SUREFIRE CYBER

The Hartford

AI 101: Understanding the Fundamentals of Artificial Intelligence

The Hartford

Jennifer Birnbaum, Cyber Underwriting Officer

Daniel Silverman, Underwriting Director, Specialized CyberTech

Keith Tagliaferri, Director, Cyber Claim Practices

SureFire Cyber

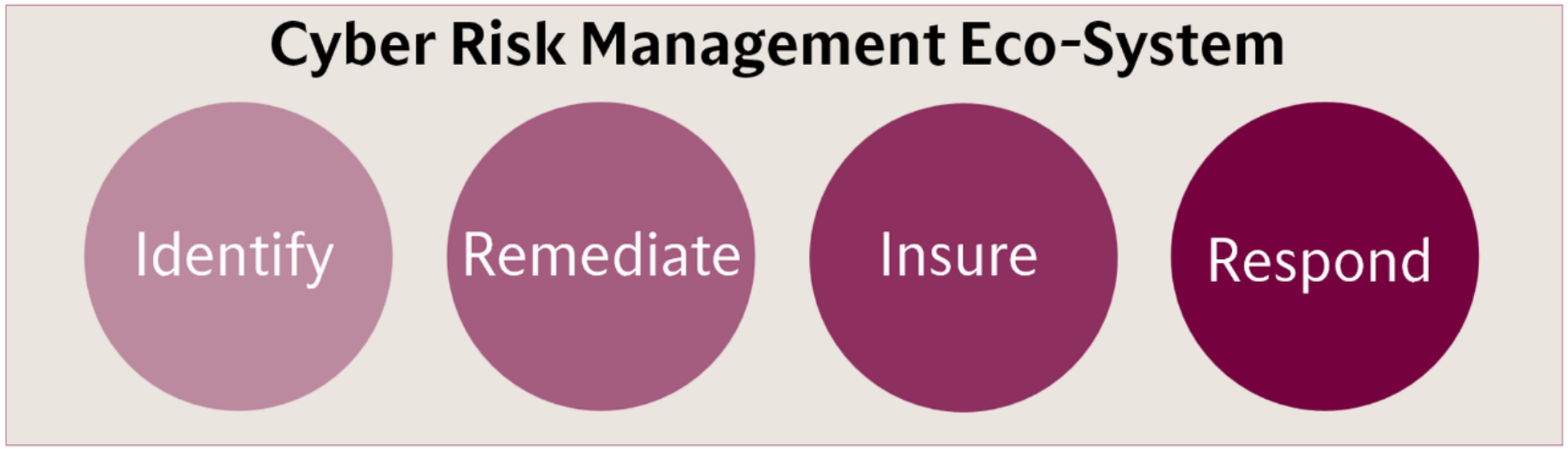
Karla Reffold, Chief Insights Officer

November 12, 2026

© 2025 The Hartford

Our Role – CyberTech Insurance & Risk Management

The Hartford



SUREFIRE CYBER

AI 101

Understanding the Fundamentals of Artificial Intelligence and Emerging Cyber Risks

November 2026

About Surefire Cyber

Surefire Cyber delivers expert-led, AI-powered incident response with quality, consistency, and speed.

We use experience and automation to help our clients manage cyber incidents decisively, reducing impact so an incident is a manageable event, not a defining moment.

Built to support insurance-driven response at scale, our end-to-end incident response platform enables faster outcomes, consistent quality, transparent pricing, claims-ready reporting, and data-enriched insights.

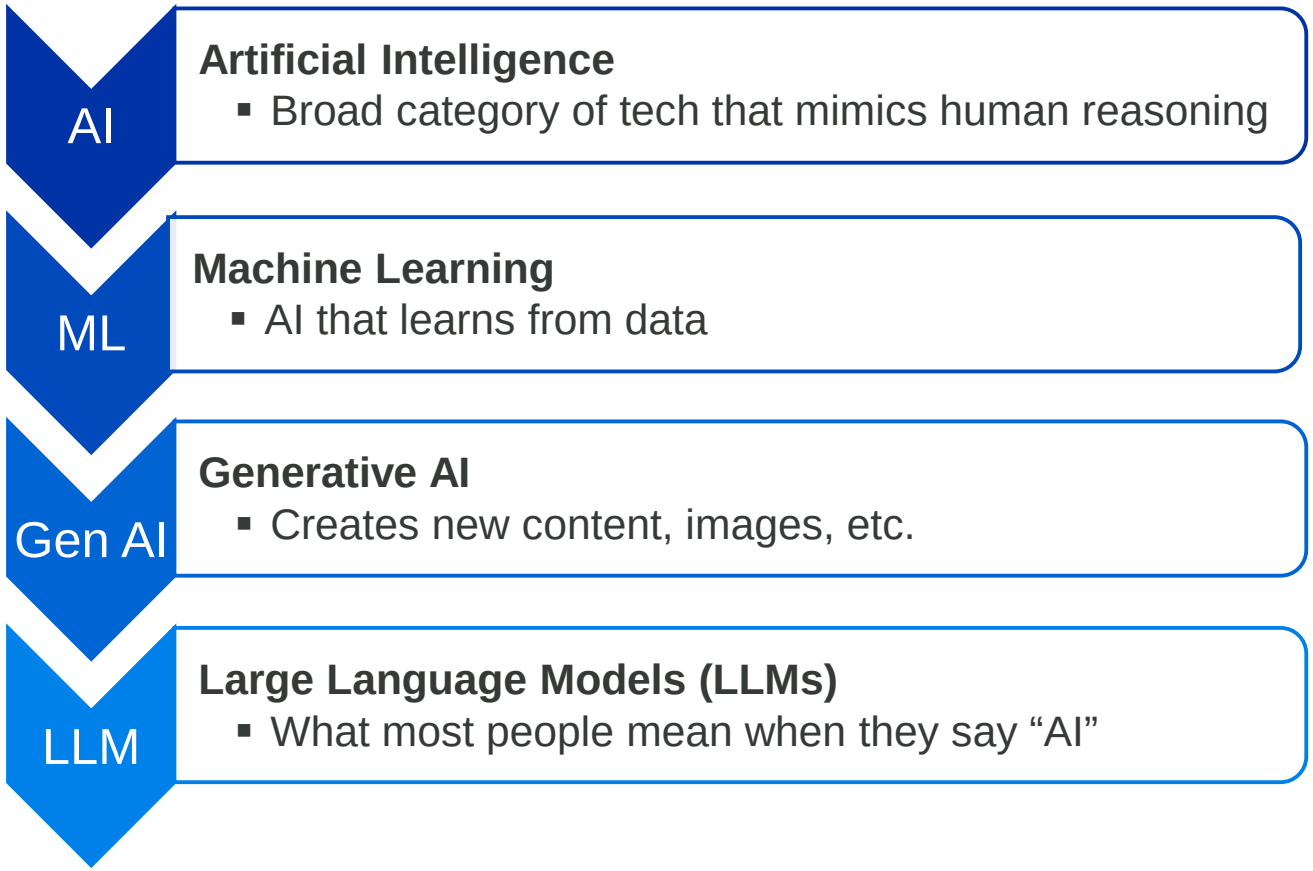
We are creating the future of incident response and changing how our clients experience cyber events, helping them plan, respond, and recover with greater confidence.

SUREFIRE CYBER

01 Demystifying AI

Terminology

SUREFIRE CYBER



Key Insight

AI is a Pattern-Matching Engine

It’s a sophisticated autocomplete that predicts what comes next based on patterns in massive datasets.

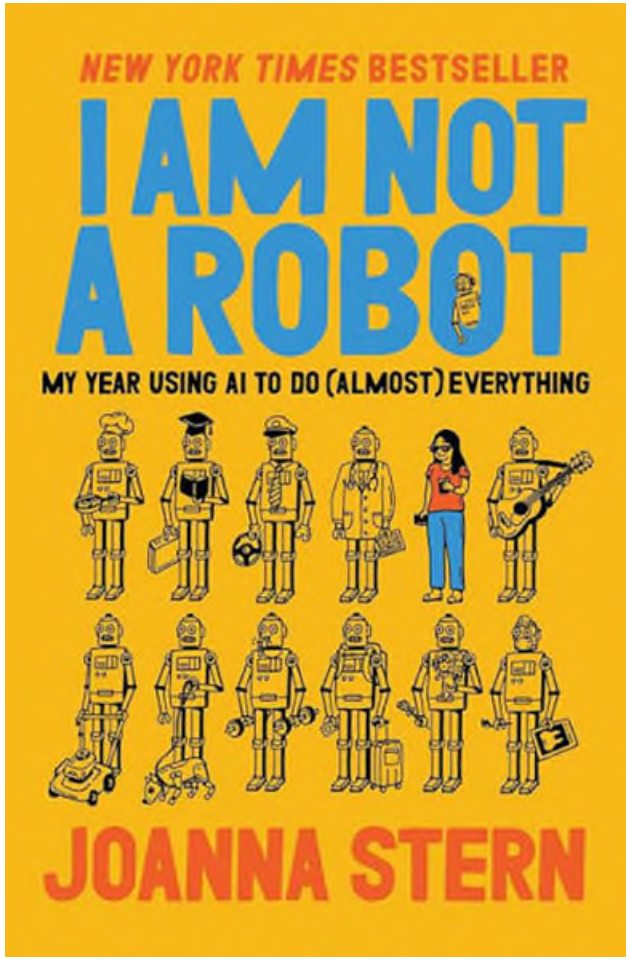
It doesn’t “think” or “understand” like humans.

When you lower the **temperature** (creativity) of outputs you can possibly get the same output of words from the same dataset.

AI Isn’t Just ChatGPT

It’s showing up in the world in many ways

- Generative AI
- Diagnostic AI
- Autonomous driving
- Robotics
- Recommendations and personalization
- Wearable and recordable AI
- Agents



I’m Not Using AI….

Businesses may be exposed in additional areas

- Calendar features that suggests times to meet
- Zoom, Teams, Otter
- Customer service chatbots
- Accounting software that match transactions
- Inventory forecasting or logistic software
- Ad platforms like Google and Meta
- Email filtering software
- Fraud detection
- Security monitoring

How Do Machines Learn?

Much like we do

- How do you teach a child to recognize a dog?
- You don’t list the rules
- You show them dogs until they learn



Training LLMs

SUREFIRE CYBER





Focusing on LLMs

SUREFIRE CYBER

✓ AI CAN

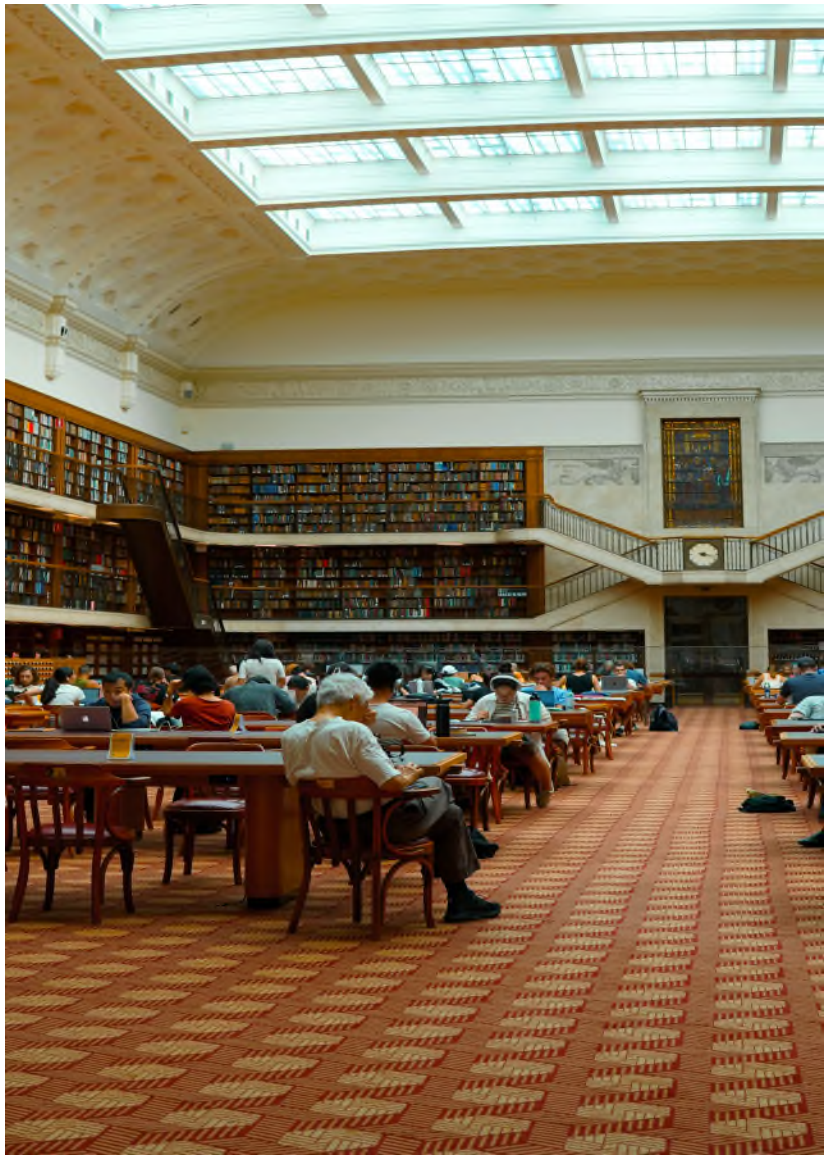
- Generate human-like text, code, images
- Summarize and analyze documents at scale
- Answer questions from trained knowledge
- Automate repetitive cognitive tasks
- Identify patterns in large datasets
- Translate between languages

✗ AI CANNOT

- Truly understand context or meaning
- Guarantee factual accuracy (hallucinations)
- Apply real-world judgment or ethics
- Access real-time information (without tools)
- Know what it doesn't know
- Take accountability for its outputs

RAG Architecture

SUREFIRE CYBER





SUREFIRE CYBER

02

AI Risk – External Actors

AI for Ransomware and Espionage

SUREFIRE CYBER

- Automated reconnaissance, faster vulnerability discovery, malware assistance, password-guessing improvements
- Increasing speed and volume
- Making it easier to become a threat actor
- Disinformation campaigns

JADEPUFFER

SUREFIRE CYBER

How the first AI-driven ransomware played out (first reported July 2026)

- Initial access via a known, year-old common vulnerability and exposure (CVE) in an internet-exposed server (Langflow, a tool for building AI apps)
- AI hunts for credentials (reconnaissance) and moves laterally
- Investigated and adapted its approach at extreme speeds
- Fixed one error in 31 seconds when encountering an issue
- Targets a production database: makes an encrypted copy, drops the originals, then drops whole databases outright
- No real path to decryption: the key was random, printed once and never saved or sent

Source: Sysdig

Why Was This Different?

SUREFIRE CYBER

Development of ransomware

Generation	Human does	Machine does
Ransomware 1.0	Everything	Executes malware
Scripted ransomware	Makes decisions	Executes predefined playbook
AI-assisted ransomware	Makes decisions	Helps write/research/execute
Agentic ransomware	Sets objective	Makes decisions + executes + adapts

Vulnerability Exploitation

Is it really getting worse?

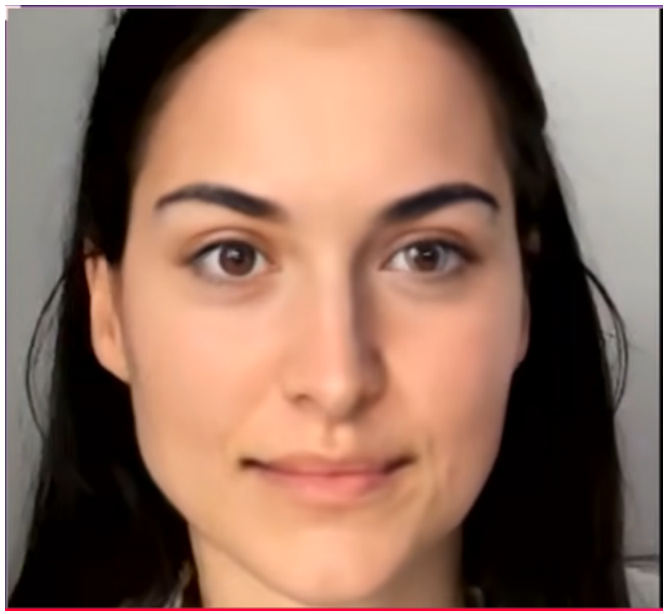
Case Year	Mean days	Median days	% exploited ≤30 days	% same-day/zero-day
2024	236.8	79	30.0%	16.7%
2025	162.8	93.5	23.1%	17.0%
2026	272.6	157	21.7%	10.9%

How to read it: days from CVE publication to exploitation in Surefire Cyber cases. 2026 is not faster, it is broader. The median has stretched, while same-day exploitation has not risen.

Source: Surefire Cyber

Deepfakes Explained

- AI-generated or altered audio, video, and images that convincingly imitate real people. Made with very little audio or video content.
- Voice is almost impossible to distinguish
- Video is catching up
- High profile fraud cases but the hidden risk is disinformation within organizations





Source: GetReal Labs

Business Email Compromise (BEC) and AI

VOLUME IS UP

- LLMs eliminate the typos and awkward grammar that used to be red flags, translate messages flawlessly, and personalize emails at scale using public information from LinkedIn and social media.

FINANCIAL LOSS

- Around 30% of Surefire Cyber BEC clients experience financial loss
- Percentage is unchanged but volume is up
- Median loss is $200,000

SOPHISTICATION

- Not yet increasing
- The business of cyber

Source: Surefire Cyber

How Organizations Can Protect Themselves

SUREFIRE CYBER

MITIGATIONS

Strong security basics and policies

Patch regularly

Phishing resistant multi-factor authentication (MFA)

Backups that are tested

Managed detection and response (MDR)

Tabletops / incident response planning

QUESTIONS TO ASK

What would happen if a suspicious voice, video, payment request, or AI-generated message is received?

How often do you patch internet facing devices?

What type of MFA is in place and where?

How often do you test your backups?

SUREFIRE CYBER

03

AI Risks – From The Inside

Internal AI Risks

Risk	Example	Mitigation
Data exposure	Shadow AI or public tools	Approved AI tool policy with data loss prevention (DLP) controls and training
Supply chain risk	AI tools access business data, insecure integrations and vendor dependency	AI vendor due diligence, monitor integrations and continuity plans
AI manipulation	Prompt injection	External content is untrusted and restrict permissions
Inaccurate outputs	Hallucinations and fabricated outputs	Verification required and define where AI needs intervention

Internal AI Risks

SUREFIRE CYBER

Risk	Example	Mitigation
Insecure technology	AI generated code containing vulnerabilities	Secure development practices and code review
Bias and accountability	Bias and overreliance on AI	Establish accountability processes and test models for bias
Operational controls	Agents deleting data or changing systems	Restrict access and permissions and implement human review
Copyright infringement	AI generated images may infringe copyright	Accountability and oversight based on business risk

SUREFIRE CYBER

04

Defending With AI

AI-Enabled Defense

SUREFIRE CYBER

FIND

- Continuously scans source code, dependencies and web apps for vulnerabilities
- Ranks what is actually exploitable, not just what exists
- Prioritizes vulnerabilities in a world where thousands are discovered

FIX

- Generates a suggested patch and opens it as a request for a developer to review
- Newer tools validate the fix in an isolated environment before proposing it
- Fixes code before deployment

DETECT EXPLOITATION

- Watches deployed code and infrastructure at runtime for exploitation attempts
- Flags behavior that is unexpected for review
- Creating timelines or initial findings

Why it matters: Discovery, speed, and prioritization helps businesses respond to the increased threat. We can be one step ahead of adversaries.

AI Threat Detection: The Bleeding Edge

SUREFIRE CYBER

WHAT IS EMERGING

- AI agents that triage and investigate alerts across endpoint detection and response (EDR), identity, cloud and email
- Correlates events that used to take a large security operations center (SOC) team, deep logging and manual stitching
- Writes up each investigation in plain English with its reasoning shown

WHY AI FITS HERE

- A hallucinated alert costs an analyst minutes, a missed one costs an incident
- The volume and spread of events across platforms is beyond what people can review
- With proper guardrails (human approval for containment, an audit trail for every verdict, ongoing tuning) false positives can be kept to a minimum

SUREFIRE CYBER

05 Takeaways

Questions To Ask

- How are you defining AI?
- What AI tools are being used?
- Do you have an AI policy and how is it monitored?
- What approval stages are in place where AI is used?

Takeaway Checklist

- Know what data you have and what AI you use
- Establish an AI policy
- Don’t enter confidential information into unapproved tools
- Maintain fundamental cyber controls
- Maintain processes for funds transfers
- Keep a human in the loop

Thank You

For immediate cyber intelligence support:
cyberintel@surefirecyber.com

Follow us:
LinkedIn | X

For immediate support or to report an incident:
response@surefirecyber.com
1-800-270-9034

This document (including any attachments) may contain privileged, confidential, or protected information intended only for the intended recipient. If you are not the intended recipient, any unauthorized review, use, disclosure, or distribution is prohibited. If you have received this document in error, you are required to notify the sender, then delete this document and any attachment from your computer and any of your electronic devices where the document is stored.

Questions?

© 2025 The Hartford

Disclaimer

The information provided in these materials is intended to be general and advisory in nature. It shall not be considered legal advice. The Hartford does not warrant that the implementation of any view or recommendation contained herein will: (i) result in the elimination of any unsafe conditions at your business locations or with respect to your business operations; or (ii) be an appropriate legal or business practice. The Hartford assumes no responsibility for the control or correction of hazards or legal compliance with respect to your business practices, and the views and recommendations contained herein shall not constitute our undertaking, on your behalf or for the benefit of others, to determine or warrant that your business premises, locations or operations are safe or healthful, or are in compliance with any law, rule or regulation. Readers seeking to resolve specific safety, legal or business issues or concerns related to the information provided in these materials should consult their safety consultant, attorney or business advisors. All information and representations contained herein are as of November 2026.

The Hartford Insurance Group, Inc., (NYSE: HIG) operates through its subsidiaries, including the underwriting company Hartford Fire insurance Company, under the brand name, The Hartford®, and is headquartered in Hartford, CT. For additional details, please read The Hartford's legal notice at www.thehartford.com.