to sumptuous cloud compute resources and a presumed moral compass. But LLaMA trained weights have already leaked, and soon powerful capabilities will be rendered on inexpensive, abundant, and uncontrollable consumer hardware. The genie is way out of the bottle, and widespread nefarious use will certainly occur.
- **Adversarial prompting and command chaining will create new classes of threats.** AgentGPT provides the ability to search for a vulnerability, query historic baseline behavior, and retrieve related events. A model can now scavenge threats and assemble those needed for a highly-targeted campaign. These new threats will be extremely difficult to detect without unsupervised learning.

Embrace it you must. Question you should.

- **Are privacy, ethics, and security concerns well understood?** Not by a long shot. Privacy is largely centered around the willing or naive insertion of confidential and/or intellectual property information into open-source large language models (LLM) like ChatGPT, where that information then becomes usable by open source deep learning algorithms. Ethical concerns were only discussed in the aggregate, but they include bias and discrimination,

Contact us
(617) 584-0581
jdinneen@acscenter.org

transparency and explainability, job displacement, surveillance and autonomy, environmental impact, and misuse/malicious use. The DoD, and the late Secretary Ash Carter, were mentioned as leaders on the ethical AI policy front. Security concerns typically revolve around privacy and data security. The recent Samsung incident was cited.

- **Should your organization block individual use of AI?** Some (but hardly a majority) major organizations are taking this step - notably Apple, Amazon and Samsung and Lockheed Martin. There is full recognition that AI use is difficult to enforce, but at least until such time as a use policy is established, blocking is viewed as a "safe haven" approach for corporate governance. Note the distinction between individual use (e.g., ChatGPT, AgentGPT, Stable Diffusion) versus corporate-sanctioned software products that have (and have had for years) embedded AI/ML capabilities.
- **Will vendor disclosure become standard fare?** At least one member indicated they are asking vendors how AI is, or is planned to be, incorporated into their products. This will get interesting when one considers that neural networks and deep learning lead to questions around "explainability."
- **Can you trust AI?** How can we know what AI tells us is right? First, virtually no company has "AI specialists" on staff. Rather, companies that are leveraging AI well are equipped with critical thinkers who are highly inquisitive. Where companies have invested these types of resources - with clear domain and industry knowledge - discoveries around work shortcuts, rules assessments, remediation strategies and more are advancing rapidly. But, these same organizations are aware you can't just blindly accept AI answers. AI can and will hallucinate, and therefore, must be kept in check.

Contact us

(617) 584-0581

jdinneen@acscenter.org

- **What is an AI policy if we can't control the controls?** Policy is often used to plug gaps where laws are missing or ill-formed. Let's say you have a control that says, "I will not let unauthorized software into production." But, if AI can build software, it can also build controls. Now, who has "policy control"? How can automated controls even be verified? We are currently worried about how to create better board level involvement around security governance and risk - which ultimately leads to controls. This is an entirely different level of control management.

- **Is AI-specific policy really even needed?** Some members argued that current privacy and security policies are sufficient, even for generative AI technology. It's possible the task-at-hand is more around helping people understand how current policies apply.

Policy or not, the legal landscape is evolving.

- **AI risk management is being defined into law.** The EU is currently leading the regulatory charge with its EU AI Act, which assigns applications of AI to three risk categories. At least one member said this legislation is worth following as, similar to GDPR, it could become a global standard.

- **Will we see an "AI Bill of Rights?"** There have been attempts by the Biden administration to create an AI bill of rights that has five components including privacy, notice, opt out, and safety. But, this too is early and still playing out.

- **Are there laws today that govern the use of AI?** Yes, there are current laws that govern "some aspects" of AI technology use. But legal participants advise that things are evolving. As far as laws exist, some are untested. As an example, there are privacy laws that govern when consent is required before using personal information. Privacy laws notwithstanding, recording technology is proliferating. Automatic note taking as a general

Contact us

(617) 584-0581

jdinneen@acscenter.org

purpose could "function as an exception." There is a question of what is legal versus what is good practice. The topic of cookies came into the discussion as an example where great lengths have been taken to acquire consent circa internet use tracking. And yet there are a number of class actions in Massachusetts and around the country that are testing party consent laws based on cookie use. A second area of legal concern is that of large dataset use for training ML algorithms. How can we, as humans, be sure bias is not introduced, e.g., AI applications that might rank employment or mortgage applicants on fuzzy logic factors that are not clearly understood or traceable?

- **Will copyright be protected?** Copyright protection is being tested right now in Northern California. Can AI applications use the works of others, in this case, artists, as part of a training dataset - without attribution, permission, or licensing - without violating copyright laws? It's actually less clear than one might think. A quick look at the Supreme Court decision involving Andy Warhol and pictures of prints reveals the complexity here. Then, there is the argument that AI/ML is evolving so rapidly that even if all license-protected art was removed from training data sets, AI will still only require a few months to begin rendering similar art on its own.

Steps You Can Take

- **Irrespective of open source, generative AI applications based on LLMs, you should embrace legitimate, embedded AI as part of your defense-in-depth.** Increase your understanding of solutions built upon true unsupervised learning. The benefits are clear and will only improve. You will not be able to scale through human talent acquisition, training and retention. Those who are successful will be the ones with teams that are prepared to exploit AI from a SOC augmentation perspective.

Contact us

(617) 584-0581

jdinneen@acscenter.org

- **Establish a policy for liability prevention.** From a legal standpoint, organizations should at least establish a standard of care. You don't want to be negligent. That means take the time to develop a reasonable policy. It doesn't have to be perfect, just reasonable. Then implement it with training. The recent situation where a lawyer (of all people) used ChatGPT to build a case - and ChatGPT in turn hallucinated case references - should be cause for alarm. With reckless abandon, the lawyer did not check its sources. Now the lawyer faces disciplinary action. Organizations should clearly establish enough policy to guard against liability risk created by its own employees.

- **Don't squelch curiosity.** Understandably, employees of all disciplines are going to be curious about AI. In particular, the need for security talent is well chronicled. So while policy control is needed, it can't be at the expense of killing curiosity. Leading organizations will need to strike a reasonable balance around acceptable use of AI technology. Code generation and vulnerability analysis is one thing. Development of legal documents is, perhaps, another.

- **Think carefully about your organization's adoption of AI.** Whether developed in-house, embedded into vendor-provided solutions, or via open/commercial apps at the fingertips of your employees, delve into the bounded and unbounded elements of how AI does what it does. Can your solution provider go into specifics around AI algorithm use? What are the processes for training a particular AI tool? What protection mechanisms are in place? What are the potential implications of misuse or vulnerability?

Contact us

(617) 584-0581

jdinneen@acscenter.org

ADVANCED CYBER SECURITY CENTER

Member Participants

- Bill Brown, Abacus Insights
- Tom Gregory, Abacus Insights
- Roman Brozyna, Commure
- Joshua Marker, Commure
- Lavonne Burke, Dell
- Ed Hagopian, Dell
- David Kitchen, Dell
- Christian Hamer, Elevate Bio
- Tom McDermott, EOTSS
- Cara Bradley, Everbridge
- Jenn Harding, Harvard University
- Michael Tran Duff, Harvard University
- Chris Hart, Foley Hoag LLP
- Colin Zick, Foley Hoag
- Dan Hoag, Jackson Laboratory
- Shahid Kayani, Jackson Laboratory
- Dianne Pacheco, Jackson Laboratory
- Jenn Cook, John Hancock
- Andres Molina-Markham, MITRE
- Deb Briggs, NetScout
- Tom Myers - MA General Counsel/CPO
- Neil Clauson, Mimecast
- Shana York, Mimecast
- Curt Heintz, MIT Lincoln Laboratory
- Scott Mancini, MIT Lincoln Laboratory
- Greg Brinkman, Munich Re
- Angela Homm, Munich Re
- Hubert Kirchgaessner, Munich Re
- John Schramm, Munich Re
- Bob Briggs, NCEP
- David Luzzi, Northeastern University
- Alexis Goltra, Northeastern University
- Kenneth Liddle, Northeastern University
- Christopher Perretta, Ind. Board Director
- John Parlee, Park Place Technologies
- Joel Jacobs, Phenomanati
- Cydnee Spillane, Point32Health
- Pat Ford, Schneider Electric
- Christine Whichard, SmartBear
- Dmitry Sergeev, SmartBear
- Don Anderson, Tiger Global
- Sohail Iqbal, Veracode
- Greg Bosworth, VHB

Research Partners

- Bruce Coughlin, ThreatWarrior
- Jonathon Rubin, ThreatWarrior
- Pete Slade, ThreatWarrior
- Peter Quirk, ThreatWarrior

Special Thanks to Presenters

- Pete Slade, Co-Founder and CTO of ThreatWarrior
- Jon Rubin, AI Lead, ThreatWarrior
- Dr. Andres Molina-Markham, MITRE
- Colin Zick, Foley Hoag
- Chris Hart, Foley Hoag

Event Sponsor

THREAT WARRIOR

Contact us

(617) 584-0581

jdinneen@acscenter.org