Cybersecurity: Risk, Defense & Recovery

Figures current as of October 2026

How attacks on computers and networks happen, what they cost, and how households and organizations prevent them, find them, and recover from them. The course covers how to think about cyber risk; who attacks and how a break-in unfolds; passwords, multifactor authentication, and passkeys; software flaws and the updates that fix them; scams and fraud; ransomware; the risks that arrive through suppliers, software, and the cloud; attacks on water, power, and hospitals; what states do in cyberspace; artificial intelligence (AI) on both sides; detection and response; backups and recovery; insurance; the law; and privacy. Two worked examples run through it, a household whose accounts are taken over and a manufacturer recovering from ransomware. A dated watch list closes the course.

MODULE 1 — Risk, Not Fear

Goal: Learn what cybersecurity protects, what attacks cost, and how to reason about risk. [Govern]

Confidentiality, integrity, and availability; risk as likelihood times impact; threats, vulnerabilities, and assets; what attacks cost Americans and companies in 2025-26, and why the counts disagree; the defenders' side, from spending to staffing; and the NIST Cybersecurity Framework's six functions, the map this course follows.

MODULE 2 — Adversaries and the Anatomy of an Intrusion

Goal: Know who attacks, why, and the steps nearly every break-in follows. [Identify]

Criminals, states, hacktivists, and insiders; the criminal economy of ransomware brands, affiliates, access brokers, and password-stealing malware; how intrusions began in the latest data, with exploited flaws ahead of stolen passwords; the chain from first access to damage, and the ATT&CK catalog of techniques; and how attackers are named and identified.

MODULE 3 — Identity: Passwords, MFA, and Passkeys

Goal: Protect accounts, and see why signing in became the main battleground. [Protect]

Why passwords fail; the forms of multifactor authentication and the phishing kits that beat them; passkeys and how widely they are used; account recovery and help desks as the weak point; stolen sessions; the identities of machines and AI agents; a household's account takeover, the first worked example; and whether passkeys are ready to replace passwords.

MODULE 4 — Vulnerabilities and Updates

Goal: Understand how software flaws are found, scored, exploited, and fixed. [Identify, Protect]

What a vulnerability is; the CVE system, severity scores, and the government's list of flaws known to be exploited; the record count of 2025 and the faster pace of 2026; why attackers favor the devices at a network's edge; patching by priority; software past its end of life; the CVE program's funding scare; and whether software makers should be liable for insecure products.

MODULE 5 — Social Engineering, Scams, and Fraud

Goal: Recognize attacks on people, and the habits and rules that blunt them. [Protect]

Phishing by email, text, voice, and QR code; business email compromise and cloned voices; investment scams and the compounds that run them, with the Prince Group case; fraud against older people; the habits that work; the household example's second chapter; and who should pay when a scam succeeds.

MODULE 6 — Ransomware and Extortion

Goal: Learn how the ransomware business works, how it is changing, and what payment buys. [Respond, Recover]

Encryption, data theft, and double extortion; the groups active in 2026; payments and payment rates; Change Healthcare as the case of scale; the first ransomware attack documented as run end to end by an AI agent; governments' responses; the start of the second worked example, a manufacturer hit by ransomware; and whether paying should be banned.

MODULE 7 — Suppliers, Software, and the Cloud

Goal: See how other companies' code and services carry risk into every organization. [Govern, Identify]

Vendors and service providers as a way in; software supply chains, from SolarWinds and XZ Utils to the Shai-Hulud worm; stolen tokens for cloud software; concentration and outages, from CrowdStrike's faulty update to a Japanese cloud's ransomware outage of October 2026; one carmaker's shutdown and its suppliers; and the tools of vendor risk, from software bills of materials to contracts.

MODULE 8 — Critical Infrastructure and Operational Technology

Goal: Understand why attacks on plants, utilities, and hospitals differ, and how they are defended. [Protect, Recover]

Operational technology and why it differs from office computing; water utilities and the controller attacks of 2023-26; power grids, from Ukraine to Volt Typhoon's footholds; hospitals; and the defenses: separate networks, nothing exposed to the internet, no default passwords, and the ability to run by hand.

MODULE 9 — States in Cyberspace

Goal: Learn what governments do with cyber operations, and the limits of law and deterrence. [Identify]

Espionage, prepositioning, sabotage, influence, and theft; the cyber front of the 2026 Iran war; China's Typhoon campaigns; Russia from NotPetya to Ukraine; North Korea's theft for the state; norms, attribution, and indictments; the United States' turn to offense; and whether private companies should be allowed to hack back.

MODULE 10 — AI on Both Sides

Goal: Weigh how AI changes attack and defense, and how AI systems are themselves attacked. [Protect, Detect]

How attackers use AI, from phishing to intrusions run by agents; how defenders use it, from finding flaws to sorting alerts; securing AI systems against prompt injection and overly broad permissions; who gets the most capable models; and whether AI favors attackers or defenders.

MODULE 11 — Detection and Response

Goal: Learn how attacks are spotted and how the first days of an incident are run. [Detect, Respond]

Logs, endpoint detection, and the security operations center; how long attackers go unnoticed; the incident-response cycle; the reporting deadlines in America, Europe, and Britain; working with the police; and the manufacturer's first three days.

MODULE 12 — Recovery: Backups and Rebuilding

Goal: Plan for getting back to work after an attack. [Recover]

Backups that survive an attack; recovery objectives; rebuilding identity systems first; restoring or rebuilding; running on paper; how Marks & Spencer and the Co-op came back from the attacks of 2025; and the manufacturer's days 4 to 60.

MODULE 13 — Cyber Insurance and the War Exclusion

Goal: Understand what cyber insurance covers, why it is hard to price, and how war is excluded. [Govern]

The market and what policies pay for; the controls insurers demand; losses that arrive together; NotPetya and Merck's claim; the war exclusions of 2023 and their first test in the Iran war; the manufacturer's claim; and who should carry catastrophic cyber risk.

MODULE 14 — Governance, Disclosure, and Regulation

Goal: Learn who answers for cyber risk and what the law requires in America, Europe, and Britain. [Govern]

The board's job; America's patchwork, from the SEC's disclosure rule to the incident-reporting law and the information-sharing law; the national cyber agency's capacity; the strategy of March 2026; Europe's NIS2, Cyber Resilience Act, and DORA; Britain's bill and its ransomware plan; the manufacturer's disclosures; and whether mandates or markets make organizations safer.

MODULE 15 — Privacy and Data Brokers

Goal: See what personal data is already exposed, and what people and laws can do about it. [Protect]

Breaches as the source of exposed data; the data-broker business; state privacy laws and California's one-request deletion platform; credit freezes, opt-outs, and other steps; the household example's last chapter; and whether governments should be able to require access to encrypted data.

MODULE 16 — Where Things Stand, and What to Watch

Goal: Close with the argument, the contradictions, and a dated watch list.

The course's argument on a page; the contradictions no single module states; what to watch from November 2026 on; and a last word.

A note on the sources: counts of cybercrime come from what victims report. The Federal Bureau of Investigation's (FBI) Internet Crime Complaint Center (IC3) and the Federal Trade Commission (FTC) tally complaints, so their totals are floors rather than measures of all crime. Statistics on breaches come from firms that investigate incidents or insure against them, and each sees its own clients: Verizon's Data Breach Investigations Report, IBM's Cost of a Data Breach study, Mandiant's M-Trends, and Coveware's reports on ransomware. Figures for payments in cryptocurrency come from Chainalysis, which traces them on public ledgers.

Each figure is dated to its report and named with its source, because reports measure different things: one counts every incident, another only confirmed losses of data, a third only its own customers. Figures a company gives about its own products, such as how many flaws a tool found, are the company's claims. The bracketed words in the outline name the functions of the framework of the National Institute of Standards and Technology (NIST) that each module serves; Module 1 explains the framework.

Two worked examples run through the course: the Harper household, whose accounts are taken over (Modules 3, 5, and 15), and Harrow Precision, a manufacturer hit by ransomware (Modules 6 and 11 to 14). Both are composites, built from real attack methods and real figures; no actual family or company is described.

The field argues over nine questions: whether passkeys are ready to replace passwords, whether software makers should be liable for insecure products, who should pay when a scam succeeds, and whether paying ransoms should be banned. It also argues over whether private companies should be allowed to hack back, whether AI favors attackers or defenders, who should carry catastrophic cyber risk, whether mandates or markets make organizations safer, and whether governments should be able to require access to encrypted data. On each, the course makes each case as its advocates make it, gives the evidence, and then states its own reading once.

Where Things Stand, as of October 31, 2026

The Measures This Course Relies On

The figures the course relies on, each with its source and date and the module that explains it. Counts of reported crime are floors, since most victims never report, and figures from companies describe their own samples.

Measure	Figure	Source and date	Module
Losses reported to the FBI	$20.9 billion in 2025, up 26%, from 1,008,597 complaints	FBI Internet Crime Complaint Center (IC3), report for 2025, April 2026	1
Fraud losses reported to the FTC	$15.9 billion in 2025, up from $12.5 billion	FTC, Consumer Sentinel data for 2025	1
Costliest kind of fraud	Investment fraud: $8.65 billion in 2025, $7.23 billion of it in cryptocurrency schemes	FBI IC3, report for 2025	5
Average cost of a breach	$4.99 million worldwide and $11.5 million in America, in a survey of 602 breached organizations	IBM and the Ponemon Institute, July 29, 2026	1
How breaches began	Exploited flaws 31%, ahead of stolen credentials at 13%; third parties involved in 48%	Verizon, Data Breach Investigations Report, May 2026	2
New flaws published	48,185 records in 2025, up 20.6%; 35,364 in the first half of 2026	Jerry Gamblin's count of the Common Vulnerabilities and Exposures (CVE) program	4
Flaws known to be exploited	245 added in 2025, for 1,484 in all	The Known Exploited Vulnerabilities catalog of the Cybersecurity and Infrastructure Security Agency (CISA)	4
Passkey sign-ins	36% of sign-ins at 15 large services, up from 26% a year earlier	FIDO (Fast Identity Online) Alliance, Passkey Index, October 15, 2026	3
Ransom payments	About $820 million in 2025; 28% of victims paid, a record low	Chainalysis, February 26, 2026	6
Victims named on leak sites	2,122 in the first quarter of 2026; Qilin led with 338	Check Point	6
Median dwell time	14 days in 2025, up from 11	Mandiant, M-Trends 2026, March 23, 2026	11
Cryptocurrency stolen by North Korea	At least $2.02 billion in 2025, a record	Chainalysis, December 18, 2025	9
Spending on security	$244 billion expected in 2026, up 11.6%	Gartner, February 5, 2026	1
Cyber insurance premiums	$15.6 billion in 2025; $16.4 billion expected in 2026	Swiss Re	13
Skills gaps	95% of security professionals report at least one in their teams	ISC2, an association of security professionals: workforce study, December 4, 2025	1
Staff at America's civilian cyber agency	About 2,200 in June 2026, from about 3,400 before 2025	Markwayne Mullin, Secretary of Homeland Security, June 2026	14
Data brokers registered in California	More than 600	California's data-broker registry, 2026	15

The State of the Art, as of October 2026

The Leading Defenses, the Newest Attacks, and Who Makes Them

This table is the course's reference point: the most advanced defense in each area as of October 31, 2026, who makes or sets it, when it arrived, and the module that explains it. A second table lists the newest methods on the attackers' side. Each module gives the figures and sources.

Area	State of the art	Who	Since	Module
Framework	The NIST Cybersecurity Framework 2.0, whose functions run from Govern to Recover	NIST	February 26, 2024	1
Sign-in	Passkeys: a key pair kept on the device, or synced across a person's devices, that a look-alike site cannot use	Standards of the FIDO (Fast Identity Online) Alliance, built into Apple's, Google's, and Microsoft's systems	Announced by the three companies in May 2022; new Microsoft accounts passwordless by default from May 1, 2025	3
Federal sign-in rule	Phishing-resistant multifactor authentication (MFA) required for federal staff	White House memorandum M-22-09	January 2022	3
Session protection	Device Bound Session Credentials, which tie a sign-in cookie to the computer's security chip	Google, in Chrome	2026; websites must add support	3
Prioritizing flaws	The Known Exploited Vulnerabilities catalog; version 4 of the Exploit Prediction Scoring System (EPSS), which predicts exploitation	The Cybersecurity and Infrastructure Security Agency (CISA); FIRST, an association of incident-response teams	Catalog from November 3, 2021; EPSS 4 from March 17, 2025	4
Detection and response	Endpoint and extended detection and response, with managed services watching around the clock	CrowdStrike, Microsoft, Palo Alto Networks, and others	Standard in large organizations; spreading to mid-size firms through insurers' demands	11, 13
Incident response	NIST Special Publication 800-61 Revision 3, tied to the framework's functions	NIST	April 2025	11
Backups	Immutable copies that cannot be changed or deleted for a set period, kept in an isolated recovery environment	Backup makers and cloud providers	Standard advice since the ransomware wave of 2020 and 2021	12
Software supply chain	Short-lived publishing credentials on npm; software bills of materials	GitHub; CISA and the EU's Cyber Resilience Act	npm's token limits from November 5, 2025; bills of materials required in the EU from December 11, 2027	7, 14
AI that finds and fixes flaws	Claude Mythos Preview in Project Glasswing; Google's Big Sleep and CodeMender; OpenAI's Codex Security	Anthropic; Google; OpenAI, whose figures are their own	Glasswing from April 7, 2026; Codex Security from March 6, 2026	10
AI in security operations	Agents that triage alerts and phishing reports and draft investigations	Microsoft's Security Copilot; similar agents from Google and CrowdStrike	Security Copilot included for Microsoft's largest business customers from November 2025	10

Area	State of the art	Who	Since	Module
Post-quantum encryption	NIST's first three standards designed to resist quantum computers, one for setting up keys and two for signatures, used alongside today's methods in browsers	NIST; browser makers; Cloudflare	Standards August 13, 2024; on by default in Chrome and Edge from version 131 and in current versions of Firefox and Apple's systems; about 70% of browser connections to Cloudflare by September 29, 2026, by its count	16

On the Attackers' Side

Method	Example	Module
Infostealers sold as a service	Lumma, the most widely used, until its disruption in May 2025	2
Phishing that steals sessions	Kits that relay a real sign-in and keep the session cookie, defeating codes and push prompts	3
Talking past help desks	Scattered Spider and ShinyHunters, whose members include teenagers	2, 3, 12
Exploiting edge devices	Flaws in virtual private networks (VPNs), firewalls, and gateways, turned into attacks soon after disclosure	4
Ransomware with data theft	Qilin, Akira, The Gentlemen, and a revived LockBit, the leaders on leak sites in early 2026	6
Poisoned packages and stolen tokens	The Shai-Hulud worm on npm from September 2025; the Salesloft Drift token thefts of August 2025	7
Hiding inside infrastructure	Volt Typhoon in American utilities; Salt Typhoon in telecommunications carriers	8, 9
Wiping through management tools	Handala's wipe of Stryker's devices through Intune, March 11, 2026	9
Deepfakes	A deepfake video call that led an Arup employee in Hong Kong to send millions in early 2024	5
AI agents running intrusions	GTG-1002's espionage, reported November 13, 2025; JadePuffer's extortion, July 2026; the mass exploitation of PaperCut servers, August and September 2026	6, 10
Industrial-scale scams	Compounds in Southeast Asia staffed by trafficked workers, such as those the Prince Group is accused of running	5

Who's Who

The Agencies, Companies, Groups, and People This Course Names, as of October 31, 2026

Governments and Agencies

Name	What it does in cybersecurity	Modules
Cybersecurity and Infrastructure Security Agency (CISA)	America's civilian cyber-defense agency, part of the Department of Homeland Security (DHS); runs the catalog of exploited flaws and will enforce the Cyber Incident Reporting for Critical Infrastructure Act (CIRCIA)	4, 6, 8, 14
Office of the National Cyber Director	White House office that coordinates cyber policy and wrote the strategy of March 2026	9
National Security Agency (NSA) and US Cyber Command	Signals intelligence and military cyber operations, under one commander	8, 9
Federal Bureau of Investigation (FBI)	Investigates cybercrime, takes down criminal servers, and runs the Internet Crime Complaint Center (IC3)	1, 5, 6, 9, 11
Justice Department	Prosecutes attackers, seizes their domains and coins, and enforces the bulk data rule	5, 9, 15
Securities and Exchange Commission (SEC)	Requires listed companies to disclose material cyber incidents and how they manage cyber risk	11, 14
National Institute of Standards and Technology (NIST)	Publishes the Cybersecurity Framework, the National Vulnerability Database, and encryption standards	1, 4, 11, 16
Federal Trade Commission (FTC)	Consumer-protection agency; tallies fraud reports and enforces the ban on selling Americans' data to foreign adversaries	1, 5, 15
Federal Communications Commission (FCC)	Runs the Cyber Trust Mark, a label for secure connected devices	14
California Privacy Protection Agency (CalPrivacy)	Runs California's data-broker registry and its one-request deletion platform	15
European Commission and the European Union Agency for Cybersecurity (ENISA)	The EU's executive, which enforces the NIS2 (Network and Information Security 2) directive and proposes cyber laws, and the EU's cybersecurity agency	4, 14
Lloyd's of London	Insurance market whose rules set the war exclusions of 2023	13

Companies and Organizations

Name	What it does in cybersecurity	Modules
Microsoft	Windows, Entra ID, Intune, and Defender; publishes a yearly digital defense report and names threat groups after weather	2, 3, 4, 7, 9, 10
Google, with Mandiant and Wiz	Chrome, Android, and Google Cloud; its Mandiant unit investigates breaches and publishes M-Trends; Wiz secures clouds	1, 2, 3, 10, 11
CrowdStrike	Endpoint detection and threat intelligence; its faulty update of July 2024 crashed millions of Windows computers	2, 7, 13
Palo Alto Networks	Firewalls and security platforms; bought CyberArk, an identity-security company	1
Anthropic and OpenAI	AI labs whose models find flaws for defenders, and whose reports document misuse by attackers	10
Verizon and IBM	Publishers of the Data Breach Investigations Report and the Cost of a Data Breach study	1, 2, 6, 11

Name	What it does in cybersecurity	Modules
Chainalysis and Coveware	Trace ransom payments on public ledgers; negotiate for victims and report on payments	6, 9
Check Point, Sysdig, GreyNoise, and Dragos	Security companies that count leak-site victims, secure clouds, watch mass exploitation, and protect industrial systems	6, 8, 10
FIDO Alliance	Industry group that writes the standards behind passkeys	3
MITRE	Nonprofit that runs the Common Vulnerabilities and Exposures (CVE) program and the ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) catalog of attack techniques	2, 4
Munich Re, Swiss Re, CyberCube, and Marsh	Reinsurers, a risk modeler, and a broker that measure and price cyber insurance	13
Cyber Monitoring Centre	British body that rates the economic damage of large cyber events	7

Threat Groups

Name	Who they are	Modules
Qilin, Akira, The Gentlemen, and LockBit	Ransomware groups that led leak-site counts in 2025 and 2026	2, 6
ALPHV (BlackCat) and RansomHub	Ransomware groups behind the attack on Change Healthcare	6
Clop	Extortion group behind the thefts through MOVEit file-transfer software in 2023	2, 6
Scattered Spider and ShinyHunters	English-speaking criminals who talk their way past help desks and steal tokens	2, 7
TeamPCP	Group behind the Shai-Hulud worm on npm; alleged members were charged in Australia in August 2026	7
UNC6395	Group behind the theft of data through Salesloft Drift tokens in August 2025	7
Volt Typhoon and Salt Typhoon	Chinese state groups, one hiding in American infrastructure, the other in telecommunications networks	8, 9, 15
i-Soon	Chinese contractor that took hacking jobs from state agencies	9
Sandworm and APT28	Units of the GRU, the military intelligence service of Russia	7, 8, 9, 10
The SVR, the foreign intelligence service of Russia	Behind the SolarWinds intrusion and the reading of Microsoft leaders' email	7, 9
Handala and CyberAv3ngers	Personas of Iran's intelligence ministry and of its Revolutionary Guards, by American findings	8, 9, 13
North Korea's state hackers	Steal cryptocurrency and place remote technology workers in foreign companies under false identities	9
NoName057(16)	Pro-Russian group that floods websites with traffic, disrupted in July 2025	2
GTG-1002 and JadePuffer	A Chinese state group and a criminal crew whose intrusions were run largely by AI agents	6, 10
Prince Group	Cambodian conglomerate accused by American prosecutors of running scam compounds	5

People

Name	Who they are	Modules
Nick Andersen	Acting director of CISA since February 26, 2026	14
Markwayne Mullin	Secretary of Homeland Security since March 24, 2026	14
Sean Cairncross	National Cyber Director, confirmed August 2, 2025	9
Gen. Joshua Rudd	Commander of US Cyber Command and director of the NSA since March 20, 2026	9
Paul Atkins	Chairman of the SEC since April 2025	14
Tulsi Gabbard	Director of national intelligence from February 2025 to June 2026	15
Sean Plankey	President Trump's nominee to lead CISA, who withdrew in April 2026	14
Rand Paul	Senator from Kentucky who has resisted a long extension of the 2015 information-sharing law	14
Andres Freund	Microsoft engineer who found the backdoor in XZ Utils in March 2024	7
Chen Zhi	Chairman of the Prince Group, sent to China in January 2026	5

Terms

The Vocabulary of Cybersecurity

Term	Meaning
Confidentiality, integrity, and availability	The three properties security protects: data seen only by those allowed, left unaltered, and there when needed
Threat, vulnerability, and risk	Who might attack; the weakness they could use; and the likelihood and cost of the harm taken together
Phishing	A message that tricks someone into giving up a password, approving a sign-in, or running malware; by phone it is voice phishing
Multifactor authentication (MFA)	A second proof of identity besides a password, such as a code, an app prompt, or a security key
Passkey	A sign-in that replaces the password with a pair of keys, the private one kept on the user's device, which a fake site cannot use
Session cookie or token	What a site issues after sign-in so that the user need not sign in again; stolen, it opens the account without password or MFA
Adversary in the middle (AiTM)	A phishing kit that relays a real sign-in and keeps the session cookie
Infostealer	Malware that copies saved passwords, cookies, and wallets and sends them to its operator
Subscriber identity module (SIM) swap	Persuading a carrier to move a victim's phone number to the attacker's SIM card, so the codes go to the attacker
Vulnerability, zero-day, and n-day	A flaw an attacker can use; a zero-day is used before any fix exists, an n-day after a fix is out
Common Vulnerabilities and Exposures (CVE)	The numbering system for publicly known flaws
Common Vulnerability Scoring System (CVSS) and EPSS	Scores for a flaw's severity and for the chance that it will be exploited soon
Known Exploited Vulnerabilities (KEV) catalog	CISA's list of flaws with evidence of use in real attacks
Edge device	Equipment at a network's boundary, such as a VPN gateway or firewall, reachable from the internet
Ransomware	Malware that encrypts a victim's files and demands payment for the key
Double extortion and leak site	Stealing data before encrypting it and threatening to publish it; a leak site is where gangs post victims who refuse to pay
Ransomware as a service (RaaS)	Developers rent ransomware to affiliates, who carry out attacks for a share of the ransom
Initial access broker	A criminal who breaks into networks and sells the access to others
Wiper	Malware that destroys data beyond recovery, sometimes disguised as ransomware
Business email compromise (BEC)	Fraud that uses a hijacked or imitated business email account to redirect payments
Pig butchering	Investment fraud that builds a relationship over weeks before steering the victim into a fake trading platform
Living off the land	Attacking with tools already installed on the victim's systems, which blend in with normal administration
Lateral movement and privilege escalation	Spreading from the first machine to others, and gaining the rights of administrators
Dwell time	How long an attacker stays inside a network before being found

Term	Meaning
Endpoint detection and response (EDR), extended detection and response (XDR), and managed detection and response (MDR)	Software on computers that detects and stops attacks; its extension to email, cloud, and network; and a service that watches it around the clock
Security information and event management (SIEM) and security operations center (SOC)	A system that gathers logs and alerts from across an organization, and the team that watches them
Out of band	Communication through channels the attacker cannot read, such as personal phones or a separate service
Immutable backup	A copy that cannot be changed or deleted for a set period, even by an administrator
3-2-1-1-0	Three copies of data, on two kinds of media, one off site, one offline or immutable, and zero errors when restores are tested
Recovery time and recovery point objectives	How fast a system must be back, and how much recent data the organization can afford to lose
Active Directory and Entra ID	Microsoft's systems that control who may sign in to what, on a company's own servers and in the cloud
Operational technology (OT)	Computers that run physical processes, such as pumps, turbines, and assembly lines
Programmable logic controller (PLC) and supervisory control and data acquisition (SCADA)	A PLC runs one machine or process; a SCADA system monitors and commands many from a control room
Supply-chain attack	An attack that reaches its victims through a supplier's software, service, or update
Managed service provider (MSP)	A firm that runs other companies' computers, often with administrator access to all of them
Software bill of materials (SBOM)	A list of the components inside a piece of software
OAuth token	A permission that lets one application act on a user's data in another without a password
Shared responsibility	The rule that a cloud provider secures its own infrastructure and the customer secures what it puts there
Prompt injection	Instructions hidden in content that an AI model reads, which hijack what it does
Attribution	Judging who carried out an attack, with a stated level of confidence
War exclusion	An insurance clause that removes losses from war and from state-backed cyberattacks
Limit and retention	The most a policy will pay, and the share of a loss the insured pays first
Data broker	A company that collects and sells information about people with whom it has no direct relationship
Credit freeze	A block on a credit file that stops lenders from opening new accounts in a person's name
End-to-end encryption	Encryption in which only the sender and the recipient hold the keys
Post-quantum cryptography	Encryption designed to resist attack by future quantum computers

Module 1 — Risk, Not Fear

What Is Being Protected, What Attacks Cost, and How to Think About Them

1a. What Cybersecurity Protects

Cybersecurity is the protection of computers, networks, and the data and services they carry against attack, misuse, and accident. It is usually described by three properties, known together, from their initials, as the CIA triad. Confidentiality means that only the right people can read the data. Integrity means that data and systems are not changed without authority. Availability means that systems work when they are needed.

Each kind of attack strikes at one of the three. A leaked medical record breaks confidentiality. A criminal who changes the bank details on a supplier's invoice breaks integrity, and the payment goes to the wrong account. Ransomware, malicious software that locks a victim's files until a ransom is paid, breaks availability: a hospital whose records are locked cannot see its patients' histories. Modern ransomware attacks two properties at once, since the gangs copy the data before they lock it and threaten to publish it (Module 6).

Two further properties matter more than they used to. Authenticity means knowing who sent a message or made a request, which is what a cloned voice on a telephone call attacks (Module 5). Safety matters wherever computers run physical things: a tampered controller at a water plant can change the dose of a chemical, and one in a car or a medical device can hurt people (Module 8). The field began by protecting computers. It now protects the services that households, companies, and countries depend on.

1b. Risk: Likelihood Times Impact

Security professionals reason about risk, not fear. A threat is anyone or anything that might cause harm: a criminal gang, a foreign intelligence service, a careless employee, a flood. A vulnerability is a weakness a threat could use, such as unpatched software, a reused password, or staff who have never been warned about fake invoices. An asset is what is worth protecting: money, data, the ability to make and ship products. A risk exists where a threat can use a vulnerability to harm an asset.

Risk is measured as likelihood times impact. A 10% chance each year of a $2 million loss is an expected loss of $200,000 a year, a figure a manager can weigh against the cost of controls. The arithmetic is rough, because both the chance and the size of a loss are estimates, and it understates rare events that could ruin a firm. Its value is that it forces a comparison between risks, so that money goes first to the likeliest and most damaging ones rather than to the most frightening.

There are four ways to handle a risk. It can be reduced with controls, such as multifactor authentication (MFA), which asks for a second proof of identity besides a password (Module 3). It can be transferred to someone else through insurance or contracts (Module 13). It can be avoided by not doing the risky thing at all, such as not storing customers' card numbers. Or it can be accepted, knowingly. No organization can remove all risk; the aim is to bring it to a level it has chosen on purpose, its risk appetite.

Fear sells products, but most losses come from a few common methods: stolen or guessed passwords, deceptive emails and calls, and devices left unpatched at the edge of a network (Modules 2 to 5). Basic controls therefore remove most of the risk, and exotic attacks are rare. Defenders also layer their controls, a practice called defense in depth, so that the failure of one does not open the way to disaster.

1c. What Attacks Cost: The Reported Numbers

The largest public count of losses in America comes from the FBI's Internet Crime Complaint Center. Its report for 2025, released in April 2026, counted 1,008,597 complaints and $20.877 billion in reported losses, 26% more than the $16.6 billion of 2024. Fraud committed online made up about 85% of the losses, and the most common complaint was phishing, messages that impersonate a trusted sender to steal passwords or money. Most victims never report, so the totals are floors.

The FTC counts differently, from the reports consumers send to its Consumer Sentinel network, and put Americans' reported losses to fraud at $15.9 billion in 2025, up from $12.5 billion in 2024. The two databases overlap, and neither covers everything: the FTC's includes fraud with no computer involved, and the FBI's includes losses by businesses. The direction is the same in both, and steeply upward.

For organizations, the most quoted figure is IBM's. Its Cost of a Data Breach study, run with the Ponemon Institute and published on July 29, 2026, put the average cost of a breach at $4.99 million worldwide, a record, and $11.5 million in America, the costliest country. The study surveys 602 organizations that suffered breaches, so its averages describe that sample rather than every company. Verizon's Data Breach Investigations Report of May 19, 2026 drew on more than 31,000 incidents, over 22,000 of them confirmed breaches, in 145 countries, from November 2024 to October 2025.

Counts disagree because they measure different things. Verizon distinguishes an incident, any event that compromises security, from a breach, an incident in which data is confirmed to have reached an unauthorized party. Insurers count claims; incident-response firms count the cases their own clients bring them; governments count what victims report. The course names the source beside each figure, so that a number is never read as more than its source can support.

1d. The Defenders' Side: Spending, Companies, and People

Spending on defense keeps rising. Gartner, a research firm, expected the world to spend $244 billion on information security in 2026, 11.6% more than in 2025 at constant exchange rates (February 5, 2026); its later quarterly updates, as reported, put the figure near $248 billion. Software is the largest part, followed by services and network equipment such as firewalls, the devices that filter traffic at a network's edge.

The industry is consolidating into a few large platforms. Google closed its $32 billion purchase of Wiz, a company that secures cloud computing, on March 11, 2026, the largest acquisition in Google's history. Palo Alto Networks, the largest company that sells only security products, completed its purchase of CyberArk, which manages employees' and machines' access to systems, on February 11, 2026. The deal was valued at about $25 billion when it was announced in July 2025; Palo Alto's filings put the purchase price at $21.1 billion on the day it closed.

People are scarcer than products. ISC2, the largest association of security professionals, surveyed a record 16,029 practitioners for its 2025 workforce study (December 4, 2025) and, for the first time, published no estimate of the global shortage of workers. Its respondents stressed skills over headcount: 95% reported at least one gap in their team's skills, with AI security (41%) and cloud security (36%) the most needed. Spending growth, however fast, does not show that organizations are safer, since most breaches still begin with basics left undone.

1e. A Map: The NIST Cybersecurity Framework 2.0

The most widely used map of the field comes from NIST, America's agency for measurement standards. Its Cybersecurity Framework first appeared on February 12, 2014, written for operators of critical infrastructure. Version 2.0, published on February 26, 2024, was written for organizations of every size and sector, and it added a sixth function, Govern, to the original five. The framework does not prescribe products. It names outcomes an organization should achieve and leaves the means to it.

Its six functions are the questions every security program must answer: Govern, Identify, Protect, Detect, Respond, and Recover.

The last three functions express a change of mind. Older security aimed to keep attackers out, as if a wall could be made perfect. Most organizations now work on the assumption, called assume breach, that some attacks will succeed, and they plan to notice them quickly and to recover. Prevention still matters most, since the cheapest incident is the one that never happens, but a program that only prevents has no answer when prevention fails. The table shows where each function is taught.

Function	What it covers	Modules
Govern	Strategy, risk appetite, roles, policy, and the oversight of suppliers	1, 7, 13, 14
Identify	What an organization owns, who threatens it, and its weaknesses	2, 4, 7, 9
Protect	Identity and access, training, data security, updates, and resilient design	3, 4, 5, 8, 10, 15
Detect	Finding attacks in progress, from logs and alerts	10, 11
Respond	Containing, investigating, communicating, and reporting an incident	6, 11
Recover	Restoring services and data, and learning the lessons	6, 8, 12

Function names and scope from NIST, *The NIST Cybersecurity Framework (CSF) 2.0*, February 26, 2024. Modules 13 to 15 also cover how risk is paid for and governed by law.

Worked example — Pricing a Risk

A 40-person accounting firm wants to know whether $30,000 a year of new controls is worth it: MFA on every account, endpoint detection software on every computer, and backups kept offline. Its insurer's broker estimates an 8% yearly chance of a serious ransomware attack, and the firm puts the cost of one at $1.2 million: three weeks without its systems in tax season, recovery work, legal advice, and lost clients. The expected loss is 0.08 times $1.2 million, or $96,000 a year. Suppose the controls cut the chance by 70%, to 2.4%, and halve the cost of an attack that still succeeds, to $600,000, because the offline backups shorten the outage. The expected loss falls to 0.024 times $600,000, or $14,400 a year. The controls save about $81,600 a year in expected losses for $30,000, so they pay for themselves more than twice over. The figures are estimates, and a range would be more honest than a point: at half the assumed chance of an attack, the saving is about $40,800, still above the cost.

Recommended Reading / Reference

Book: Ross Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems* (third edition, 2020), the standard account of how security works and fails, by a Cambridge professor who died in 2024

Reference: NIST, *The NIST Cybersecurity Framework (CSF) 2.0* (February 26, 2024), the framework's text, with its functions and categories

Module 1 Key Points

Risk, Not Fear

• Security protects confidentiality, integrity, and availability; ransomware now attacks two at once, copying data and locking it.

• Risk is likelihood times impact; it can be reduced, transferred, avoided, or accepted, but never removed.

• Victims reported $20.9 billion of losses to the FBI for 2025, 26% more than for 2024, and reports are floors.

• Gartner expected $244 billion of security spending in 2026; Google's $32 billion purchase of Wiz closed in March 2026.

• NIST's framework has six functions; Detect, Respond, and Recover assume that prevention sometimes fails.

Module 1 Quiz

Check Your Understanding

1. Which property does ransomware that locks a hospital's records attack most directly?

A) Confidentiality, since outsiders can now read the patients' files
B) Integrity, because the records are permanently rewritten
C) Availability
D) Authenticity of the hospital's messages

2. How does the course define risk?

A) The likelihood that a threat uses a vulnerability, multiplied by the harm it would do to an asset
B) Any flaw in software
C) The number of attacks a year
D) An attacker's level of skill

3. What did ISC2 leave out of its 2025 workforce study for the first time?

A) Its salary survey
B) An estimate of the global workforce gap
C) The count of its respondents, which fell sharply that year
D) Certification data

4. Which function did version 2.0 of NIST's framework add in February 2024?

A) Detect
B) Recover, so that organizations plan to restore their services
C) Identify, covering assets and risks
D) Govern, on strategy and roles

5. A firm faces a 10% chance each year of a $2 million loss. What is its expected annual loss?

A) $200,000
B) $2 million, the full size of the loss
C) $20,000
D) Zero until an attack actually happens

Module 2 — Adversaries and the Anatomy of an Intrusion

Who Attacks, Why, and How a Break-In Unfolds

2a. Four Kinds of Attacker

Criminals carry out most attacks, and they want money. They steal it directly, extort it with ransomware, or sell what they steal: passwords, card numbers, and access to companies' networks. Criminal groups range from lone fraudsters to organizations with payrolls, managers, and help desks for their victims. Most work from countries that will not hand them over, Russia and its neighbors above all, which is why so few are ever tried.

States attack for intelligence, for advantage in a future conflict, and sometimes for money (Module 9). Their operations are better funded and more patient than most criminals', and they often hide among criminal activity. Hacktivists attack for a cause, usually by flooding websites with traffic or defacing them; some are fronts for intelligence services. A coordinated European operation in July 2025, Operation Eastwood, disrupted NoName057(16), a pro-Russian group that had flooded government and company websites across Europe for three years.

Insiders are employees, contractors, or partners who misuse access they were given. Some steal or sabotage on purpose; far more cause harm by mistake, by sending data to the wrong person or leaving a storage bucket open to the internet. A newer kind of insider is the fake employee, such as the remote technology workers North Korea places in foreign companies (Module 9). The four kinds differ more in motive than in method, and the same defenses stop most of them.

2b. The Criminal Economy

Cybercrime is specialized, like any mature industry. Developers write ransomware and rent it to affiliates, the criminals who break into victims and run the extortion, under an arrangement called ransomware as a service; the affiliate typically keeps most of each ransom and the brand takes a cut. Initial access brokers break into networks and sell the entry to others. Other specialists supply servers that ignore complaints, launder cryptocurrency, recruit the money mules who move stolen funds through their own bank accounts, or write phishing kits.

Infostealers are the supply line. This malicious software, often hidden in pirated programs or fake browser updates, copies the passwords, cookies, and cryptocurrency wallets stored on an infected computer and sends them to its operators, who sell them in bundles called logs. A stolen session cookie, the token a website issues after someone signs in, can let a buyer into an account without the password or the second factor (Module 3). In May 2025 Microsoft and the Justice Department disrupted Lumma, the most widely used infostealer, after Microsoft counted more than 394,000 infected Windows computers in two months; about 2,300 of its web domains were seized or blocked.

Groups split, rebrand, and return. In February 2024 an operation led by Britain's National Crime Agency took over the servers of LockBit, then the most prolific ransomware brand, and published its affiliates' details. LockBit opened a new leak site in November 2025 and has posted victims since, though far fewer than before. A younger, English-speaking scene has grown alongside the Russian-speaking gangs, known by

names such as Scattered Spider and ShinyHunters. Its members, some still teenagers, specialize in talking their way past help desks; British police arrested four people in July 2025 over attacks on three British retailers.

2c. How Intrusions Begin

Verizon's report of May 2026 found that exploited vulnerabilities, flaws in software that attackers can use, began 31% of the breaches it studied, up from 20% a year earlier. It was the first time in the report's 19 years that exploitation had led the list. Credential abuse, signing in with stolen or guessed passwords, began 13% of breaches but appeared somewhere in about 39% of attack chains, often after the first step. Third parties, such as suppliers, software makers, or service providers, were involved in 48% of breaches, up from 30% a year earlier (Module 7).

Other datasets rank the routes differently because they see different victims. Microsoft's Digital Defense Report of October 1, 2026 found phishing to be the way in for 23% of the cases its incident responders handled, up from 7% the year before. That sample is weighted toward large companies that use Microsoft's services, and it measures the cases Microsoft was called to, not all attacks. The safe conclusion from all the sources is that three routes dominate: flaws in internet-facing systems, stolen credentials, and deceived people.

Exploitation rose because attackers found flaws in the devices that face the internet, which they can attack at scale (Module 4). The model was Clop, a ransomware group that in 2023 used an unknown flaw in MOVEit Transfer, a file-transfer product, to steal data from more than 2,700 organizations, by the count of Emsisoft, a security firm, without encrypting anything. Finding and weaponizing such flaws is getting faster, partly because AI tools help (Module 10).

2d. The Attack Chain

Most intrusions follow the same chain of steps, whoever runs them. Each step leaves traces and offers defenders a chance to stop it, and an attack stopped at any link fails. The table follows the chain.

Step	What the attacker does	Example	A defense
Reconnaissance	Finds people, systems, and exposed devices	Scans the internet for a remote-access device with a known flaw	Knowing one's own exposure first
Initial access	Gets a foothold	An exploited flaw, a stolen password, a phishing email, a supplier's access	Patching, phishing-resistant sign-in, limits on vendors' access
Execution and persistence	Runs tools and makes sure it can return	Plants a web shell, a small program that takes commands through a website	Endpoint detection; allowing only approved programs
Privilege escalation	Becomes an administrator	Copies passwords from a computer's memory, or tricks a help desk	Separate administrator accounts; strict control of privileged access
Discovery and lateral movement	Maps the network and spreads	Uses the victim's own administration tools	Dividing the network into segments; watching administrators' activity
Collection and exfiltration	Gathers data and sends it out	Compresses files and uploads them to a cloud storage service	Alerts on unusual outbound traffic

Step	What the attacker does	Example	A defense
Impact	Achieves its goal	Encrypts or wipes systems, steals money, or spies quietly	Offline backups and practiced recovery

Steps simplified from the catalog of attackers' tactics and techniques described below.

Defenders describe the chain in a shared vocabulary. MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge), maintained by MITRE, a nonprofit that runs research centers for the federal government, catalogs the tactics, or goals, of each step and hundreds of techniques for achieving them, each with a number. A report that says an attacker used technique T1078, valid accounts, tells every reader the same thing, and a company can check whether its tools would detect each technique on the list.

Skilled attackers increasingly live off the land: they use the tools already built into the victim's systems, such as the administration utilities that come with Windows, rather than bringing malware that security software might flag. Their activity then looks like an administrator's routine work. The most patient state groups operate this way for months or years (Module 8).

2e. Naming and Identifying Attackers

Each security company names attackers in its own scheme, which confuses newcomers. Since April 2023 Microsoft has used weather: Typhoon for groups working for China, Sandstorm for Iran, Blizzard for Russia, Sleet for North Korea, and Tempest for criminals, with a temporary Storm number for groups not yet placed. CrowdStrike uses animals, such as Panda for China, Kitten for Iran, Bear for Russia, Chollima for North Korea, and Spider for criminals. Mandiant, owned by Google, numbers them: APT, for advanced persistent threat, for state groups; FIN for financially motivated criminals; and UNC, for uncategorized, for clusters it has not yet tied to a known group.

Attribution, deciding who was behind an attack, draws on the malware, the servers, the working hours, the mistakes, and intelligence from outside the computer: informants, intercepts, and seized devices. It is a judgment with a level of confidence, not a proof, and attackers plant false clues. Russian military hackers who disrupted the opening of the 2018 Winter Olympics in South Korea, with malware later named Olympic Destroyer, left code designed to point at North Korea and China. Governments sometimes publish their attributions and act on them (Module 9).

Worked example — Colonial Pipeline, Step by Step

On May 7, 2021, Colonial Pipeline, which carries about 45% of the fuel used on America's East Coast, found a ransom note on its systems. The attackers, affiliates of a ransomware brand called DarkSide, had signed in through an old remote-access account for its virtual private network (VPN), a service that gives outside users entry to an internal network. The account was no longer in use, had no MFA, and its password had appeared in a leak; that was initial access with a stolen credential. They spread through the company's office systems and copied about 100 gigabytes of data before encrypting computers, the steps of lateral movement, exfiltration, and impact. The pipeline's own control systems were not reached, but Colonial shut the pipeline for about five days as a precaution, partly because it could not bill for fuel, and drivers queued at empty stations. Colonial paid 75 bitcoin, then about $4.4 million, and on June 7 the Justice Department announced that it had recovered 63.7 of them, then worth about $2.3 million. A second factor on one unused account would have stopped the chain at its first link.

Recommended Reading / Reference

Book: Scott J. Shapiro, *Fancy Bear Goes Phishing: The Dark History of the Information Age, in Five Extraordinary Hacks* (2023), which explains how attacks work through five famous cases

Reference: MITRE ATT&CK (attack.mitre.org), the catalog of attackers' tactics and techniques that defenders use to map their detections

Module 2 Key Points

Adversaries and the Anatomy of an Intrusion

• Criminals commit most attacks; states, hacktivists, and insiders differ from them more in motive than in method.

• Cybercrime is specialized: access brokers, password stealers, and ransomware affiliates each sell one step of an attack.

• Exploited flaws began 31% of breaches in Verizon's 2026 report, ahead of stolen credentials for the first time in 19 years.

• Third parties were involved in 48% of breaches, up from 30% a year earlier.

• Most intrusions follow one chain: access, persistence, escalation, spreading, then theft, encryption, wiping, or fraud.

Module 2 Quiz

Check Your Understanding

1. In Verizon's 2026 report, what most often began a breach?

A) Phishing emails
B) Exploited software vulnerabilities
C) Stolen passwords
D) Physical theft of laptops and phones from offices

2. What does an initial access broker do?

A) Writes ransomware
B) Launders crypto payments
C) Negotiates ransoms
D) Breaks into networks and sells that entry to other criminals, such as ransomware affiliates

3. In Microsoft's naming scheme, which country do 'Typhoon' groups work for?

A) Russia, as with the military unit behind NotPetya
B) Iran, whose groups attacked American water utilities
C) China
D) North Korea, the most active thief of cryptocurrency

4. How did attackers first get into Colonial Pipeline's network in 2021?

A) A phishing email
B) A supplier's update
C) A password for an unused VPN account with no MFA
D) A flaw in the pipeline's control systems, reached directly over the internet from abroad

5. What does 'living off the land' mean in an intrusion?

A) Using the tools already built into the victim's systems, such as administration utilities, so that the activity blends in with normal work
B) Stealing electricity
C) Hiding in cloud storage
D) Working from a country that will not extradite

Module 3 — Identity: Passwords, MFA, and Passkeys

The Front Door, and a Household That Lost Its Keys

3a. Why Passwords Fail

Signing in is where most attacks on people and companies meet their target, because a valid account opens doors that no flaw needs to be found for. Passwords fail in four main ways. People reuse them, so a password leaked by one website unlocks another. Attackers try the leaked pairs of usernames and passwords at other sites by the million, a technique called credential stuffing. They try a few common passwords against many accounts, called password spraying, which stays under the limits that lock an account after repeated failures. And they steal passwords directly, with phishing pages and infostealers (Module 2).

Microsoft's Digital Defense Report of October 2025 found that more than 97% of the attacks on identities it saw were password attacks, mostly spraying and brute force. The same report found that MFA blocks more than 99% of attempts to break into accounts, even when the attacker holds the right password. Few controls in security buy so much for so little.

The old advice made matters worse. Rules that demanded a capital letter, a number, and a symbol, and changes every 90 days, pushed people toward predictable patterns such as Summer2026!. NIST's digital identity guidelines (Special Publication 800-63B) advise against both rules and favor length instead: a long passphrase that is changed only when there is evidence it has leaked. A password manager, an application that creates and stores a different random password for every site, removes reuse altogether.

3b. Multifactor Authentication, and How It Is Beaten

MFA asks for two or more kinds of proof: something you know, such as a password; something you have, such as a phone or a hardware key; or something you are, such as a fingerprint. Its forms differ in strength. A code sent by text message can be intercepted by an attacker who persuades a mobile carrier to move the victim's number to a new SIM card, the chip that ties a phone to its number, a trick called SIM swapping. Codes from an authenticator app avoid that. A push notification asking the user to approve a sign-in can be defeated by sending dozens of requests until a tired user taps yes, which is called MFA fatigue.

The more serious threat is the adversary-in-the-middle (AiTM) phishing kit. The victim clicks a link to a convincing copy of a sign-in page, which quietly passes everything through to the real site. The user types the password and the code, the real site signs the user in, and the kit captures the session cookie the site issues in return. The attacker then uses the cookie, and the account is open. Kits of this kind are sold as services. Microsoft's Digital Defense Report of October 2026, as summarized in coverage, counted AiTM techniques in 44.6% of the phishing it identified and found password attacks down 26% in a year, a sign that attackers are moving past the password to the session.

Phishing-resistant MFA defeats these kits because the credential checks which website it is talking to. Hardware security keys and passkeys built on the standards of the FIDO (Fast Identity Online) Alliance, an industry group, will answer only the genuine site, so a look-alike page gets nothing it can use. Smart cards

issued to federal employees work the same way. Since a memorandum of January 2022 (M-22-09), federal agencies have been required to move their staff to phishing-resistant methods.

3c. Passkeys

A passkey replaces the password with a pair of cryptographic keys. When a user creates one, the device makes a private key, which never leaves the user's control, and gives the website a matching public key. To sign in, the site sends a challenge, the device signs it with the private key after the user unlocks the device with a fingerprint, face, or personal identification number (PIN), and the site checks the signature with the public key. There is nothing to reuse, nothing to type into a fake page, and nothing worth stealing from the website's servers, since a public key opens nothing on its own.

Passkeys come in two kinds. Synced passkeys are copied across a person's devices through a cloud account, such as Apple's iCloud Keychain or Google Password Manager, so a lost phone does not lose them; their safety then rests on that cloud account. Device-bound passkeys stay on one piece of hardware, such as a security key, and suit administrators and other high-risk users. Microsoft made new consumer accounts passwordless by default on May 1, 2025, and Amazon said on October 20, 2026 that more than 500 million of its customers had turned passkeys on.

The FIDO Alliance's Passkey Index of October 15, 2026, pooled from 15 large consumer services, found that 36% of successful sign-ins used a passkey, up from 26% a year earlier, and that 38% of accounts had one. Passkey sign-ins succeeded 93% of the time, against 78% for passwords with traditional MFA, and took a third as long, saving 27 seconds each. The services' figures are confidential and unweighted, so a small service counts as much as Amazon.

3d. Recovery and the Help Desk

Every sign-in method has a back door: the way back in for a user who has lost it. Attackers aim there. A forgotten-password link sent to an email account the attacker controls, a text code sent to a swapped SIM, or a help desk that resets credentials over the telephone all bypass whatever strong method guards the front. The account is only as strong as its weakest recovery path.

The costliest attacks of recent years began with a telephone call. In September 2023 attackers took over systems at MGM Resorts, the casino company, after calling its information technology (IT) help desk and posing as an employee whose details they had found online, by the attackers' own account, which MGM has not confirmed. MGM estimated the cost at about $100 million, and its slot machines, hotel keys, and booking systems were disrupted for days. Similar calls opened the British retail attacks (Module 12). Help desks defend themselves by calling back on a number already on file, checking identity by video against an employee photo, and refusing to reset administrators' credentials over the phone at all.

3e. Sessions, Tokens, and Machine Identities

After sign-in, a website issues a session token, usually a cookie, that stands in for the user until it expires. Whoever holds it is treated as the user, so a stolen token bypasses passwords and passkeys alike. Infostealers harvest these tokens from browsers, and AiTM kits capture them in transit. In 2024 criminals used credentials stolen by infostealers, some of them years old, to sign in to customers' accounts at Snowflake, a cloud data company, where MFA was not required. About 165 organizations were exposed, among them AT&T, whose records of calls and texts by nearly all its mobile customers were taken.

The defenses shorten and tie down sessions. Sites can expire them sooner, ask for a fresh sign-in before sensitive changes such as a new email address or a bank transfer, and revoke every session when a password changes. Google's Chrome browser began in 2026 to bind sessions to the device's security chip, a feature called Device Bound Session Credentials, so that a cookie copied to another computer no longer works; websites must add support for it. Microsoft's October 2026 report found that 52.2% of the intrusions that began with a valid account went on to steal more credentials, the chain this section describes.

Machines have identities too. Programs sign in to each other with keys, tokens, and service accounts, and these non-human identities far outnumber employees in most companies. They rarely have MFA, often have broad rights, and are seldom rotated, which makes them prized by attackers (Module 7). AI agents that act for users, reading email or changing systems, add a new kind of machine identity (Module 10).

Worked example — The Harper Household, Part 1: The Takeover

The Harpers live in Sacramento, California: Dana, an accountant; Marcus, a teacher; their son Leo, 15; and, a few streets away, Dana's father Walt, 74. In August 2026 Leo installs a free cheat tool for an online game on the family desktop, where everyone shares one administrator account. The tool carries an infostealer, which copies the browser's saved passwords and live session cookies for 23 accounts and sends them to its operator, who sells the bundle for a few dollars. Two weeks later a buyer loads Dana's webmail cookie into his own browser and is inside her mailbox with no password and no MFA prompt. He sets a rule that forwards any message containing 'bank', 'statement', or 'code' to an outside address and deletes the original. He uses the forgotten-password link at Dana's brokerage, which sends the reset to her email, changes her phone number on file, links a new bank account, and requests a transfer of $18,400. Separately, a reused password from an old breach opens Marcus's account at a retailer, where $640 of gift cards are ordered. The brokerage holds transfers to new accounts for three days and texts the old phone number about the change, which is how Dana finds out. Dana calls the brokerage, which freezes the transfer. She signs out every session on her email, deletes the forwarding rule, changes the password, adds a passkey, and checks her recovery address and phone. The desktop is wiped and reinstalled, every password is changed through a password manager, and Leo gets his own account without administrator rights. Dana reports the crime to the FBI's IC3 and to the FTC's IdentityTheft.gov. Passkeys alone would not have stopped this attack, because the infostealer took sessions, not passwords. A rule at the brokerage requiring a fresh passkey sign-in before any change of contact details would have stopped it.

3f. Contested: Are Passkeys Ready to Replace Passwords?

The case for. Passkeys remove the weaknesses that cause most account takeovers (3c). They are faster and more reliable for users, as the Index's success rates show, and the largest platforms now offer or default to them. Advocates argue that every service that keeps passwords as an equal option keeps the weakest door open, and that the change should be pushed hard while attention is high.

The case against. Skeptics answer that most services still fall back to email links or text codes, the recovery paths that attackers target (3d). In the Index, eight of the ten services that report both measures had a passkey share of sign-ins below their share of enrolled accounts: users had passkeys but were not routed to them. Moving passkeys between Apple, Google, and Microsoft accounts is still awkward, shared family devices complicate them, and companies cite old systems (38%), budgets (35%), and recovery (33%) as the main obstacles, in the FIDO Alliance's survey of May 2026.

The evidence and the course's reading. The evidence supports both sides on different points: passkeys work where they are used, and their use is still well short of their availability. The course's reading is that

passkeys are ready to be the default sign-in for consumers and the required one for administrators, but not yet able to retire passwords, because recovery and old systems keep the weaker paths open. The security gain comes when those paths are hardened too, and when sensitive changes require a fresh passkey, as the Harper case shows.

Recommended Reading / Reference

Report: FIDO Alliance, *Passkey Index 2026* (October 15, 2026), the pooled sign-in data of large consumer services, eight pages

Reference: NIST, Special Publication 800-63B, *Digital Identity Guidelines: Authentication and Authenticator Management*, the federal standard for passwords, MFA, and authenticators

Module 3 Key Points

Identity: Passwords, MFA, and Passkeys

• Microsoft found over 97% of identity attacks were password attacks, and that MFA blocks over 99% of attempts.

• Phishing kits that relay a sign-in steal the session, beating text and app codes; only phishing-resistant methods stop them.

• Passkeys carried 36% of successful sign-ins at the services in FIDO's October 2026 index, up from 26% a year earlier.

• Attackers target recovery paths and help desks; MGM's 2023 breach, costing about $100 million, began with a call.

• Stolen session cookies bypass passwords and passkeys alike, so sensitive changes should require a fresh sign-in.

Module 3 Quiz

Check Your Understanding

1. What does a stolen session cookie let a criminal skip?

A) Only the username, since the password is still needed
B) The recovery questions, but not the password
C) Paying for the stolen bundle of data
D) Signing in

2. What share of successful sign-ins used a passkey in FIDO's index of October 2026?

A) About 5%
B) About 36%
C) About 74%, most sign-ins at large services
D) About 93%, the same as the success rate

3. Why do passkeys resist phishing?

A) They expire daily
B) Each is bound to the website that created it
C) They are longer
D) They are kept on the website's server, where only the company can read them, never on the user's device

4. How did attackers get into MGM Resorts' systems in 2023?

A) By calling its help desk
B) Through a flaw in a slot machine
C) With a thumb drive
D) By bribing an employee in the casino's security office

5. What does adversary-in-the-middle phishing capture that defeats one-time codes?

A) The victim's card number, typed into a fake shop
B) The device, through a fake charging cable
C) The session
D) Only the username, leaving the password safe

Module 4 — Vulnerabilities and Updates

Flaws, Patches, and the Race Between Them

4a. What a Vulnerability Is

A vulnerability is a flaw in software or hardware that lets someone do what the designers never meant to allow: read memory that belongs to another program, run commands on a server, or skip a password check. An exploit is the code or method that uses the flaw. Some flaws are mistakes in code, such as a program that copies more data into a buffer than the buffer holds, letting an attacker overwrite what lies beyond it. Others are mistakes in design or setup, such as a database left open to the internet with no password.

Memory-safety errors, in which a program reads or writes memory it should not touch, have been the largest single class of serious flaws. Microsoft said in 2019 that about 70% of the flaws it fixed each year were of this kind, and Google's Chromium browser project reported a similar share. Languages that check memory use automatically, such as Rust, remove most of them: Google reported in September 2024 that memory-safety flaws had fallen from 76% of Android's vulnerabilities in 2019 to 24% in 2024 as new code moved to such languages.

A zero-day is a flaw that attackers know about before the maker has a fix, so defenders have had zero days to prepare. Most attacks, though, use flaws whose fixes already exist but have not been applied, sometimes called n-days. Researchers who find a flaw usually tell the maker privately and publish once a fix is out, a practice called coordinated disclosure; many companies pay them bug bounties for it.

4b. Naming and Scoring Flaws

Publicly known flaws get an identifier from the Common Vulnerabilities and Exposures (CVE) program, such as CVE-2025-5777, so that everyone refers to the same flaw. The program, started in 1999, is run by MITRE with funding from the Cybersecurity and Infrastructure Security Agency (CISA), America's national cyber-defense agency. Hundreds of software makers and research groups, called CVE Numbering Authorities, assign the numbers for their own products: 502 of them by March 31, 2026. NIST's National Vulnerability Database (NVD) then adds details, such as which products are affected.

Each flaw also gets a severity score from 0 to 10 under the Common Vulnerability Scoring System (CVSS), whose version 4.0 was released on November 1, 2023. Severity is not risk: a flaw scored 9.8 in software nobody exposes to the internet may matter less than a 7.5 in a device every company runs at its edge. The Exploit Prediction Scoring System (EPSS), run by FIRST, an international association of incident-response teams, estimates instead the chance that a flaw will be exploited in the next 30 days. Its fourth version arrived on March 17, 2025.

The most useful list is CISA's catalog of Known Exploited Vulnerabilities (KEV), started on November 3, 2021, which lists only flaws with reliable evidence of exploitation in real attacks. A binding directive requires federal civilian agencies to fix each listed flaw by a deadline, and companies use the list to set their own priorities. CISA added 245 flaws in 2025, 20% more than in 2024, bringing the catalog to 1,484 entries at the end of the year; it marked 24 of the year's additions as used by ransomware gangs.

4c. The Flood of 2025 and 2026

The number of published flaws is rising fast. The CVE program published 48,185 new records in 2025, 20.6% more than in 2024, by the count of Jerry Gamblin, an independent analyst who has tracked the figures for years. In the first half of 2026 it published 35,364, 49.5% more than in the same months of 2025, a pace that points to more than 70,000 for the year. More numbering authorities, the Linux kernel's among them since February 2024, and AI tools that help researchers find flaws (Module 10) explain much of the rise.

Only a small fraction of these flaws are ever used in attacks. Of the CVEs published in 2025, 184 had been confirmed as exploited by August 2026, fewer than one in 250, in an analysis by Axis Intelligence, a research group. Volume is therefore a poor guide to danger. A security team that tries to fix everything at once fixes nothing quickly, and the flaws that matter wait in the same queue as the ones that never will.

The databases are straining under the load. In April 2025 NIST marked tens of thousands of older records awaiting its analysis as deferred, and on April 15, 2026 it stopped fully analyzing every new CVE. It now gives full treatment only to flaws that are exploited, used by federal agencies, or in critical software, and about 29,000 records published before March 2026 were left unscheduled. Defenders who relied on the NVD's details now draw more on makers' own advisories and on CISA's catalog.

4d. The Edge: Why Attackers Favor VPNs, Firewalls, and Gateways

The devices that sit at the edge of a network, facing the internet, have become attackers' favorite way in. They include VPN gateways, firewalls, email gateways, and file-transfer servers. They are reachable by anyone, they hold the keys to the network behind them, and most cannot run the endpoint detection software that watches ordinary computers (Module 11), so an attacker inside one often goes unseen. Many are patched slowly because updating them interrupts the connections a business runs on.

A single flaw in a popular device can expose thousands of organizations within days, as with a flaw in Citrix's remote-access gateways, CVE-2025-5777, exploited in 2025 and nicknamed CitrixBleed 2. Attackers scan the whole internet for vulnerable devices within hours of a flaw becoming known. Microsoft's report of October 2026, as summarized in coverage, put the median time to turn a newly found flaw into a working attack at well under 24 hours. Edge devices should be patched first, in days rather than weeks, and those that can no longer be updated should be replaced.

4e. Patching by Priority, and Software Past Its Life

Patching starts with an inventory, since no one can update a device they do not know they own. Microsoft releases its fixes on the second Tuesday of each month, called Patch Tuesday, and other makers on their own schedules. Mature programs rank flaws by whether they are on CISA's list, whether the system faces the internet, and the EPSS score, and set deadlines by rank. Where a fix cannot be applied quickly, they reduce the exposure another way, by switching off the vulnerable feature or blocking access to it.

Software past its end of life gets no more fixes, so every flaw found later stays open. Microsoft ended support for Windows 10 on October 14, 2025, while hundreds of millions of computers still ran it. It offers home users a year of Extended Security Updates, free with its Windows Backup service or Rewards points or for $30, and on June 25, 2026 it extended that consumer program to October 12, 2027; businesses can buy updates for longer. Routers, cameras, and other connected devices often stop receiving updates after a few years, and many never received any.

For households, the rules are simpler: turn on automatic updates, restart when asked, replace routers and devices their makers no longer support, and remove software no longer used. Phones and browsers update themselves; the laggards are usually the router in the closet and the old laptop in the drawer.

4f. Who Keeps the Catalog Running

The CVE program nearly stopped in 2025. On April 15 a letter from MITRE warned that its contract with the government to run the program would expire the next day. CISA exercised an option on the contract that night, keeping MITRE in place for 11 more months, to March 16, 2026; federal records valued the contract at $57.8 million. Members of the program's board set up the CVE Foundation to argue for funding from more than one government, and ENISA, the European Union's cybersecurity agency, launched a European Vulnerability Database in May 2025.

By that deadline the crisis had passed: CISA had told the program's board in January that there would be no funding cliff, and the contract was renegotiated, though its terms were not published. The episode showed how much of the world's defense rests on one American contract. The CVE system remains run by MITRE under CISA's sponsorship, with the alternatives in Europe and elsewhere as backups if that changes.

4g. Contested: Should Software Makers Be Liable for Insecure Products?

The case for. Advocates of liability argue that software makers are best placed to prevent flaws, while their customers cannot inspect the code they buy and bear the losses when it fails, a market failure that only law can correct. They point to the Cyber Safety Review Board, a federal panel, which concluded in April 2024 that a 'cascade of security failures' at Microsoft let Chinese hackers read the email of senior American officials in 2023. They note that CISA's voluntary Secure by Design pledge, signed by 68 companies at its launch in May 2024, has no penalties, and that Europe's Cyber Resilience Act now imposes duties with fines (Module 14).

The case against. Opponents answer that liability would raise costs, slow new products, and fall hardest on small firms and on the unpaid volunteers who maintain open-source software. 'Reasonable security' is hard to define, they argue, and contracts and insurance already share the risk. The federal government has moved their way. The March 2023 national strategy had proposed shifting liability to software makers, but the panel's members were dismissed in January 2025. Executive Order 14306, of June 6, 2025, also dropped the requirement that federal software suppliers attest to secure development, calling it burdensome.

The evidence and the course's reading. The evidence shows that makers' choices change outcomes: Android's move to memory-safe code sharply cut its memory-safety flaws (4a). It does not yet show how liability would work, since Europe's law has yet to take full effect (Module 14). The course's reading is that liability for clearly bad practices is justified, among them default passwords, products with no way to update, and known exploited flaws left unfixed. Broad liability for any flaw would be premature.

Worked example — A Month of Patches, Ranked

A distribution company with 1,200 computers, 90 servers, and four internet-facing devices receives 120 new security fixes in one month. Two of the flaws are already in CISA's catalog of exploited vulnerabilities, and one of those affects its VPN gateway. Its rules: anything on the catalog that faces the internet is fixed within 72 hours, other catalog entries and flaws with an EPSS score above 10% within 14 days, and everything else within 30 days. The VPN gateway is patched the same night, after hours, with the vendor's advice to check for signs of compromise; the team also looks for accounts or files the attackers might have planted before the fix. The second catalog flaw, in a desktop program, goes out to all computers within a week. Of the remaining 118 fixes, 9 have EPSS scores above 10% and are scheduled for the second week, and the other 109 go out with the normal monthly cycle. The team spends its first hours on three fixes rather than spreading them across 120, and the one flaw most likely to be used against it is closed before most attackers have scanned for it.

Recommended Reading / Reference

Book: Bruce Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World* (2018), which argues that insecure products are a market failure that law must correct

Reference: CISA, Known Exploited Vulnerabilities Catalog, the list of flaws with evidence of real attacks, updated as new ones are confirmed

Module 4 Key Points

Vulnerabilities and Updates

• 48,185 CVEs were published in 2025, and the first half of 2026 ran 49.5% ahead; well under 1% are exploited.

• CISA's catalog of exploited flaws had 1,484 entries at the end of 2025; fix those and internet-facing devices first.

• Edge devices such as VPNs and firewalls face the internet, hold network access, and usually cannot run detection software.

• NIST stopped fully analyzing every CVE in April 2026, keeping full analysis for exploited, federal, and critical software.

• Free Windows 10 updates for home users now end October 12, 2027; unsupported devices keep every flaw found later.

Module 4 Quiz

Check Your Understanding

1. What does CISA's Known Exploited Vulnerabilities catalog list?

A) Every CVE ever published
B) Flaws with a CVSS score of 10
C) Software past its end of life
D) Flaws with reliable evidence of exploitation in real attacks, which federal agencies must fix by set deadlines

2. How many new CVE records were published in 2025, by the count the course uses?

A) 48,185
B) 4,818
C) About 4.8 million, one for every program in use
D) 481,850, ten times the 2024 count

3. Why do attackers favor VPN gateways, firewalls, and other edge devices?

A) They are cheap
B) They store no data
C) They face the internet, hold network access, and often cannot run detection software
D) Their makers publish no fixes at all, so every flaw stays open for years after it is found and reported

4. Until when did Microsoft extend its consumer security updates for Windows 10?

A) October 14, 2025, the end of support
B) October 13, 2026, as first announced
C) Indefinitely, for any computer enrolled before 2026
D) October 12, 2027

5. What happened to the CVE program in April 2025?

A) It moved to Europe
B) Its contract nearly lapsed, until CISA used an option to keep MITRE running it for 11 more months
C) It stopped issuing numbers
D) It merged with the NVD

Module 5 — Social Engineering, Scams, and Fraud

Attacks on People Rather Than Machines

5a. Why People Are the Target

Social engineering means manipulating people into doing what an attacker wants: clicking a link, reading out a code, approving a payment, or handing over access. It is usually cheaper than breaking software, and it works against organizations with excellent technical defenses. The levers are old ones. Authority makes a request seem to come from a boss, a bank, or the police. Urgency leaves no time to check, fear makes people act to avoid a loss, and secrecy keeps them from asking anyone else.

Most scams follow the same arc: a contact that seems legitimate, a story that creates pressure, a request to move money or reveal credentials, and an instruction to tell no one. Recognizing the arc matters more than recognizing any particular script, since the scripts change every few months while the arc stays the same.

5b. Phishing in Every Channel

Phishing began with email and has moved to every channel. Smishing uses text messages, such as the waves of texts in 2025 and 2026 claiming an unpaid road toll or an undelivered parcel. Vishing uses voice calls, sometimes on workplace platforms such as Microsoft Teams, where a caller poses as the IT department. Quishing hides a malicious link in a quick response (QR) code, which a phone's camera opens without the address being read; Microsoft's report of October 2026 counted more than 145 million phishing attacks using QR codes in a year.

A newer trick, known as ClickFix, shows a fake error or a fake 'prove you are human' box on a website and tells the visitor to fix it by pasting a command into the computer, which installs malware. It works because it turns the victim into the attacker's installer and gets past tools that block suspicious downloads. Generative AI has also removed the clumsy grammar that once gave phishing away (Module 10).

5c. Business Email Compromise and Cloned Voices

Business email compromise (BEC) targets companies' payments. A criminal takes over or imitates the email of an executive, a supplier, or a lawyer and asks for a transfer, often a change of the bank details on a real invoice. The FBI received 24,768 complaints of BEC in 2025, with losses of $3.047 billion, making it one of the costliest crimes it tracks. The defense is procedural rather than technical: any change of payment details is confirmed by calling the supplier on a number already on file, never one in the email, and large payments need two people's approval.

Cloned voices and faces have raised the stakes. In early 2024 an employee in the Hong Kong office of Arup, a British engineering firm, joined a video call with what appeared to be the company's chief financial officer and several colleagues. All of them were deepfakes, videos generated by AI. On their instructions the employee made 15 transfers totaling 200 million Hong Kong dollars, about $25 million. A few seconds of a person's recorded voice is enough to clone it convincingly, so a familiar voice on the telephone no longer proves who is speaking.

5d. Investment Scams and the Compounds That Run Them

Investment fraud was the costliest crime in the FBI's report for 2025: $8.649 billion, of which $7.228 billion involved cryptocurrency investment schemes. Cryptocurrency figured in $11.37 billion of the year's losses in all. The FTC's data agree on the ranking: investment scams caused the largest losses reported to it, more than $7.9 billion, while impostor scams were the most often reported, with more than a million reports.

The typical scheme is known as pig butchering, after the practice of fattening an animal before slaughter. A stranger starts a friendship or romance by text or on social media, sometimes after a 'wrong number' message, and over weeks introduces an investment platform. The victim deposits money and sees fake profits on a sham trading site, deposits more, and finds out only when withdrawals are blocked or new 'taxes' are demanded. The losses often include retirement savings and loans taken out to invest.

Many of these scams are run from compounds in Myanmar, Cambodia, and Laos, where trafficked workers are forced to work the phones. The United Nations' human-rights office estimated in August 2023 that at least 120,000 people were held in such compounds in Myanmar and about 100,000 in Cambodia; in February 2026 it put the total across Southeast Asia at at least 300,000. The UN Office on Drugs and Crime estimated in July 2026 that the networks took $88.3 billion to $114.1 billion from victims in 2025, and it found people from more than 80 countries in the compounds.

The largest case so far began on October 14, 2025, when American prosecutors in Brooklyn charged Chen Zhi, chairman of Cambodia's Prince Group, with running forced-labor scam compounds. The Justice Department also filed a civil forfeiture complaint against about 127,271 bitcoin, then worth about $15 billion, the largest forfeiture action in its history, and America and Britain imposed sanctions on the group. Cambodia arrested Chen and sent him to China on January 7, 2026. The bitcoin remains in American custody, and a court must still order its forfeiture before any of it can go to victims.

5e. Fraud Against Older People

People aged 60 and over filed 201,266 complaints with the FBI in 2025 and reported losses of $7.748 billion. Older people are targeted because many have savings, home equity, and landlines, and because they are often more trusting of authority. Tech-support scams, in which a pop-up or caller claims a computer is infected, cost Americans of all ages $2.135 billion in 2025, and they are often the first step of a longer con.

The FBI described the most damaging version in 2023 as the phantom hacker scam. A fake technician says the victim's computer has been hacked. A fake bank employee then says the accounts are compromised, and a fake government agent instructs the victim to move savings to a 'safe' account to protect them. Each caller confirms the others' story, and the victim is told that the bank itself cannot be trusted. Victims have lost entire retirement accounts in a single transfer.

5f. Habits and Controls That Work

A few habits defeat most scams. Never move money, buy gift cards, or read out a code because someone who contacted you asked; hang up and call back on a number you already have, such as the one on a bank card. Treat secrecy and urgency as warning signs in themselves. Agree a family code word for emergencies, since a cloned voice cannot know it. Use the controls banks offer: alerts on every transfer, limits on new payees, and a trusted contact whom the bank may call if it sees unusual activity.

Speed matters after a loss. The FBI's IC3 runs a team that works with banks to freeze transfers reported within a few days, and a bank told at once can sometimes recall a wire. Reports to the IC3 and the FTC also build the cases that lead to arrests and forfeitures, such as the Prince Group's. Shame keeps many victims silent, and scammers rely on it.

Worked example — The Harper Household, Part 2: The Call

Ten days after the takeover in Module 3, Dana's father Walt gets a call. The caller ID shows his bank's real number, which is easily faked. The caller knows Walt's name, that Dana is his daughter, and that her brokerage account was attacked last week, all details taken from Dana's mailbox, which held messages with her father and from the brokerage. He says Walt's accounts are next, that a bank employee is involved, and that Walt must move his $42,000 of savings to a 'protected account' today and tell no one, not even Dana, until a federal agent calls him. A second man, claiming to be from the Treasury, calls an hour later to confirm the story. After Part 1 the family had agreed two rules: any call about money is ended and returned on the number printed on a card, and anyone asking for secrecy about money is to be reported to the rest of the family. Walt hangs up, calls the bank on the number on his card, and learns that nothing is wrong. Dana then has the bank add a 24-hour hold on new payees and name her as Walt's trusted contact. The lesson is that one breach feeds the next: the data stolen from Dana made the scam against her father convincing.

5g. Contested: Who Should Pay When a Scam Succeeds?

The case for making banks and platforms pay. Advocates argue that banks run the payment systems and the fraud-detection tools, so they are best placed to stop scam payments, and that they will invest in stopping them only if they bear the losses. Social media companies and telecoms carry the scam ads, messages, and calls. Britain made reimbursement mandatory from October 7, 2024: banks must repay victims of authorized push payment fraud, payments the victim was tricked into making, up to £85,000, with the sending and receiving banks sharing the cost. Australia passed a law in 2025 obliging banks, telecoms, and digital platforms to prevent scams, with compensation where they fail.

The case against. Opponents warn of moral hazard: customers who expect a refund may take less care. Some claims are false, filed by people who made the payment knowingly. Costs are passed on to every customer, and banks cannot see the social media chat or the telephone call in which the scam took place. American law reflects this view: the Electronic Fund Transfer Act's rules protect consumers against transfers they did not authorize, not against ones they made while deceived. The Consumer Financial Protection Bureau sued the operator of Zelle and three large banks over scam losses in December 2024, then dropped the case in March 2025.

The evidence and the course's reading. Britain's Payment Systems Regulator reported that 88% of the money lost to such fraud was repaid in the 18 months to March 31, 2026, £316 million, against about 66% before the rules. It reported no collapse in care by customers, though the scheme is young and the regulator is being folded into the Financial Conduct Authority. The course's reading is that losses should be shared among banks, telecoms, and platforms, with a cap and an exception for gross negligence. Leaving them with victims leaves the firms best placed to stop scams with little reason to.

Recommended Reading / Reference

Book: Kevin D. Mitnick and William L. Simon, *The Art of Deception: Controlling the Human Element of Security* (2002), the classic account of social engineering by a once-notorious hacker

Report: FBI Internet Crime Complaint Center, *2025 Internet Crime Report* (April 2026), with losses by type of crime, age, and state

Module 5 Key Points

Social Engineering, Scams, and Fraud

• Investment fraud was the costliest crime in the FBI's 2025 report, $8.65 billion, $7.23 billion of it in cryptocurrency schemes.

• Business email compromise cost $3.05 billion in 2025; a callback on a number already on file defeats most of it.

• The UN's human-rights office estimated in 2026 that at least 300,000 people were held in Southeast Asia's scam compounds.

• Americans aged 60 and over reported $7.75 billion of losses to the FBI in 2025.

• Never move money because a caller asks; hang up and call back on a number you already have.

Module 5 Quiz

Check Your Understanding

1. Which kind of fraud caused the largest losses in the FBI's report for 2025?

A) Tech-support scams
B) Business email compromise, the costliest for companies
C) Investment fraud
D) Romance

2. How much did Arup's Hong Kong office lose to a deepfaked video call in 2024?

A) About $2,500
B) About $250,000
C) About $25 billion, the largest corporate fraud on record
D) About $25 million, in 15 transfers

3. What sets a 'pig-butchering' investment scam apart from a quick phishing scam?

A) The scammer builds a relationship over weeks or months and shows fake profits on a sham trading site before taking everything
B) It targets only farms
C) It uses no money
D) It is legal in most countries

4. What share of the money lost to authorized push payment fraud did British banks repay under the rules of October 2024?

A) Almost none of the money lost
B) About 88%
C) About 66%, unchanged from before the rules
D) Only what victims recovered by suing in court

5. What is the safest response to a caller who says your bank account is under attack and asks you to move money?

A) Move it quickly
B) Ask the caller for an ID
C) Follow the caller's instructions but keep the receipts, so that the bank can refund you later
D) Hang up and call the number on your card

Module 6 — Ransomware and Extortion

The Business Model That Took Over Cybercrime

6a. How Ransomware Works

Ransomware encrypts a victim's files with a key only the attacker holds, then demands payment, usually in cryptocurrency, for the tool that unlocks them. Early ransomware hit individual computers for a few hundred dollars. Modern ransomware is a planned intrusion against a whole organization: the attacker spends days or weeks inside, finds and destroys the backups, and only then encrypts everything at once, often in the early hours of a weekend or holiday when few staff are watching.

Since 2019 the gangs have also stolen data before encrypting it, a practice called double extortion. Victims who could restore from backups were then threatened with publication of their data on the gang's leak site. Some gangs add pressure by flooding the victim's websites with traffic or by contacting its customers and patients directly. A growing number skip encryption altogether and simply steal and threaten, which is quicker and harder to detect (Module 2).

A related weapon is the wiper, malware that destroys data and only pretends to be ransomware, so that no payment can restore anything. States have used wipers disguised as ransomware to cause damage while muddying who was responsible (Module 9).

6b. The Ransomware Business in 2026

Ransomware runs on the model described in Module 2. Affiliates move between brands, and brands vanish and reappear after takedowns, disputes, and exit scams, in which the operators disappear with the affiliates' share. Leak sites give the only public count of victims, and it is a partial one: a victim appears only if it refused to pay, and not every group posts.

By the count of Check Point, a security company, more than 70 leak sites posted 2,122 new victims in the first quarter of 2026. Qilin led for the third quarter running, with 338. The market had splintered to a record 85 active groups in the third quarter of 2025, then consolidated: in the first quarter of 2026 the ten largest posted 71% of victims, and Qilin, Akira, The Gentlemen, and a revived LockBit together posted 41%. The FBI's report for 2025 listed Akira and Qilin as the variants most often reported to it.

Ransomware's share of breaches keeps growing. Verizon's report of May 2026 found ransomware in 48% of the breaches it studied, up from 44% a year earlier, and 96% of its ransomware victims were small and mid-size organizations. Large companies make the news, but the typical victim is a regional manufacturer, a school district, a dental group, or a town.

6c. What Payment Buys

Fewer victims pay than ever. Chainalysis estimated on February 26, 2026 that victims paid about $820 million in ransoms traceable on public blockchains in 2025, a figure likely to approach $900 million as more payments are traced, against a revised $892 million in 2024. Only 28% of victims paid, a record low. Verizon found that 69% of the ransomware victims in its data did not pay, up from 65%, and that the median ransom among those who did was $139,875. Coveware, which negotiates for victims, put the median payment at $150,000 in the second quarter of 2026 and called the payment rate a record low.

Victims pay less because more of them can recover from backups and because of the takedowns and sanctions described in 6f. A payment buys a decryption tool, which may be slow or faulty, and a promise to delete the stolen data. The promise is worth little: data from paying victims has resurfaced on other sites, and some victims have been extorted a second time. Paying also funds the next attack.

The FBI's report counted 3,611 ransomware complaints in 2025, with $32.3 million of losses, far below the true cost, because it counts only what victims report as lost and excludes downtime, recovery, and lost business. The costs that matter are those of the outage. For a large victim they run to hundreds of millions of dollars (Module 12), and for a small one they can close the business.

6d. The Case of Scale: Change Healthcare

On February 21, 2024, an affiliate of ALPHV, also called BlackCat, encrypted systems at Change Healthcare, a unit of UnitedHealth Group that processes about 15 billion medical claims a year. The attackers had entered nine days earlier through a remote-access portal, made by Citrix, that did not require MFA. Pharmacies could not check insurance, and doctors' offices went weeks without payments; some borrowed to make payroll. UnitedHealth paid a ransom of $22 million, and its chief executive told Congress that the missing second factor had let the attackers in.

The payment did not end the extortion. ALPHV's operators disappeared with the money without paying their affiliate, who then took the stolen data to another gang, RansomHub, which demanded a second payment and posted samples. In July 2025 UnitedHealth put the final count of people whose data were exposed at about 192.7 million, the largest breach of health data in American history, and it had projected costs of about $2.9 billion for 2024.

6e. Ransomware Run by an AI Agent

On July 1, 2026, Sysdig, a company that secures cloud computing, published what it called the first documented ransomware attack carried out end to end by an AI agent (Module 10). It named the operation JadePuffer. The agent broke into an internet-facing server running Langflow, an open-source tool for building AI applications, through a flaw that had been fixed more than a year earlier.

The agent then gathered credentials, explored the network, found a storage server still using its factory password, and encrypted 1,342 configuration records before deleting the originals. When one of its own programs failed, it diagnosed the error and fixed it in about 31 seconds, without human help. On July 20 Sysdig reported that the same operator had returned with ransomware built to destroy AI models and their data. No single step was new or sophisticated. What changed was that a machine strung them together (Module 10).

6f. How Governments Respond

Governments fight ransomware with takedowns, sanctions, and rules on payments. Police in several countries have seized the servers of gangs and handed decryption keys to victims, and America's Treasury has imposed sanctions on gang members and on cryptocurrency exchanges that launder for them, which makes paying a sanctioned group illegal. More than 70 countries and organizations belong to the Counter Ransomware Initiative, begun by the United States in 2021, whose members have pledged that their governments will not pay ransoms.

Rules on payments are spreading. Since May 30, 2025, Australian businesses with turnover above A$3 million must report any ransom payment within 72 hours; at least 75 reported paying in the first eight months, according to records released under freedom of information. Britain announced on July 22, 2025 that it would ban payments by public bodies and operators of critical infrastructure and require other organizations to report before paying. No bill had been introduced by October 2026. North Carolina and Florida have barred their state agencies and local governments from paying since 2022. In America, an obligation to report payments awaits a federal rule (Modules 11 and 14).

Worked example — Harrow Precision, Part 1: The Break-In

Harrow Precision makes machined parts for carmakers and medical-device companies. It has 900 employees in three plants in Ohio and Indiana, sales of $310 million a year, and a 12-person IT team. On a Thursday in October, a flaw in its VPN gateway is added to CISA's catalog of exploited vulnerabilities; Harrow's rules give all patches 30 days, so the fix is scheduled for the next maintenance weekend. An access broker exploits the flaw that night, plants a web shell, and sells the entry to a ransomware affiliate for $15,000. Over the next eight days the affiliate dumps passwords from a server's memory, finds an old service account with administrator rights and a weak password, and spreads across the network using Windows' own tools. It copies 380 gigabytes to a cloud storage service, among them engineering drawings, customer contracts, and personnel files holding the Social Security numbers of 3,100 current and former employees. Harrow's backups sit on a server joined to the same network and signed in with the same administrator accounts, so the affiliate deletes them. At 2:10 a.m. on a Saturday it pushes ransomware to every Windows machine at once. By 2:40, 1,400 of the company's 1,650 computers and servers are encrypted, including the systems that schedule production and ship orders. The machine tools on the plant floors still run, but no one can tell them what to make. The ransom note asks for $4.2 million in bitcoin and promises to publish the stolen files in seven days if the company does not pay. Modules 11 to 14 follow what happens next.

6g. Contested: Should Paying Ransoms Be Banned?

The case for a ban. Supporters argue that ransomware exists because it pays, and that a ban removes the incentive: criminals would stop attacking targets that cannot pay. Every payment funds more attacks, they note, and the promise to delete stolen data is often broken. Public bodies spending taxpayers' money on criminals is especially hard to defend, which is why Britain chose to start with them, and payments to sanctioned groups are already illegal.

The case against. Opponents answer that some victims face harm far worse than the payment: a hospital whose systems are down, or a small firm that will close. A ban would punish the victim, push payments underground through intermediaries, and discourage victims from reporting at all, which would leave governments knowing less. Criminals might switch to targets abroad or to plain theft and wiping rather than give up. Enforcing a ban against companies already in crisis would be hard, and fines would add to the damage the attack has done.

The evidence and the course's reading. The payment rate has fallen to record lows without any national ban, which shows that backups and preparation let victims refuse. Nevada, hit statewide in August 2025, refused to pay and restored about 90% of its data within 28 days, at a cost of about $1.5 million. There is too little evidence yet on whether state bans have reduced attacks. The course's reading is that public bodies should be barred from paying, and every organization required to report payments promptly, since reporting helps police trace the money. A general ban for private firms should wait until recovery is within reach of most of them.

Recommended Reading / Reference

Book: Renee Dudley and Daniel Golden, *The Ransomware Hunting Team: A Band of Misfits' Improbable Crusade to Save the World from Cybercrime* (2022), on the volunteers who break ransomware and free its victims

Report: Chainalysis, *The 2026 Crypto Crime Report*, ransomware chapter (February 26, 2026), the main estimate of payments traced on public blockchains

Module 6 Key Points

Ransomware and Extortion

• Ransomware gangs now steal data before encrypting it; some skip encryption and simply threaten to publish.

• Only 28% of victims paid in 2025, a record low, by Chainalysis's estimate; traced payments were about $820 million.

• Change Healthcare's 2024 breach began at a portal without MFA and exposed data on about 192.7 million people.

• Sysdig documented in July 2026 the first ransomware attack it judged to be run end to end by an AI agent.

• Paying buys a decryption tool and a promise; it does not guarantee that stolen data is deleted.

Module 6 Quiz

Check Your Understanding

1. What is double extortion in ransomware?

A) Charging twice the ransom
B) Stealing a copy of the victim's data before encrypting it, then threatening to publish the data if the ransom is not paid
C) Attacking two companies at once
D) Encrypting the backups as well

2. Which group led ransomware leak-site postings for a third straight quarter in early 2026, by Check Point's count?

A) Qilin
B) LockBit, after its takedown in 2024
C) Clop, the group behind MOVEit
D) DarkSide, which hit Colonial Pipeline

3. What share of ransomware victims paid in 2025, by Chainalysis's estimate?

A) 4%
B) About 70%, as in most years since 2019
C) 28%
D) Half, the usual share

4. How did attackers get into Change Healthcare in February 2024?

A) A phishing email
B) A flaw in a medical device
C) Through a remote-access portal, made by Citrix, that did not require MFA
D) A rogue insider

5. What made JadePuffer notable in July 2026, by Sysdig's account?

A) An AI agent ran every step of the attack
B) It paid its victims
C) It used a new zero-day
D) It was the largest ransomware attack ever recorded, hitting thousands of companies in a single day

Module 7 — Suppliers, Software, and the Cloud

Other People's Code and Other People's Computers

7a. A Vendor's Access Is an Attacker's Access

Every organization runs on other organizations. A mid-size company may rely on hundreds of suppliers that hold its data or can reach its systems: the cloud services it rents, the software it installs, the payroll firm, the help-desk contractor, and the managed service provider (MSP) that runs its computers for it. Each of them is a door. Third parties figured in a growing share of the breaches in Verizon's latest report (Module 2), and an attacker who breaks into one supplier can reach all of its customers at once.

The economics favor such attacks. On July 2, 2021, a ransomware gang called REvil used a flaw in VSA, software from Kaseya that MSPs use to manage their clients' computers, to push ransomware through about 60 MSPs to about 1,500 businesses downstream, by Kaseya's estimate. In Sweden the Coop supermarket chain closed about 800 stores because its tills stopped working. One intrusion produced a thousand victims.

7b. Software Supply Chains: SolarWinds and XZ Utils

A software supply-chain attack plants malicious code in a product before it reaches its users, so that the victims install the attack themselves through a trusted update. The model case was disclosed in December 2020. Russia's foreign intelligence service, the SVR, had inserted a backdoor into updates of Orion, network-management software from SolarWinds. About 18,000 customers installed the tainted update, and the attackers chose about 100 companies and nine federal agencies for deeper intrusions, by the White House's count.

Open-source software, written and shared freely and maintained by volunteers, sits inside nearly every commercial product, and it is a target too. On March 29, 2024, Andres Freund, an engineer at Microsoft, noticed that logins to a test machine were taking about half a second longer than they should. He traced the delay to a backdoor in versions 5.6.0 and 5.6.1 of XZ Utils, a compression library used throughout Linux. Someone using the name Jia Tan had spent about two years earning the trust of its overworked volunteer maintainer before slipping the backdoor in. It was caught before it reached the stable releases most servers run, largely by luck.

A flaw in a common component spreads the same way without any malice. When a flaw in Log4j, a widely used open-source library for logging, was disclosed in December 2021, companies spent weeks working out whether and where they used it, since it was buried inside thousands of products. The question 'are we affected?' is often harder to answer than the fix.

7c. The Shai-Hulud Worm and TeamPCP

Developers build software from packages, ready-made components fetched from public registries; npm, the registry for the JavaScript language, holds millions of them. In September 2025 a worm named Shai-Hulud, after the giant sandworms of the novel *Dune*, began spreading through npm. It stole developers' credentials from their computers and build systems, then used their tokens to publish infected versions of the other

packages they maintained, which infected the next developers who installed them. It compromised more than 180 packages in its first wave.

A second wave in late November 2025, which its authors called the Second Coming, compromised more than 700 packages within hours and created more than 27,000 malicious repositories on GitHub, the code-hosting service, according to the Cloud Security Alliance, an industry group. Smaller waves through 2026 hit packages from AI and security companies, among them LiteLLM and Aqua Security's Trivy scanner. On August 26, 2026, Australian police, working with the FBI, arrested two men aged 21 and 23 near Perth and charged them as members of TeamPCP, the group behind the campaign. Police said its malware had stolen more than 500,000 credentials from more than 1,000 organizations; the charges have not been tested in court.

The registry changed its rules in response. GitHub, which runs npm, stopped issuing its old long-lived tokens on November 5, 2025, limited the tokens that can publish packages to 90 days, and revoked all the old tokens on December 9. It now requires a second factor to publish and encourages trusted publishing, in which packages are published from build systems with short-lived credentials rather than from developers' laptops. A stolen token is now worth much less, and for a much shorter time.

7d. Stolen Tokens for Cloud Software

Companies connect their cloud software to each other with OAuth tokens, permissions that let one application act on a user's data in another without sharing a password. Such tokens are long-lived and rarely watched, and they bypass MFA by design. From at least August 8 to 18, 2025, attackers used tokens stolen from Drift, a chatbot made by Salesloft that plugged into companies' Salesforce accounts, to export data from hundreds of those accounts, then searched it for passwords and keys. Google tracks the group as UNC6395.

Education got its own version in 2026. Instructure, whose Canvas software runs online classes at thousands of schools and universities, detected an intrusion on April 29. On May 7 a group calling itself ShinyHunters defaced the Canvas sign-in pages of about 330 institutions and claimed to hold about 275 million records. On May 11 Instructure said it had reached an agreement with the attacker and received confirmation that the data had been destroyed. It did not say whether it had paid.

Cloud services divide security work under what the providers call the shared responsibility model. The provider secures its data centers, hardware, and core software; the customer secures its own settings, identities, and data. Most cloud breaches come from the customer's side of that line: a storage bucket left public, an account without MFA, or a token granted more access than it needed.

7e. Concentration and Outages

When many organizations depend on one supplier, a single failure becomes a national event, whether or not an attacker is involved. On July 19, 2024, CrowdStrike, a leading maker of endpoint security software, pushed a faulty update to its Falcon software. It crashed about 8.5 million Windows computers, by Microsoft's estimate, grounding flights and disrupting hospitals and broadcasters around the world. Delta Air Lines alone cancelled about 7,000 flights over five days.

Cloud computing concentrates the risk further. Amazon, Microsoft, and Google rent out nearly two-thirds of the world's cloud infrastructure, by the count of Synergy Research Group, a market-research firm, for the second quarter of 2026. On October 20, 2025, a failure in Amazon Web Services' main region in northern

Virginia knocked thousands of applications offline for most of a day. Regional clouds carry the same risk at smaller scale. From about 3:40 a.m. on October 7, 2026, a ransomware attack shut down the East Japan Region 1 of IDC Frontier's IDCF Cloud, a subsidiary of SoftBank, cutting off services for 495 companies and local governments.

7f. When One Firm Stops: Jaguar Land Rover

A large manufacturer's suppliers feel an attack as if it were their own. Jaguar Land Rover (JLR), Britain's largest carmaker and a unit of India's Tata Motors, shut its systems after discovering an intrusion at the start of September 2025. Its plants in Britain and abroad stopped for about five weeks, with production restarting in phases by October 8. Its suppliers, many of them small firms that live on its orders, had nothing to make and no one to bill, and some laid off staff.

On September 29 the British government announced a £1.5 billion loan guarantee for JLR, partly so that it could pay its suppliers. JLR booked £196 million of direct costs from the attack in its July-September quarter. The Cyber Monitoring Centre, an independent British body that rates the severity of cyber events, estimated on October 22, 2025 that the event cost the British economy at least £1.9 billion and affected more than 5,000 British organizations, the costliest cyber event in the country's history. The attack's chief cost was the shutdown itself.

7g. Managing Supplier Risk

Managing supplier risk begins with a list: every supplier, what data it holds, what systems it can reach, and what would stop if it failed. Suppliers are then ranked, and the most critical are checked most closely, through questionnaires, certifications, and evidence of controls. Contracts set the rules: security requirements, notice of any breach within a fixed number of hours, the right to audit, and a plan for leaving. Access is kept to the minimum, protected with MFA, granted only when needed, and watched.

For software, the key tool is a software bill of materials (SBOM), a list of every component and library inside a product, with versions, like the ingredients label on food. With SBOMs, a company can answer 'are we affected?' by searching a database rather than by asking every vendor. America's federal guidance has encouraged SBOMs since 2021, and Europe's Cyber Resilience Act will require them for products sold there (Module 14).

Concentration needs its own plan. A company that cannot do without a particular cloud or security product should know how it would work for a day or a week without it, whether through a second provider, an offline mode, or manual procedures (Module 12). Regulators of banks in Europe now oversee the largest technology suppliers directly for this reason (Module 14).

Worked example — Answering 'Are We Affected?' in an Afternoon

When news of the XZ Utils backdoor broke on a Friday, a software company that keeps an SBOM for each of its 40 products and internal systems searched its inventory for the backdoored versions. Within 20 minutes it had its answer: two internal build containers had pulled the bad version during testing that week, and none of its shipped products contained it. It rebuilt the two containers from an earlier version, checked them for signs of misuse, and told its customers that afternoon that they were unaffected. A peer company without SBOMs spent the following week asking each of its engineering teams and suppliers whether XZ Utils appeared anywhere in their work, and its customers waited for an answer that it could not give with

confidence. The SBOMs did not prevent the risk; they turned a week of uncertainty into an afternoon of work.

Recommended Reading / Reference

Report: Cyber Monitoring Centre, statement on the Jaguar Land Rover cyber event (October 22, 2025), the estimate of the event's cost to the British economy and its method

Reference: OpenSSF, Supply-chain Levels for Software Artifacts (SLSA, slsa.dev), a framework of steps for securing how software is built and published

Module 7 Key Points

Suppliers, Software, and the Cloud

• Russian intelligence hid a backdoor in SolarWinds updates that about 18,000 customers installed; a vendor's access is an attacker's access.

• The Shai-Hulud worm stole publishing tokens to infect npm packages; two alleged members of its group were arrested on August 26, 2026.

• CrowdStrike's faulty update of July 19, 2024 crashed about 8.5 million Windows computers without any attacker.

• Jaguar Land Rover's five-week shutdown in 2025 cost Britain at least £1.9 billion and reached more than 5,000 organizations.

• A software bill of materials lets an organization answer 'are we affected?' in hours rather than weeks.

Module 7 Quiz

Check Your Understanding

1. What did CrowdStrike's faulty update of July 19, 2024 show?

A) That attackers can hide in security updates for years
B) Failures need no attacker
C) That cloud outages affect only small firms that cannot afford backups
D) That Windows computers cannot be updated safely at all

2. How did the Shai-Hulud worm spread through npm?

A) By email attachments
B) Through infected thumb drives left in parking lots
C) By exploiting a flaw in Windows that let it jump between computers on a network
D) By stealing publishing tokens

3. Who found the backdoor in XZ Utils in March 2024?

A) The NSA
B) GitHub's automated scanners, before any release
C) A Microsoft engineer
D) The attacker, who confessed after two years of work on the project

4. What does a software bill of materials list?

A) A product's price
B) The components and libraries inside a piece of software, with their versions, like an ingredients label
C) Every customer of a product
D) The bugs found in each release

5. What did the Cyber Monitoring Centre estimate Jaguar Land Rover's 2025 shutdown cost Britain?

A) £19 million
B) £190 million
C) £19 billion, about 1% of Britain's economic output
D) At least £1.9 billion

Module 8 — Critical Infrastructure and Operational Technology

When Attacks Reach Water, Power, and Hospitals

8a. Operational Technology

Operational technology (OT) is the computing that runs physical processes: the pumps and valves of a water plant, the breakers of a power grid, the robots of a factory. Its building blocks include programmable logic controllers (PLCs), small rugged computers that switch motors and read sensors; supervisory control and data acquisition (SCADA) systems that watch and command them from a control room; and the screens on which operators run them. America defines 16 sectors of critical infrastructure, among them energy, water, health care, communications, and manufacturing.

OT differs from office computing in ways that make it hard to defend. Its first priorities are safety and uptime, not keeping data secret. Its equipment runs for 15 to 30 years, far longer than a laptop, and much of it was designed before anyone expected an attacker. Common industrial protocols, such as Modbus, which dates from 1979, accept commands from anyone who can reach the device, with no password at all. Patching often means stopping the plant, so it waits for a scheduled shutdown that may be a year away.

OT used to sit on separate networks, sealed off from the internet. That separation has eroded as companies connected plants to their office systems for monitoring and billing, and as equipment makers added remote access to service their machines. Engineers describe the ideal as a stack of layers, from business systems at the top to field devices at the bottom, with tightly controlled crossings between them. Many plants fall far short of it.

8b. Water Utilities

America has about 50,000 community water systems, and most are small: a town's water may be run by two or three operators with no IT staff. In November 2023 hackers calling themselves CyberAv3ngers, linked by the American government to Iran's Revolutionary Guards, took over Unitronics controllers, made in Israel, that utilities had left reachable from the internet with their default passwords. One was at a booster station of the water authority of Aliquippa, Pennsylvania, which switched to manual control. In May 2024 the Environmental Protection Agency (EPA) said that more than 70% of the water systems it had inspected since September 2023 did not fully meet basic cybersecurity requirements.

The attacks grew in 2026, the year of the Iran war (Module 9). On April 7 six agencies, CISA, the FBI, the National Security Agency (NSA), the Energy Department, the EPA, and US Cyber Command, warned jointly that Iranian actors were exploiting controllers made by Rockwell Automation and had disrupted water, energy, and government services. On July 26 and 27 attackers reached the control systems of more than 30 community water systems in Minnesota. On July 30 the FBI and EPA reported incidents in at least seven states involving Rockwell's Allen-Bradley MicroLogix controllers exposed to the internet.

CISA said some of the July attacks had forced boil-water notices and sustained manual operation, and it urged utilities to take exposed controllers off the internet at once. No agency had formally attributed the July attacks by October 2026; officials suspected a group aligned with Iran. In every case the attackers needed no secret flaw, only a controller anyone could reach.

8c. Power Grids

The first confirmed attack to cut power came on December 23, 2015, in Ukraine. Hackers of Sandworm, a unit of Russia's military intelligence service, took control of three distribution companies' systems, opened breakers remotely, and left about 225,000 customers without electricity for several hours. A year later a second attack, using malware later named Industroyer, darkened part of Kyiv. In April 2022, weeks into Russia's full-scale invasion, Ukrainian defenders stopped an attempt with a new version, Industroyer2, before it could switch off substations.

America's grid has not suffered such an attack, but it has been prepared for one. On February 7, 2024, CISA, the NSA, and the FBI warned that Volt Typhoon, a group working for China's state, had placed itself inside the IT networks of American communications, energy, transportation, and water operators. It had kept access to some of them for at least five years. Its aim, the agencies said, was not to steal data but to be ready to disrupt those services in a crisis, such as a conflict over Taiwan. It lived off the land (Module 2), which made it very hard to spot.

Dragos, a firm that specializes in industrial security, reported in 2025 that Volt Typhoon had spent about 300 days in 2023 inside the network of the electric and water utility of Littleton, Massachusetts, a town of about 10,000 people. Small utilities matter to such campaigns because they may serve military bases or sit beside larger systems. The bulk power system's operators must meet mandatory security standards, known as the Critical Infrastructure Protection (CIP) standards of the North American Electric Reliability Corporation (NERC), but most water utilities and many smaller power distributors face far lighter rules.

8d. Hospitals

Hospitals are critical infrastructure in the most literal sense. When ransomware strikes one, ambulances are diverted, scans and lab results are delayed, and staff fall back on paper. On June 3, 2024, Qilin's ransomware hit Synnovis, which ran the pathology laboratory services for several large hospitals in south-east London. Thousands of operations and appointments were postponed, and hospitals appealed for donations of the universal blood type because they could not match blood quickly.

In June 2025, King's College Hospital confirmed that a long wait for a blood-test result caused by the attack was among the factors that contributed to a patient's death. Synnovis estimated its own direct losses for 2024 at £32.7 million. Attacks on the companies that serve hospitals, such as laboratories and claims processors (Module 6), can do as much harm as attacks on the hospitals themselves.

8e. Defending Plants and Utilities

The defenses that matter most are unglamorous. Know every controller and how it connects. Take controllers off the open internet, where search engines built to find connected devices list them for anyone to see. Replace default passwords. Allow remote access only through a gateway that requires MFA and logs every session, and keep vendors' remote access switched off until it is needed. Separate the plant network from the office network so that ransomware in one cannot reach the other, and watch the traffic between them.

Above all, keep the ability to run by hand. Operators who can work pumps and valves manually, with written procedures and practiced drills, turn an attack on controllers into an inconvenience. The dependency runs the other way too: when office systems fail, a plant may have to stop even if its controllers are untouched, as

Colonial Pipeline did (Module 2), so plants need ways to keep billing and scheduling going without them. CISA offers small utilities free scanning of their internet-facing systems.

Worked example — A Small Water Utility Closes Its Front Door

A town of 9,000 people runs its water with two operators. A free scan by CISA in August 2026 finds that a controller at one of its pumping stations answers on the internet and still uses the password it shipped with. The operators call their integrator, the firm that installed the system, which comes the next morning. They disconnect the controller from the internet, change every default password, and replace the vendor's always-on remote connection with a small gateway that requires MFA and is switched on only when the vendor calls ahead. They write a one-page procedure for running the station's pumps by hand and practice it once with the town's public-works crew. The work costs about $12,000, mostly for the gateway and two days of the integrator's time. The utility has not become secure against a determined state, but it is no longer the easiest target in its state, which is what the July attacks selected for.

Recommended Reading / Reference

Book: Andy Greenberg, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (2019), on the Russian unit that darkened Ukraine's grid and released NotPetya

Reference: CISA, EPA, and FBI, *Top Cyber Actions for Securing Water Systems* (February 2024), a two-page list of the basic steps for utilities

Module 8 Key Points

Critical Infrastructure and Operational Technology

• Plant systems put safety and uptime first, run for decades, and often use protocols with no authentication.

• Attacks on internet-exposed controllers hit more than 30 Minnesota water systems in July 2026; some forced boil-water notices.

• Volt Typhoon hid in American infrastructure networks to be ready to disrupt them in a crisis, not to steal data.

• A 2024 ransomware attack on a London pathology provider contributed to a patient's death, a hospital confirmed in 2025.

• The cheapest defenses: no controllers on the open internet, no default passwords, and staff who can run things by hand.

Module 8 Quiz

Check Your Understanding

1. What do operators of plants and utilities usually put first?

A) Safety and uptime
B) Keeping data confidential above all, as banks do
C) Installing every patch the day it appears
D) Moving all control systems to the public cloud

2. What did attackers exploit at American water utilities in July 2026?

A) Pipes
B) A flaw in the utilities' billing websites that let them change customers' bills
C) Employees' home routers, reached by phishing
D) Internet-exposed controllers

3. What was Volt Typhoon doing in American infrastructure networks, by the February 2024 advisory?

A) Stealing customers' card numbers for sale on criminal markets
B) Demanding ransoms from water utilities in several states
C) Hiding, ready to disrupt
D) Mining cryptocurrency on the utilities' servers

4. What happened in Ukraine on December 23, 2015?

A) A dam failed
B) Hackers cut power to about 225,000 customers
C) A pipeline exploded
D) Russian hackers wiped every computer in the country's banks and ministries at once

5. Which is the cheapest effective defense for a small utility's controllers?

A) Taking them off the open internet and replacing default passwords, with remote access only through a gateway that requires MFA
B) Buying cyber insurance
C) Hiring a full security team
D) Replacing every controller

Module 9 — States in Cyberspace

Espionage, Sabotage, Theft, and War

9a. What States Do With Cyber Operations

Governments use cyber operations for five main purposes. Espionage, the theft of secrets, is by far the most common, and every major power does it. Prepositioning plants access in another country's infrastructure for use in a future crisis, as Volt Typhoon did (Module 8). Sabotage damages or destroys: the Stuxnet worm, widely reported to have been built by America and Israel and never officially acknowledged, wrecked centrifuges at Iran's Natanz enrichment plant around 2010. Influence operations steal and leak documents or spread propaganda to sway opinion abroad.

The fifth purpose, theft of money, is pursued at scale by one state, North Korea (9e). Governments also buy commercial spyware, tools sold by private firms that break into phones, and some have turned it on journalists, lawyers, and dissidents. America placed NSO Group, the Israeli maker of the Pegasus spyware, on a trade blacklist in November 2021 for this reason.

States often work through others. China hires contractors, Russia tolerates and sometimes directs criminal gangs, and Iran presents its operations as the work of independent hacktivists. This blurring gives governments deniability, and it means that a 'criminal' or 'activist' attack may serve a state's ends.

9b. The Iran War's Cyber Front

The war that began with American and Israeli strikes on Iran on February 28, 2026 was fought in cyberspace too. By press reports, cyber operations accompanied the opening strikes, aimed at Iran's command and communications networks. A ceasefire on April 7 led to talks, but the truce collapsed in July, and fighting resumed from August 31. Iran-linked groups attacked throughout, with floods of web traffic, defacements, data theft, and wipers aimed at Israel, the Gulf states, and America.

The most damaging attack in America fell on Stryker, a large maker of medical devices. On March 11, 2026, attackers who had taken over its account in Intune, Microsoft's service for managing companies' devices, used it to send remote-wipe commands to the company's computers and phones. Stryker told the Securities and Exchange Commission (SEC) that its entire Microsoft environment had been affected. A group calling itself Handala claimed responsibility and said it had wiped more than 200,000 systems and taken 50 terabytes of data, claims no one has verified.

On March 19 the Justice Department announced that Handala was a fake activist persona operated by Iran's Ministry of Intelligence, and it seized four web domains the group used. After the April ceasefire, Handala declared that the cyber war would not end with any military ceasefire. The controller attacks on American water systems that summer (Module 8) fell within the same period, though no agency had formally attributed them by October.

9c. China: Typhoons and Contractors

China runs the largest state hacking effort, aimed above all at intelligence. Salt Typhoon broke into the networks of major telecommunications carriers, in America and abroad, reaching the systems carriers use to carry out court-ordered wiretaps and the call records of senior officials. On August 27, 2025, in an advisory joined by agencies of 13 countries, the FBI said the campaign had reached at least 200 American companies and targets in 80 countries.

Much of China's hacking is done by contractors. A leak of files from one of them, i-Soon, in February 2024 showed a market in which companies took hacking jobs from state agencies, priced their services, and complained about pay. American prosecutors charged employees of i-Soon in March 2025. Another Chinese group used a stolen Microsoft signing key to read the email of senior American officials (Module 4).

9d. Russia: From NotPetya to the War in Ukraine

Russia's military intelligence service, the GRU, carried out the most destructive cyberattack yet recorded. On June 27, 2017, its Sandworm unit hid malware, later named NotPetya, in an update to M.E.Doc, accounting software used by most companies in Ukraine. The malware looked like ransomware but could not be undone, a wiper in disguise, and it spread from Ukraine around the world in hours. It shut down the shipping line Maersk, the drug maker Merck, the delivery firm TNT, and the food company Mondelez, among many others. The White House put the damage above $10 billion.

When Russia invaded Ukraine on February 24, 2022, it struck first at Viasat's KA-SAT satellite network. That attack disabled tens of thousands of modems in Ukraine and across Europe, including the remote monitoring of about 5,800 wind turbines in Germany. Cyberattacks played a smaller part in the war than many had feared, partly because Ukraine's defenders, helped by Western governments and companies, were well prepared. Russia's foreign intelligence service, the SVR, behind SolarWinds (Module 7), went on spying; in January 2024 it read the email of Microsoft's senior leadership.

9e. North Korea: Theft for the State

North Korea steals money to finance a state under sanctions, including its weapons programs. Chainalysis estimated on December 18, 2025 that hackers linked to North Korea stole at least $2.02 billion in cryptocurrency in 2025, a record and 51% more than in 2024, out of about $3.4 billion stolen worldwide. Its total since tracking began reached about $6.75 billion. The largest single theft in history came on February 21, 2025, when its hackers took about $1.5 billion in Ether from Bybit, an exchange based in Dubai, after compromising a developer's computer at the firm that supplied Bybit's wallet software.

North Korea also places thousands of its people in remote technology jobs abroad under stolen or false identities, sending their wages home and sometimes stealing data or planting access. On June 30, 2025 the Justice Department announced searches of 29 'laptop farms' in 16 states, houses where American helpers kept company computers running for the workers. It seized 29 financial accounts, 21 websites, and about 200 computers. In one case, workers had used more than 80 stolen identities to get jobs at more than 100 American companies. The UN's panel of experts that tracked North Korea's sanctions-busting closed in April 2024, after a Russian veto ended its mandate.

9f. Norms, Attribution, and Consequences

States have agreed in principle on rules for cyberspace, but not on enforcing them. A UN group of governmental experts set out 11 voluntary norms in 2015, among them that states should not damage each other's critical infrastructure in peacetime or attack the teams that respond to incidents, and the General Assembly endorsed them. An open-ended working group of all UN members adopted its final report on July 11, 2025. The Assembly then created a permanent Global Mechanism, which held its first plenary session from July 20 to 24, 2026. Legal experts' views on how international law applies are collected in the Tallinn Manual 2.0 (2017).

When a state is caught, the usual answers are public attribution, often by several allied governments at once; indictments of named officers who will never stand trial; sanctions; and the seizure of servers and domains. Such steps raise costs but rarely change behavior. America's military answer since 2018 has been a doctrine of 'defend forward', under which US Cyber Command operates inside adversaries' networks to disrupt attacks before they land. Since March 20, 2026, Cyber Command and the NSA have been led by General Joshua Rudd, confirmed by the Senate on March 10 after nearly a year in which both had acting chiefs.

9g. America's Turn to Offense

The Trump administration released its cyber strategy, titled 'President Trump's Cyber Strategy for America', on March 6, 2026. The document is seven pages long and has six pillars: shaping adversaries' behavior, 'common sense' regulation, modernizing federal networks, securing critical infrastructure, keeping superiority in emerging technologies, and building talent. Its emphasis falls on imposing costs on attackers rather than on new rules for companies, a shift from the previous strategy (Module 4). Sean Cairncross, confirmed as National Cyber Director on August 2, 2025, leads it.

The same day the President signed Executive Order 14390, aimed at cybercrime and fraud. It ordered agencies to review the tools available against transnational criminal groups and to set up an operational cell in the National Coordination Center, and it told prosecutors to prioritize scam centers and sextortion. It also asked the Attorney General to recommend a program to compensate victims from seized and forfeited money, such as the Prince Group's bitcoin (Module 5).

A National Security Presidential Memorandum of August 12, 2026 went further. Titled 'Expanding Capabilities to Combat Transnational Cyber-Enabled Crime', it created a program under which vetted American companies may conduct offensive operations against foreign criminal groups. They must contract with the Justice Department or the Department of Homeland Security (DHS), post bonds of at least $1 million, and obtain written approval from both departments for each operation. Targets must be foreign criminal groups not part of a government; operations likely to kill, seriously injure, or amount to a use of force are barred. Its operating procedures were due within 60 days and had not been published by late October.

9h. Contested: Should Private Companies Be Allowed to Hack Back?

The case for. Supporters argue that the government cannot keep up: criminal gangs and scam networks operate from countries that will not act, police have too few staff, and private security firms often see attackers' infrastructure first and understand it best. Letting them disrupt that infrastructure, under rules, would raise criminals' costs quickly. They note that Microsoft's Digital Crimes Unit has used civil court orders

since 2010 to seize the domains that run botnets, networks of hijacked computers, and that such actions, like the Lumma takedown (Module 2), have worked.

The case against. Critics answer that attackers route their operations through other victims' computers, so a counterattack may hit a hospital or a small business rather than the criminal. Attribution can be wrong, and a private strike on a group that turns out to work for a state could escalate into a crisis between governments. Firms that sell security would profit from the operations they run, and allies may object to American companies attacking servers in their countries. Congress never brought the earlier proposals, the Active Cyber Defense Certainty Act of 2017 and 2019, to a vote, and the Computer Fraud and Abuse Act still bars hacking outside the program.

The evidence and the course's reading. Court-ordered disruptions by companies have a record of success with little collateral damage, while the memorandum's program has no record yet: its procedures were unpublished in October 2026, and no operation has been reported. The course's reading is that private disruption under judicial or federal authorization, with deconfliction and liability for mistakes, is defensible and is already practiced. Unilateral hacking back by companies acting on their own judgment is not. The program should be judged by its procedures and first operations.

Recommended Reading / Reference

Book: Kim Zetter, *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon* (2014), the fullest account of the sabotage of Iran's enrichment plant

Reference: Council on Foreign Relations, Cyber Operations Tracker, a database of publicly known state-sponsored cyber operations since 2005

Module 9 Key Points

States in Cyberspace

• States mostly spy, but also preposition, sabotage, influence, and steal; North Korea stole a record $2.02 billion in 2025.

• Iran's intelligence ministry runs Handala, which wiped Stryker's devices on March 11, 2026, the Justice Department said.

• NotPetya, a Russian wiper disguised as ransomware, did more than $10 billion of damage in 2017, by the White House's estimate.

• UN norms against attacking infrastructure in peacetime are voluntary; attribution, indictments, and sanctions are the usual answers.

• An August 2026 memorandum lets vetted American firms attack foreign criminal groups, each operation approved by two departments.

Module 9 Quiz

Check Your Understanding

1. Who operates the 'hacktivist' persona Handala, according to the Justice Department in March 2026?

A) A group of Israeli students protesting the war
B) A Russian criminal gang that rents out its malware
C) China's Ministry of State Security, working through contractors
D) Iran's intelligence ministry

2. How much cryptocurrency did hackers linked to North Korea steal in 2025, by Chainalysis's estimate?

A) $20 million
B) $202 million
C) At least $2.02 billion
D) About $20 billion, more than the country's whole economy

3. What did NotPetya, released in June 2017, turn out to be?

A) A hoax
B) A wiper posing as ransomware
C) A tool for stealing bank passwords, sold on criminal forums
D) A test of Ukrainian accounting software by its own developers that escaped by accident

4. What does the memorandum of August 12, 2026 allow?

A) Vetted private companies to conduct offensive cyber operations against foreign criminal groups, under federal approval and control
B) Any company to hack back freely
C) Attacks on foreign governments
D) Payments to ransomware gangs

5. What did Volt Typhoon and Salt Typhoon have in common?

A) Both were run by Russia's military intelligence service
B) Both were hoaxes
C) Both demanded ransoms from American telecom carriers and utilities
D) Both were attributed to China

Module 10 — AI on Both Sides

Faster Attacks, Faster Defense, and the Security of AI Itself

10a. How Attackers Use AI

Most criminal use of AI is ordinary. Large language models, the systems behind chatbots, write fluent phishing in any language, translate stolen documents, research targets, and help write code. The FBI's report for 2025 had a section on AI for the first time: it logged 22,364 complaints that involved AI, with losses of $893.3 million, many of them fraud using cloned voices or fake videos (Module 5). IBM's study of July 2026 found that one in four malicious breaches in its sample involved AI, up 56% in a year, and that they cost $6 million on average.

Malware has begun to call AI models while it runs. Google's Threat Intelligence Group reported on November 5, 2025 that APT28, a unit of Russia's military intelligence service, had used malware called PROMPTSTEAL against Ukraine. The malware asked an openly available AI model, through an online service, to write the commands it would run on each victim's computer. Google called it the first malware it had seen querying a language model in live operations. It also described PROMPTFLUX, an experimental program that asked Google's Gemini model to rewrite its own code to avoid detection, which had not been shown to infect anyone.

10b. Agents That Attack

The larger change is the AI agent, a model given tools and allowed to act on its own: to scan, run commands, read the results, and decide what to do next. Anthropic, which makes the Claude models, reported on November 13, 2025 that a group it judged to be sponsored by China's state, which it named GTG-1002, had used its Claude Code tool in mid-September against about 30 organizations, including technology companies, banks, chemical makers, and government agencies. The model carried out 80 to 90% of the tactical work, with humans approving the key steps, and a small number of the intrusions succeeded. The model also invented data at times, such as credentials that did not work.

A year later agents ran a campaign of their own. PaperCut, which makes software for managing printers, disclosed two flaws on August 27, 2026. From August 31 to September 2, an operator ran hundreds of agents, built on OpenAI's Codex tooling but running a model from DeepSeek, a Chinese AI company, against servers that had not been patched. In a report of September 9, GreyNoise, a firm that monitors internet traffic, counted at least 440 servers compromised at 395 organizations in 48 countries. The agents took under four hours to reach their first working attack and then hit 11 organizations in 26 seconds.

At one American high school the agents went from first access to control of the whole network in seven minutes. Even so, they reached that level of control at only 12 victims, and they partly ignored the operator's own list of 28 countries to avoid. Microsoft's report of October 2026, as summarized in coverage, judged that most campaigns it saw were still directed by humans. It also cited a test by Britain's AI Security Institute in which two frontier models, a Claude Mythos model from Anthropic and OpenAI's GPT-5.5, each completed a 32-step attack on a mock company network, with no defenders, in some of their runs. Mandiant found that the time between one gang's break-in and its handoff to a second group had fallen to 22 seconds in 2025, from more than eight hours in 2022.

10c. How Defenders Use AI

The same abilities find flaws before attackers do. In August 2025 the finalists of the AI Cyber Challenge, run by the Pentagon's Defense Advanced Research Projects Agency (DARPA), showed systems that found 54 of 70 flaws planted in real open-source software and patched 43, by DARPA's figures. They also found 18 flaws no one had known about, at an average cost of $152 a task; Team Atlanta won the $4 million first prize. Google says its Big Sleep agent, which runs in Chrome's build system every day, and its CodeMender tool, which writes fixes, helped Chrome fix 1,072 security bugs in two releases in June 2026, more than in the previous 23 releases together.

On April 7, 2026, Anthropic began Project Glasswing, giving Claude Mythos Preview, a model it judged too capable at finding and exploiting flaws for general release, to a group of defenders. The partners were 11 companies and organizations, among them Amazon Web Services, Apple, Cisco, CrowdStrike, Google, JPMorganChase, the Linux Foundation, and Microsoft, plus more than 40 maintainers of critical software. Anthropic committed up to $100 million in usage credits and $4 million to open-source security groups. It said the model had found thousands of previously unknown flaws, among them one 27 years old in the OpenBSD operating system. On June 2 it claimed more than 10,000 critical ones and added about 150 organizations in more than 15 countries.

OpenAI turned its flaw-finding agent, announced in October 2025 as Aardvark, into Codex Security, a research preview for business customers, on March 6, 2026, and claimed it had found about 800 critical flaws in testing. These figures are the companies' own and have not been independently counted. In security operations, AI assistants now sort alerts and draft investigations: Microsoft included its Security Copilot for its largest business customers from November 2025, with agents that triage phishing reports and alerts, and Google and CrowdStrike offer similar agents.

10d. Securing AI Systems

AI systems are targets as well as tools. Prompt injection hides instructions in content an AI system reads, such as a web page, an email, or a document, so that it follows the attacker instead of its user: an agent asked to summarize an inbox may find a message telling it to forward the user's files elsewhere. The Open Worldwide Application Security Project (OWASP), a nonprofit, ranks prompt injection first among the risks of applications built on language models. No complete fix exists, because the models cannot reliably tell instructions from data.

The damage an injected instruction can do depends on what the agent is allowed to do. Agents connect to email, files, code, and company systems through tool connectors, many using the Model Context Protocol (MCP), an open standard published by Anthropic in November 2024. A connector with broad permissions turns a tricked agent into an insider. The defenses are the old ones applied to new identities: give each agent its own account with the narrowest rights (Module 3), require a human's approval for consequential actions, and log what agents do.

The software used to build AI applications has flaws like any other, as JadePuffer's entry through Langflow showed (Module 6). Models themselves can be tampered with: model files shared on public hubs can carry malicious code, and training data can be poisoned to plant hidden behavior.

10e. Who Gets the Most Capable Models

AI companies now restrict their most capable tools for security work: besides Glasswing and Codex Security's limited preview (10c), Google limits a version of CodeMender to governments and trusted partners. The companies also watch for misuse of their public models and close the accounts involved, as Anthropic did in the GTG-1002 case.

Governments have entered the question of access too. Anthropic released two new models, Claude Fable 5 and Claude Mythos 5, on June 9, 2026. On June 12 it suspended access to both to comply with export controls imposed by the Commerce Department. The Department lifted the controls on June 30, and Anthropic restored access on July 1 (Anthropic's statement: https://www.anthropic.com/news/fable-mythos-access). Open models, whose weights anyone can download, sit outside such controls, as the DeepSeek model in the PaperCut campaign showed.

10f. Contested: Does AI Favor Attackers or Defenders?

The case that it favors attackers. Attackers need to find one weakness; defenders must close all of them. AI lowers the skill needed to attack, as Sysdig put it after JadePuffer, to whatever it costs to run an agent, and it lets one operator act against hundreds of targets at machine speed. The PaperCut campaign turned a disclosed flaw into hundreds of break-ins within days, faster than most organizations patch. Microsoft's report of October 2026, as reported, said the advantage for now lay with attackers.

The case that it favors defenders. Defenders have more data, more computing, and the cooperation of the AI companies, which can see and stop misuse of their models. AI that finds and fixes flaws before release, as in DARPA's challenge and Google's Chrome figures, removes the raw material of attacks at the source. Agents attacking live networks are noisy and make mistakes, and the PaperCut agents fully succeeded at only a small share of the organizations they reached (10b). Automated triage lets small teams handle alert volumes that once needed large ones.

The evidence and the course's reading. The attacks documented so far used old techniques against unpatched or misconfigured systems, at new speed: JadePuffer's flaw had long been fixed, and PaperCut's victims had not yet patched. The course's reading is that in the near term AI favors attackers against organizations that patch slowly and watch little, and that over time it favors defenders who use it to find and fix flaws first. The balance will be set less by the models than by how fast defenders apply what the models find.

Recommended Reading / Reference

Book: Bruce Schneier, *A Hacker's Mind: How the Powerful Bend Society's Rules, and How to Bend Them Back* (2023), which ends with what happens when AI systems start to find and exploit weaknesses

Reference: OWASP, *Top 10 for Large Language Model Applications* (2025), the most cited list of risks in applications built on language models, led by prompt injection

Module 10 Key Points

AI on Both Sides

• Most attackers use AI for ordinary work, such as fluent phishing; the FBI logged 22,364 AI-related complaints in 2025.

• AI agents exploited PaperCut print servers at 395 organizations in 48 countries within three days in 2026.

• Defenders' AI finds flaws at new scale: DARPA's finalists found 54 of 70 planted flaws; companies claim thousands more.

• Prompt injection hides instructions in what an AI reads; agents need narrow permissions and their own identities.

• AI companies restrict their most cyber-capable models; Anthropic gave Mythos Preview only to chosen defenders in April 2026.

Module 10 Quiz

Check Your Understanding

1. What did Anthropic report about the group it called GTG-1002 in November 2025?

A) It stole Anthropic's model

B) A Chinese state-sponsored group used Claude Code to carry out most of the tactical work of intrusions against about 30 organizations

C) It was a hoax

D) It attacked only Russian banks

2. What did the agents in the PaperCut campaign of 2026 run on, by GreyNoise's account?

A) A DeepSeek model

B) Only Anthropic's Claude models, through stolen accounts

C) Hand-written scripts with no AI at all, despite the reports

D) Google's Gemini, through its public chat app

3. What did the finalists' systems in DARPA's AI Cyber Challenge do in August 2025?

A) Ran ransomware

B) Wrote phishing emails

C) Found 54 of 70 planted flaws and patched 43

D) Broke into the Pentagon's networks during the contest, which was halted early

4. What is prompt injection?

A) A faster way to train models

B) A vaccine for computers

C) A typing shortcut

D) Hiding instructions in content an AI system reads, such as a web page or email, so that it follows the attacker instead of its user

5. Why did Anthropic limit Claude Mythos Preview to selected partners in April 2026?

A) Because it could not write code, and partners wanted to test it first

B) Its skill at finding flaws

C) Because regulators had banned its sale in every country

D) Because it was too slow for ordinary customers to use

Module 11 — Detection and Response

Seeing an Attack, and Handling the First Days

11a. Seeing an Attack

Detection starts with records. Computers, network devices, and cloud services keep logs: who signed in, from where, what programs ran, what data moved. Logs matter only if they are collected, kept long enough, and read. Many breaches are understood only months later, when investigators find that the logs that would have shown the start were never switched on or had already been deleted.

The main tool on each computer is endpoint detection and response (EDR) software. It records what programs do, flags behavior typical of attacks, such as a program dumping passwords from memory or encrypting thousands of files a minute, and lets defenders cut a machine off from the network with one click. Extended detection and response (XDR) adds email, identity, and cloud data to the same view. A security information and event management (SIEM) system collects logs from everywhere and correlates them, so that a sign-in from abroad followed by an administrator's password change raises one alert rather than two unnoticed ones.

Someone has to watch. A security operations center (SOC) is the team that triages alerts around the clock, and most companies cannot staff one: a 24-hour watch needs at least eight to ten trained people. Managed detection and response (MDR) providers sell the service instead, watching many clients' alerts from shared centers. Cheap tripwires help too: canaries, fake passwords or files that no legitimate user would ever touch, sound an alarm the moment an intruder tries them.

11b. How Long Attackers Go Unnoticed

Dwell time is the time between an attacker's arrival and its discovery. Mandiant's M-Trends report of March 23, 2026, drawn from its investigations of 2025, put the global median at 14 days, up from 11 a year earlier. Exploited flaws were the most common way in for the sixth year running, at 32% of its cases, followed by voice phishing (11%) and access left over from an earlier compromise (10%). The rankings differ from Verizon's (Module 2) because Mandiant sees the larger, more serious cases its clients call it in for.

Medians hide two different worlds. Ransomware announces itself, so its dwell times are short: the attacker wants to be found once the damage is done. Espionage is built to stay hidden, and state groups have stayed inside networks for years (Module 8). IBM's study measures something else again, the time to both identify and contain a breach, which averaged 247 days in its 2026 sample.

11c. The Incident-Response Cycle

NIST's guide to incident response, revised in April 2025 as Special Publication 800-61 Revision 3, ties response to the functions of its framework (Module 1). The work runs in a cycle: prepare; detect and analyze; contain, eradicate, and recover; and learn the lessons. Preparation decides most of the outcome. A plan names who leads, an incident commander, and who decides on shutdowns, payments, and public statements. It lists outside help on call: an incident-response firm on retainer, outside lawyers, the insurer's hotline, and the local FBI field office.

Communication must go out of band, through channels the attacker cannot read, because attackers inside a network often read the victim's email and chat to follow the response. Responders keep contact lists on paper and meet on personal phones or a separate service. Lawyers usually hire the investigators, so that their findings are protected by legal privilege, though courts have not always upheld that protection. Teams rehearse all of this in tabletop exercises, walking through a scenario around a table before a real one arrives.

Containment involves hard choices. Disconnecting systems stops the spread but alerts the attacker and halts the business; watching quietly gathers evidence but risks more damage. Evidence must be preserved before machines are wiped: copies of memory, disks, and logs. Eradication removes every foothold the attacker planted and resets every credential it might hold, including the special Windows account whose key signs network logins, which is reset twice so that forged tickets stop working. Leaving one backdoor means doing it all again.

11d. The Reporting Deadlines

An organization hit by an attack may owe notice to several regulators at once, each on its own clock. The deadlines start at different moments: when the incident is discovered, when it is judged significant, or, for America's securities regulator, when the company decides it is material to investors. The table sets out the main ones as of October 2026.

Regime	Who must report	Deadline	Status, October 2026
SEC, Form 8-K Item 1.05	Companies with publicly traded securities in America	Four business days after deciding an incident is material	In force since December 18, 2023; trade groups have asked the SEC to repeal it (Module 14)
Federal banking regulators	American banks	36 hours after deciding an incident is significant	In force since 2022
Health Insurance Portability and Accountability Act (HIPAA) breach rule	Health-care providers, insurers, and their contractors	To patients without unreasonable delay, at most 60 days after discovery	In force
CIRCIA	Covered critical-infrastructure entities, about 316,000 expected	72 hours for covered incidents; 24 hours after paying a ransom	Final rule under White House review since October 1, 2026; not yet in force
Defense Federal Acquisition Regulation Supplement (DFARS) 252.204-7012	Defense contractors and their suppliers	72 hours, to the Defense Department	In force
State breach-notification laws	Anyone holding residents' personal data, in all 50 states	Varies; often 30 to 60 days	In force
General Data Protection Regulation (GDPR)	Anyone handling Europeans' personal data	72 hours, to the data-protection regulator	In force since 2018
NIS2	Essential and important entities in the European Union	24-hour early warning, 72-hour notification, final report within a month	In national law in most member states

Regime	Who must report	Deadline	Status, October 2026
Digital Operational Resilience Act (DORA)	Banks, insurers, and other financial firms in the European Union	First notice within 4 hours of classifying an incident as major and no later than 24 hours after detection; then 72 hours; final report within a month	In force since January 17, 2025
Cyber Resilience Act	Makers of digital products sold in the European Union	24-hour early warning and 72-hour notification of exploited flaws and severe incidents; final report 14 days after a fix, or a month for an incident	In force since September 11, 2026
UK Cyber Security and Resilience Bill	Operators of essential services, managed service providers, and data centers	24-hour initial notice and full report within 72 hours	Before Parliament (Module 14)

11e. Working With the Police

Victims often hesitate to call the FBI, fearing that it will take over the response or make the attack public. In practice the FBI works alongside the victim's team. It can share what it knows about the attackers, check whether a payment would breach sanctions, and sometimes recover money, as it did after Colonial Pipeline paid (Module 2). From July 2022 to January 2023 the FBI secretly took over the servers of a ransomware group called Hive and gave decryption keys to more than 300 victims, sparing them about $130 million in ransoms, by the Justice Department's estimate.

Worked example — Harrow Precision, Part 2: The First Three Days

At 2:14 a.m. on the Saturday of the attack in Module 6, Harrow's EDR software, installed on its laptops and desktops but not yet on its servers, flags mass encryption. Its MDR provider phones the on-call IT manager at 2:31. By 2:50 she has disconnected the internet links at the plants and headquarters and shut down the VPN. At 3:30 she calls the insurer's breach hotline, which assigns an incident-response firm and an outside law firm from its panel by 6:00. The executives meet at 7:00 on a conference line set up from personal phones, assuming that company email is being read. They stop the production lines, which cannot be scheduled, and tell staff by text message and printed notices at the gates. On Sunday the responders copy the memory of key machines before anything is wiped, and they find the web shell on the VPN gateway, where the intrusion began, and the old service account. They disable it, reset every administrator's password, and reset the Windows sign-in key twice. Nearly all of Harrow's backups are gone, but one copy of the production-planning database, kept by its software vendor in a separate cloud account that cannot be altered for 30 days, is 26 hours old and intact. On Monday the firewall's logs show the upload, so the theft is confirmed. The chief executive decides not to pay while recovery is assessed, the lawyers check the gang against sanctions lists, and the FBI's Cincinnati field office is told. By the end of the third day Harrow has told its two carmaker customers and a defense contractor, as its contracts require, and told its employees that their Social Security numbers were taken. Module 12 picks up from here.

Recommended Reading / Reference

Book: Clifford Stoll, *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage* (1989), the first great detection story, in which a 75-cent accounting error led to a spy

Reference: NIST, Special Publication 800-61 Revision 3, *Incident Response Recommendations and Considerations for Cybersecurity Risk Management* (April 2025)

Module 11 Key Points

Detection and Response

• Endpoint detection, central log analysis, and a round-the-clock watch, often hired as a managed service, are the core of detection.

• Mandiant's global median dwell time rose to 14 days in 2025, up from 11.

• Isolate first, preserve evidence, and communicate out of band, since attackers may be reading the victim's email.

• Deadlines differ: four business days for SEC filers, 72 hours under GDPR and NIS2, 24 hours for EU product warnings.

• Calling the FBI early can bring keys, intelligence, and sanctions checks; Hive's seized keys spared more than 300 victims.

Module 11 Quiz

Check Your Understanding

1. What was the global median dwell time in Mandiant's M-Trends 2026 report?

A) 2 days
B) About 200 days, as in the early 2010s
C) 14 days
D) Under one hour, because attackers now move so fast

2. What does endpoint detection and response (EDR) software do?

A) Records activity on each computer and can isolate it
B) Encrypts email
C) Blocks all downloads
D) Scans the internet for exposed servers belonging to the company and reports them to its managers each week

3. How long does a company with publicly traded securities have to file a Form 8-K after deciding a cyber incident is material?

A) 24 hours
B) Four business days
C) 60 days, as for health data under HIPAA
D) One year, in the annual report rather than in a separate filing

4. Why should incident responders communicate out of band?

A) Because regulators require all calls to be recorded on paper
B) Because email is too slow for a crisis lasting several days
C) Because the insurer forbids using company systems during a claim
D) Attackers may be reading

5. What did the FBI's operation against the Hive ransomware group give victims between July 2022 and January 2023?

A) Cash refunds
B) Free insurance
C) Decryption keys taken from Hive's own servers, which spared more than 300 victims from paying ransoms
D) New computers

Module 12 — Recovery: Backups and Rebuilding

Getting Back to Work After the Worst Happens

12a. Backups That Survive an Attack

A backup is a copy of data kept so that it can be restored after a loss. The traditional rule is 3-2-1: three copies of important data, on two kinds of storage, with one copy off site. Ransomware has forced an extension, 3-2-1-1-0: one copy offline or immutable, meaning it cannot be changed or deleted for a set period even by an administrator, and zero errors when restores are tested. A backup that has never been restored is a hope, not a plan.

Attackers look for backups first, because a victim who can restore will not pay. They delete the snapshots that Windows keeps, wipe the backup servers, as at Harrow (Module 6), and use stolen administrator passwords to empty cloud backup accounts. Backups survive when they are walled off: kept under separate credentials protected by MFA, outside the main network's directory of users, in storage set to refuse deletion, with an offline copy where possible. Many organizations also prepare an isolated recovery environment, a clean room of fresh systems in which to rebuild without the attacker watching.

12b. Recovery Objectives

Two measures set the shape of a recovery plan. The recovery time objective (RTO) is how long a system can be down before the business suffers serious harm. The recovery point objective (RPO) is how much data, measured in time, the business can afford to lose: a nightly backup gives an RPO of up to a day. A business impact analysis sets both for each system, and the answers sort systems into tiers, so that the order of recovery is decided calmly in advance rather than argued over during the crisis.

Planners often define a minimum viable company: the few systems and people needed to keep the most important work going, such as taking orders, paying staff, and shipping products. Everything else waits. Shorter objectives cost more, so the targets are a business decision as much as a technical one.

12c. Identity First, and Rebuild Rather Than Restore

Most corporate networks run on a directory, usually Microsoft's Active Directory or its cloud version, Entra ID, which holds every user, password, and permission. An attacker who controlled it during the intrusion may have planted extra administrator accounts or forged credentials. Recovery therefore starts by rebuilding the directory from a source known to predate the intrusion and resetting every privileged credential (Module 11). Systems restored before that can be retaken through the old keys.

Restoring a whole server from a backup can restore the attacker's backdoor with it, if the backup was taken after the intrusion began. The safer path is to rebuild servers from clean standard images and then restore only the data, scanned before it is loaded. Rebuilding takes longer, so prepared organizations keep those images current and practice with them.

12d. Running on Paper

While computers are rebuilt, the business still has customers. Downtime procedures keep it going: printed order forms, offline copies of price lists and contact numbers, telephone trees, tills that can take payment offline, and handwritten shipping labels. Hospitals practice paper charting so that care continues when their records system is down. They work only if staff have practiced them.

12e. How Two Retailers Came Back

Two British retailers were hit in April 2025 by the same young English-speaking crews (Module 2), and their recoveries show the costs of each path. At Marks & Spencer (M&S), attackers who got in by deceiving staff at a third party deployed ransomware over the Easter weekend. M&S stopped taking online orders on April 25 and resumed them about seven weeks later, on June 10; click-and-collect returned in August. Customers' personal data, though not their payment details, were stolen.

The attack cost M&S £324 million of lost sales in the first half of its financial year, and its statutory profit before tax for those six months fell to £3.4 million, from £391.1 million a year earlier. It received £100 million from its insurers, and by November 2025 it said that practically all of its systems had been recovered. For the year to March 28, 2026, it booked £131.3 million of costs related to the incident.

The Co-op, attacked by the same campaign, spotted the intruders while they were still working and shut down parts of its own systems, which by reports kept ransomware from encrypting them. The price of shutting down was weeks of empty shelves in its food stores, and the attackers still took the personal data of all 6.5 million of its members. The Co-op put the hit to sales at £206 million in the first half of 2025 and to underlying earnings at £80 million.

Worked example — Harrow Precision, Part 3: Days 4 to 60

On day 4 the responders build a clean room in a new cloud account and rebuild Harrow's directory from scratch, issuing hardware security keys to the 40 people who will hold administrator rights. On day 6 the planning database is restored from the vendor's intact copy (Module 11); one day of orders is lost and re-entered from emails and phone notes. On day 8 the first plant restarts, running on schedules printed from the restored database and handwritten shipping labels. On day 9 the gang posts samples of the stolen files, and the company keeps its decision not to pay: paying would not undo the theft, and its recovery no longer depends on the attackers. Laptops and desktops are wiped and reimaged at about 120 a day by its managed service provider, and email and file shares return on day 15. Two plants reach 80% of normal output by day 21 and all three are at full production by day 35. By day 60 the VPN has been replaced by remote access that requires passkeys, backups sit in a separate account that refuses deletion, EDR covers every server, and internet-facing flaws on CISA's catalog must be fixed within 72 hours. The bill: about $6.1 million of lost gross profit, $1.4 million for investigators and lawyers, $1.3 million for new equipment, software, and overtime, $0.9 million of customer penalties and expedited freight, and $0.3 million to notify employees and pay for credit monitoring, about $10 million in all, more than twice the ransom demanded (Module 6).

Recommended Reading / Reference

Book: W. Curtis Preston, *Modern Data Protection: Ensuring Recoverability of All Modern Workloads* (2021), a practical guide to backup and recovery, ransomware included

Reference: CISA and partner agencies, *#StopRansomware Guide* (updated 2023), with checklists for preparing for and recovering from ransomware

Module 12 Key Points

Recovery: Backups and Rebuilding

• Attackers go for backups first; keep one copy offline or immutable, under separate credentials, and test restores.

• Recovery time and recovery point objectives set which systems come back first and how much data may be lost.

• Rebuild identity systems first and from clean sources, or the attacker walks back in.

• Marks & Spencer lost £324 million of sales in six months after its April 2025 attack and recovered £100 million from insurers.

• Practiced paper procedures keep orders moving and patients treated while computers are rebuilt.

Module 12 Quiz

Check Your Understanding

1. What does the extra '1' in the 3-2-1-1-0 backup rule add?

A) One copy kept offline or immutable
B) One more cloud
C) One backup a year
D) One person responsible for checking that every backup has been encrypted with a different key

2. What does a recovery point objective measure?

A) Cost
B) How many systems must be restored before the company can reopen its doors
C) How long a system can be down before the business suffers serious harm
D) How much data can be lost

3. Why should identity systems such as Active Directory be rebuilt first?

A) They are the cheapest
B) Law requires it
C) Because attackers who keep their keys can walk back into everything restored afterward
D) Because they hold the company's customer data, which must be notified to regulators within a single business day of discovery

4. What happened to Marks & Spencer's profit for the first half of its year after the April 2025 attack?

A) Statutory profit before tax fell to £3.4 million, from £391.1 million a year earlier
B) It doubled
C) It was unchanged
D) It fell by about 10%

5. What did the Co-op do in April 2025 that, by reports, kept ransomware from encrypting its systems?

A) It paid the ransom before any files were locked
B) Shut systems down early
C) It moved all its computers to a different cloud in one night
D) It ran every store on paper permanently

Module 13 — Cyber Insurance and the War Exclusion

Paying for a Risk That Arrives All at Once

13a. The Market

Cyber insurance is a young line of business. Munich Re, the world's largest reinsurer, estimated in 2026 that the global market had reached nearly $15 billion of premiums in 2025. Swiss Re, another reinsurer, put 2025's premiums at $15.6 billion and 2026's at $16.4 billion, with North America writing about two-thirds of the total. By its estimates, 60% to 70% of corporations with revenues above $1 billion buy cyber cover, but only 10% to 20% of small businesses do, and about 40% of large firms buy limits too low for the losses they face.

A typical policy pays for the victim's own losses and for claims against it. The first part covers the incident-response firm, lawyers, notifying and monitoring affected people, restoring data, and income lost while systems are down, usually after a waiting period of several hours; many policies also cover a ransom where paying is legal. The second part covers lawsuits, regulators' investigations, and fines where the law allows them to be insured. Policies set a limit, the most they will pay, and a retention, the share the insured pays first, and many cap particular kinds of loss at lower amounts.

Insurance shapes how victims respond. Most policies require the insured to use the insurer's panel of approved response firms and lawyers, and to call its hotline at once; the insurer's experts are often the first outsiders on the scene. Marks & Spencer's insurance payout (Module 12) shows what a large policy can pay. Jaguar Land Rover, by press reports, had not completed the purchase of a cyber policy when it was attacked (Module 7).

13b. What Insurers Demand

After ransomware losses soared in 2020 and 2021, insurers raised prices and began refusing cover to companies without basic controls. Applicants are now asked whether they use MFA, especially for email and remote access; whether they run endpoint detection; whether their backups are offline or immutable and tested; how fast they patch internet-facing devices; and whether administrators' access is tightly controlled. Many insurers also scan applicants' internet-facing systems themselves before quoting.

Brokers credit the demands with spreading these controls among mid-size firms, many of which need cover to win contracts from large customers. Prices then fell as losses steadied and competition returned. By the broker Marsh's index, global cyber rates fell 4% in the third quarter of 2026, their thirteenth quarterly decline in a row. Even the large attacks of 2025 did not reverse the trend, partly because some of the biggest victims were uninsured or underinsured.

13c. Losses That Arrive Together

Insurance works by pooling risks that are independent: a fire at one house rarely burns down houses in another city. Cyber risk is not independent. One flaw in a widely used product, such as the file-transfer software behind Clop's data thefts (Module 2), or one outage at a major supplier, such as CrowdStrike's faulty update (Module 7), can produce claims from thousands of policyholders at once. CyberCube, a firm

that models cyber risk for insurers, estimated insured losses from the CrowdStrike outage at up to $1.5 billion.

Insurers manage this aggregation risk by modeling scenarios such as the failure of a major cloud, limiting how much they write in any one sector, and passing part of the risk to reinsurers. Since 2023 some have also issued catastrophe bonds for cyber risk, which transfer losses above a threshold to investors in capital markets. The worst scenarios remain the ones that no private market wants to hold, and war is the largest of them.

13d. NotPetya and Merck's Claim

The question of war came to court after NotPetya (Module 9). Merck, the drug maker, claimed about $1.4 billion for its losses under its property insurance, which covered all risks except those excluded. Its insurers refused, citing the policies' exclusion of 'hostile or warlike action' by a government. A New Jersey court ruled for Merck in January 2022, and an appeals court agreed in May 2023, finding that the exclusion's traditional wording referred to military action, not a cyberattack on a civilian company. Merck and its insurers settled on confidential terms in January 2024, before the state's highest court could rule.

The case showed that old exclusions written for physical war did not reliably exclude a state's cyberattack. The food company Mondelez, which had claimed about $100 million for its own NotPetya losses, settled with its insurer, Zurich, in 2022. Insurers concluded that they needed exclusions written for cyber.

13e. The War Exclusions of 2023, and the Iran War

Lloyd's of London, the largest market for specialty insurance, told its members in a bulletin of August 16, 2022 that from March 31, 2023 every standalone cyber policy must exclude losses from war and from state-backed cyberattacks that significantly impair a country's ability to function or its security. Model clauses drafted by the market's trade association offer several versions, from narrow to broad. They let the parties rely mainly on attribution by the government of the country where the affected computers sit, and some versions keep cover for companies hit as bystanders outside the targeted country.

The Iran war of 2026 brought the exclusions their first test in a hot war. When Iran's intelligence ministry, operating as Handala, wiped Stryker's devices (Module 9), brokers and insurers had to ask whether such a loss would fall under the state-backed exclusion. The answer turns on details: whether a government has attributed the attack, which the Justice Department had done (Module 9), and whether an attack on one company meets the threshold of major harm to a state's functioning, which many clauses require. Stryker's own case tests little: its annual report of February 2026 said it was self-insured for cybersecurity losses, managing part of them through a captive insurer it owns. No insurer's decision or court ruling under the exclusions had been reported by October 2026; the first will show companies what their policies cover when a state attacks them.

Worked example — Harrow Precision, Part 4: The Claim

Harrow's policy, bought a year before the attack, has a $5 million limit for all claims combined, a $250,000 retention, and a waiting period of 12 hours before lost income is covered. Its covered losses from Module 12's bill come to $8.4 million: the investigators, lawyers, employee notices, and credit monitoring in full; $0.7 million of the software and overtime needed to restore systems; and $6.0 million of lost gross profit, the rest falling within the waiting period. Two items fall outside the cover: about $0.6 million for new equipment and upgrades, which the policy treats as improvements rather than repairs, and the customer penalties and

expedited freight. After the retention, the covered total is $8.15 million, but the limit caps the payment at $5 million. Harrow recovers about half of its cost. The limit had been chosen on the assumption that an attack would stop the plants for about two weeks; full production returned only after five. The lesson its board takes away is to buy limits sized to a realistic worst case, measured by the time it would take to rebuild, rather than to the premium budget.

13f. Contested: Who Should Carry Catastrophic Cyber Risk?

The case for leaving it to the market. Supporters of a private solution argue that insurers price risk and, through their demands, push companies to defend themselves, which a government backstop would weaken. Capacity is growing through reinsurance and catastrophe bonds. Exclusions for war keep insurers solvent so that they can pay ordinary claims, and large firms can hold their own reserves for the extreme cases. A public backstop would invite companies to underinvest in security, expecting to be rescued.

The case for a public backstop. Others answer that the exclusions leave the largest scenario, a state's attack on a whole economy, uninsured by design, so the losses fall on companies, and through rescues on taxpayers anyway. Britain's loan guarantee for Jaguar Land Rover (Module 7) was such a rescue, improvised after the event. Terrorism offers a model: America's Terrorism Risk Insurance Act of 2002 shares the largest losses between insurers and the government, and Britain's terrorism reinsurer, Pool Re, extended its cover in 2019 to physical damage caused by terrorists using cyber means. America's Treasury has studied a federal response since 2022, and in March 2026, for a report on its terrorism insurance program, it asked for comments on cyber losses; it has proposed no program.

The evidence and the course's reading. Most small and mid-size firms remain uninsured, the exclusions are untested in court, and governments have already stepped in after large events without a plan. The course's reading is that the market should carry ordinary losses, while a public backstop should cover catastrophic, state-scale events that the exclusions remove from the market. The backstop should come with minimum security standards and a first layer of loss that insurers keep, so that it does not reward neglect.

Recommended Reading / Reference

Book: Josephine Wolff, *Cyberinsurance Policy: Rethinking Risk in an Age of Ransomware, Computer Fraud, Data Breaches, and Cyberattacks* (2022), on how cyber insurance grew and what it can and cannot do

Reference: Lloyd's Market Bulletin Y5381, *State backed cyber-attack exclusions* (August 16, 2022), the rule that set the market's war exclusions

Module 13 Key Points

Cyber Insurance and the War Exclusion

• Cyber insurance premiums were about $15 billion to $16 billion worldwide in 2025, about two-thirds of them in North America.

• Insurers now require MFA, endpoint detection, tested offline backups, and fast patching of edge devices before they offer cover.

• One flaw or outage can hit thousands of policyholders at once, which makes cyber risk hard to diversify.

• Lloyd's has required exclusions for state-backed cyberattacks since March 2023; none had been tested in court by October 2026.

• A limit sized for a short outage can leave half of a long outage's losses uninsured.

Module 13 Quiz

Check Your Understanding

1. Roughly how large was the global cyber insurance market in 2025, by Munich Re's estimate?

A) About $1.5 billion, a small specialty line
B) About $150 billion, larger than property insurance
C) About $1.5 trillion, the size of all insurance worldwide
D) Nearly $15 billion

2. What did New Jersey's courts decide in Merck's claim for its NotPetya losses?

A) For the insurers
B) The war exclusion did not apply
C) That Merck had caused the attack itself through careless software updates
D) That cyberattacks cannot be insured under any policy written in America

3. What has Lloyd's required of standalone cyber policies since March 31, 2023?

A) Lower premiums
B) Coverage of all wars
C) An exclusion for losses from war and from state-backed cyberattacks that badly impair a country's functioning or security
D) Payment of all ransoms

4. Why are cyber losses harder to insure than house fires?

A) One flaw or outage can hit thousands of policyholders at once
B) Computers rarely fail
C) Hackers are always caught
D) Because no company has ever lost money in a cyberattack that it could not recover within a week from its own backups

5. Why did Harrow's insurance pay only about half of its losses in the worked example?

A) It paid the ransom
B) The war exclusion
C) Its policy limit was set for a shorter outage
D) Because the insurer refused every claim made by manufacturers after the attacks on carmakers in 2025

Module 14 — Governance, Disclosure, and Regulation

Who Answers for Cyber Risk, and What the Law Requires

14a. The Board's Job

Cyber risk is a business risk, and the board of directors answers for it like any other. The Govern function of NIST's framework (Module 1) describes the job: set how much risk the organization will accept, assign responsibility, fund the program, oversee suppliers, and check that the plans work. Good boards ask a few plain questions. What are our most important systems and data? Who would attack us, and how? How quickly would we know? How long would it take to recover, and what would it cost? How much of that would insurance pay?

The law increasingly asks boards to show that they asked. In Delaware, where most large American companies are incorporated, directors have a duty to make a good-faith effort to monitor major risks, though few lawsuits over cyber failures have succeeded. The SEC's rule of July 2023 requires companies with listed securities to describe in each annual report how they assess and manage cyber risk and how their boards oversee it, besides the filing due after a material incident (Module 11). Europe's NIS2 goes further, making managers approve and oversee security measures, take training, and answer personally for failures.

14b. America's Patchwork

America has no general cybersecurity law. Rules come sector by sector: banking regulators for banks, the health department's security rule for health care, mandatory standards for the bulk power grid (Module 8), and the Transportation Security Administration's directives for pipelines, issued after the Colonial attack (Module 2). States add their own. New York's financial regulator requires the firms it supervises to run security programs, and since November 1, 2025 it has required MFA for anyone accessing their information systems, unless the firm's security chief approves an equally strong alternative in writing, and an inventory of their technology.

The patchwork overlaps. A report to Congress by DHS in 2023 counted 45 federal requirements to report cyber incidents, run by 22 agencies, often with different definitions and deadlines. Congress tried to tidy them with the Cyber Incident Reporting for Critical Infrastructure Act (CIRCIA) of 2022, which told CISA to write a single rule for critical infrastructure (Module 11). CISA missed its targets of May and September for the final rule, which is now under White House review (Module 11); reports expected it to be published before the end of 2026.

The SEC's rule is under pressure. Paul Atkins, its chairman since April 2025, invited public comment on January 13, 2026 on reforming Regulation S-K, the rules for what companies must disclose, which include the annual cyber item. On April 10, five financial trade groups, among them the American Bankers Association, urged the SEC to rescind both the annual item and the incident filing, and CTIA, the wireless industry's trade group, made the same case on October 2. The SEC had published no formal proposal to change them by October 2, when CTIA wrote.

A law that lets companies share threat information with each other and the government without legal risk, the Cybersecurity Information Sharing Act of 2015, lapsed on October 1, 2025 and has since been extended

in short stretches; the stopgap budget law of September 2026 carries it to December 11, 2026. The House voted on July 22, 2026, by 216 to 212, for a defense bill that would extend it to 2036, but the Senate's version had no such provision when the House voted. Senator Rand Paul has resisted a long extension unless CISA is barred from work on disinformation.

Voluntary programs have moved slowly. The Federal Communications Commission (FCC) launched the Cyber Trust Mark, a label for connected devices that meet basic security standards, in January 2025. Its first lead administrator, UL Solutions, withdrew in December 2025 after an FCC investigation into its ties to China, and the FCC chose the ioXt Alliance, a nonprofit, to replace it on April 13, 2026.

14c. The National Cyber Agency's Capacity

CISA, created in 2018 within DHS, advises companies and protects federal civilian networks; it has no power to regulate most companies. It has had no Senate-confirmed director since January 2025. Three officials have led it in an acting capacity: Bridget Bean, then Madhu Gottumukkala from May 2025, and Nick Andersen, the head of its cybersecurity division, since February 26, 2026. The President's nominee, Sean Plankey, withdrew in April 2026 after 13 months without a vote of the full Senate, and no new nominee had been named by late October.

The agency shrank from about 3,400 staff before 2025 to about 2,200 by June 2026, by the count of Markwayne Mullin, the Homeland Security Secretary since March 24, 2026. He told Congress that month that CISA would ideally have about 2,800 and that rebuilding would take about a year once a new director is confirmed; about 250 hires were awaiting security clearances in September. The budget request for fiscal 2027, released on April 3, 2026, would cut CISA's funding by $707 million and eliminate about 867 positions, while the whole government ran on a stopgap. Fewer staff means fewer free scans, assessments, and advisories for the small utilities, hospitals, and towns that rely on them most (Module 8).

14d. Europe: NIS2, the Cyber Resilience Act, and DORA

The European Union regulates far more directly. Its NIS2 directive, due in national law by October 17, 2024, covers 18 sectors, divides firms into essential and important entities, requires security measures and incident reporting (Module 11), and sets fines of up to €10 million or 2% of worldwide turnover for essential entities. Countries were slow. Germany's law took effect on December 6, 2025, raising the number of firms its cyber agency supervises from about 4,500 to about 29,500. On July 8, 2026 the European Commission referred Ireland, Spain, France, and the Netherlands to the EU's Court of Justice for failing to transpose the directive. By a tracker's count in late October, national laws were in force in 24 of the 27 member states, all but France, Ireland, and Spain.

The Cyber Resilience Act regulates products rather than firms: hardware and software sold in the EU with digital elements, from routers to apps. It entered into force on December 10, 2024. Its reporting duties (Module 11) run through a Single Reporting Platform managed by ENISA, which went live the day they took effect. From December 11, 2027 makers must design products securely, handle and fix vulnerabilities for a support period, and document their components in an SBOM (Module 7). Fines reach €15 million or 2.5% of worldwide turnover.

The Digital Operational Resilience Act (DORA, Module 11) covers banks, insurers, and other financial firms, and it lets European regulators oversee their most important technology suppliers directly. On November 18, 2025 the EU's financial supervisors designated 19 such critical providers, among them Amazon Web Services, Google Cloud, Microsoft, Oracle, SAP, and Deutsche Telekom. The Commission proposed a

revised Cybersecurity Act on January 20, 2026, which would widen ENISA's role and create a way to restrict high-risk suppliers from countries of concern.

14e. Britain: The Cyber Security and Resilience Bill

Britain's Cyber Security and Resilience (Network and Information Systems) Bill, introduced on November 12, 2025, would update its 2018 rules. It would bring managed service providers other than the smallest, data centers of 1 megawatt or more, and suppliers designated as critical into the regime, and it would set fines of up to £17 million or 4% of worldwide turnover. The bill cleared the House of Commons in June 2026. In the House of Lords, its report stage was scheduled for October 26 and its third reading for November 2; it had not received Royal Assent by the end of October. The ban on ransom payments by public bodies that Britain announced (Module 6) is not in the bill and would need one of its own.

Worked example — Harrow Precision, Part 5: What It Had to Tell, and to Whom

Harrow is privately owned, so it files nothing with the SEC; had its shares been listed, it would have had to file a Form 8-K soon after deciding that the attack was material (Module 11), which it would probably have done within a day or two. Its other duties came from four directions. Its contracts with its carmaker customers required notice of any security incident within 72 hours, and it met them. One plant makes parts for a defense contractor, so the DFARS clause in its subcontract required it to report to the Defense Department on the same clock, which it did through the department's reporting portal. The theft of current and former employees' Social Security numbers (Module 6) triggered state breach-notification laws: its lawyers mailed letters with offers of credit monitoring within the 45 days that Ohio and Indiana allow, and notified the attorneys general where state law required it. Finally, once CIRCIA's rule takes effect, a critical-manufacturing firm of Harrow's size would probably have to report such an attack, and any ransom payment, to CISA on the rule's deadlines (Module 11). One incident, four sets of rules, each with its own definitions: the company's lawyers kept a single timeline to show which clock had started when.

14f. Contested: Mandates or Market?

The case for mandates. Advocates argue that insecurity is a classic market failure: a breach harms customers, suppliers, and the public, who do not set the company's security budget, and buyers cannot judge security before they buy. They point out that some sectors changed only when required to, such as pipelines and New York's financial firms (14b). Reporting rules, they add, give defenders the data they need, and Europe's approach shows that broad duties can be written.

The case for markets. Opponents answer that regulation produces checklists, and compliance is not security. Duties fall hardest on small firms with the fewest resources, and regulators lag years behind attackers' methods. The overlapping American reporting rules already waste effort in the first hours of an incident, when every hour counts. Insurers, large customers' contracts, and the threat of lawsuits already push companies toward the controls that matter, as Module 13 describes. The administration's strategy (Module 9) chose lighter, 'common sense' regulation on these grounds.

The evidence and the course's reading. The evidence is thinner than either side claims. Studies of American breach-notification laws found that they reduced identity theft modestly, by about 6% in one widely cited study, while Europe's newest laws are too recent to judge. The sectors with the weakest rules, such as water, are also among the weakest in practice (Module 8). The course's reading is that targeted mandates

work best: binding minimums for operators of critical services and for software makers' worst practices (Module 4), plus a single harmonized reporting regime. Broad checklists for everyone else do not.

Recommended Reading / Reference

Book: Jeff Kosseff, *Cybersecurity Law* (third edition, 2022), a guide to American cybersecurity law across sectors and states

Reference: Directive (EU) 2022/2555 (NIS2), on measures for a high common level of cybersecurity across the Union, on EUR-Lex

Module 14 Key Points

Governance, Disclosure, and Regulation

• Boards own cyber risk: NIS2 makes managers liable, and the SEC requires listed companies to describe their oversight every year.

• America regulates cybersecurity sector by sector; DHS counted 45 federal incident-reporting requirements in 2023.

• Financial and wireless-industry trade groups have urged the SEC to rescind its cyber disclosure rules; no change had been proposed by October 2026.

• CISA has had no confirmed director since January 2025 and shrank from about 3,400 staff to about 2,200 by June 2026.

• Europe's NIS2, Cyber Resilience Act, and DORA impose duties with fines; Britain's bill had not become law by October 2026.

Module 14 Quiz

Check Your Understanding

1. Who led CISA in October 2026?

A) An acting director
B) A judge
C) A Senate-confirmed director appointed in 2025 who left after six months
D) The Secretary of Defense, after CISA moved into the Pentagon in 2026

2. What does the EU's Cyber Resilience Act require from September 11, 2026?

A) Free software for all
B) That makers of digital products report actively exploited flaws and severe incidents, with an early warning within 24 hours
C) A ban on encryption
D) Licenses for programmers

3. What does the SEC's 2023 rule require of listed companies besides the four-day filing?

A) A public list of every vulnerability in their software, updated monthly
B) A ban on paying ransoms to any group, sanctioned or not
C) Insurance against all cyber losses, at a minimum set by the SEC
D) Yearly reports on cyber risk and board oversight

4. Which country's NIS2 law took effect on December 6, 2025, multiplying the number of supervised firms?

A) Germany
B) France, whose bill passed in 2024
C) Ireland, home to many tech firms' European offices
D) Spain, after its referral to the Court of Justice

5. How many federal cyber-incident reporting requirements did a DHS report to Congress count in 2023?

A) Three
B) Forty-five
C) About 4,500, one for each regulated company
D) Exactly one, CIRCIA, which replaced all the others in 2022

Module 15 — Privacy and Data Brokers

What Is Already Out There, and What Can Be Done

15a. Breaches: Where Exposed Data Comes From

Most of the personal data that criminals hold was stolen in breaches, much of it from companies the victims never dealt with directly. Earlier modules followed several, from the Snowflake data thefts (Module 3) to the health records taken from Change Healthcare (Module 6). Each breach adds to a pool of records that are bought, sold, combined, and reused for years.

One of 2026's largest exposures fell on a company most of its victims had never heard of. IDScan.net, a New Orleans firm whose software scans driver's licenses for bars, shops, and car dealers to check age and identity, said in a notice first posted on September 4, and updated on September 29, that an intruder had been in its cloud storage between April 4 and September 2. A new identity-theft service called Nexus was offering the data on a Russian-language criminal forum by August 31. It claimed more than 153 million scans of American and Canadian driver's licenses, along with 10 million other ID cards, 3 million travel documents, and 579,000 medical cards; the figures are the seller's, and the FBI's New Orleans office was investigating.

Identity data of this kind cannot be changed. A password can be reset in a minute, but a Social Security number, a date of birth, and a face on a license last a lifetime. In 2024 a background-check company, National Public Data, lost a database that researchers found held about 272 million unique Social Security numbers, and much of it was posted free on a criminal forum that August; the company filed for bankruptcy in October. The practical conclusion is to assume that identity data is already public, and to make it useless to a thief rather than try to keep it secret.

15b. The Data-Broker Business

Data brokers collect and sell information about people with whom they have no direct relationship. They draw on public records such as property deeds and court filings; on purchases, loyalty cards, and warranty forms; on apps that share location; and on one another. Their customers include marketers, lenders, landlords, insurers, and employers screening applicants, and, through people-search websites, anyone with a few dollars. California's registry lists more than 600 brokers, and Vermont, Texas, and Oregon also make them register.

Brokers feed attacks in three ways. Scammers buy their profiles, with relatives' names and old addresses, to make a call sound as if it comes from a bank (Module 5). Brokers' own databases are breached, as National Public Data's was. And location data can reveal visits to clinics, places of worship, and homes: in 2024 and January 2025 the FTC barred several brokers, among them X-Mode and Gravy Analytics, from selling data that shows visits to sensitive places.

Washington also treats bulk personal data as a national-security risk. The Protecting Americans' Data from Foreign Adversaries Act of 2024 bars brokers from selling Americans' sensitive data to China, Russia, Iran, or North Korea, or to companies they control, with fines of up to $53,088 per violation; the FTC, which enforces it, sent warning letters to 13 brokers on February 9, 2026. A Justice Department rule in force since April 8, 2025 goes further. It restricts any American company from selling or transferring bulk genetic,

biometric, health, financial, and location data to six countries of concern. Its last duties took effect on October 6, 2025, and no enforcement action had been announced by late October 2026.

15c. State Privacy Laws and California's One Request

America has no general federal privacy law; Congress has debated one for years without passing it. The states have filled the gap. By one tracker's count in October 2026, 24 states had passed comprehensive privacy laws and 20 of them were in force, a count that includes Florida's narrower law. Most give residents the right to see the data a company holds about them, correct it, delete it, and opt out of its sale and of targeted advertising. Several also require companies to honor Global Privacy Control, a browser setting that sends an opt-out to every site visited. The rights are real but tiring to use, because each request goes to one company at a time, and the brokers a person has never heard of are the hardest to find.

California's Delete Act of 2023 tackles that problem. Since January 1, 2026, residents have been able to file one request through the state's Delete Request and Opt-out Platform (DROP). Since August 1 every registered broker has had to check the platform at least every 45 days, delete the matching records, and keep them deleted, or pay $200 per request for each day it fails. More than 345,000 requests had been filed by early August. The state's privacy agency, which calls itself CalPrivacy, has also fined brokers that failed to register: one, a Texas firm called Datamasters, which had sold lists of people with health conditions such as Alzheimer's disease, was fined $45,000 in December 2025 and was ordered to stop selling Californians' data.

15d. Steps That Work

A household cannot recall stolen data, but a few free steps make most of it useless to a thief:

- **Freeze credit** at Equifax, Experian, and TransUnion. Freezes have been free by federal law since September 21, 2018, and a bureau must place one within a business day of an online or phone request and lift it within an hour. A freeze blocks new credit in a person's name without affecting existing accounts or credit scores, and parents can freeze a child's file until age 16.
- **Get an Identity Protection PIN** from the Internal Revenue Service (IRS), a six-digit number issued anew each year, without which no one can file a federal tax return in that person's name. Since 2021 anyone who can verify their identity online can get one.
- **Lock the phone number** with a port-out PIN or number lock from the mobile carrier, the defense against the SIM swaps of Module 3.
- **Watch for misuse** through transaction alerts on bank and card accounts and free weekly credit reports from AnnualCreditReport.com. After a serious breach, a fraud alert, free and renewable each year, asks lenders to verify identity before opening credit.
- **Shrink the profile.** Californians can use DROP; others must opt out of people-search sites one at a time or pay a removal service, whose results vary.
- **Report theft quickly** at IdentityTheft.gov, which produces a recovery plan and an identity theft report that lets victims have fraudulent accounts blocked from their credit reports.

Worked example — The Harper Household, Part 3: Making the Data Useless

In September 2026 a letter tells Dana and Marcus that scans of their driver's licenses, taken when a supermarket checked their IDs electronically, were stolen from the scanning company's cloud storage. After the stolen sessions of Module 3 and the call to Walt in Module 5, the family assumes its identity data is public and spends a Saturday morning making it useless. Dana, Marcus, and Walt each freeze their credit at all three bureaus online, about 20 minutes apiece, with Dana helping her father. Leo's freeze takes a letter to each bureau, with copies of his birth certificate and Social Security card and proof that Dana is his parent. The three adults get IRS identity protection PINs and each file one DROP request. Marcus asks the carrier to lock all four phone numbers against transfer, and Dana turns on alerts for every card. The steps cost nothing and take about four hours. In October a bank writes to Marcus that it could not open the credit card he applied for, because his credit file is frozen; he had applied for none. The thief had his name, address, date of birth, and license number, and still got nothing.

15e. Contested: Should Governments Be Able to Require Access to Encrypted Data?

End-to-end encryption, in which only the sender and the recipient hold the keys, now protects most messaging, from WhatsApp and Signal to Apple's iMessage, and increasingly the backups that phones keep in the cloud. Police and intelligence agencies say it keeps them from reading messages even with a warrant, a problem they call 'going dark'. Two fights tested how far governments can go.

Britain used its Investigatory Powers Act to serve Apple with a secret order to provide access to encrypted iCloud data. Apple responded on February 21, 2025 by withdrawing Advanced Data Protection, its strongest form of iCloud encryption, from users in Britain. Tulsi Gabbard, then America's director of national intelligence, said on August 19, 2025 that Britain had dropped the demand as it affected Americans, but a new order limited to British users was later reported. The Investigatory Powers Tribunal heard arguments on September 17, 2026 about how much of the case can be made public; by late October it had not ruled, and the feature had not returned in Britain.

In the European Union, a proposed law against child sexual abuse material, which critics call 'chat control', would have let authorities order messaging services to scan users' messages, including encrypted ones. EU governments agreed a position on November 26, 2025 that dropped mandatory scanning orders, and negotiations with the European Parliament have run since December 9, 2025. A temporary rule allowing services to scan voluntarily lapsed on April 3, 2026 and was later restored until April 3, 2028.

The case for access. Supporters, among them many police chiefs and child-protection groups, argue that a warrant should reach every kind of evidence, as it always has for letters, phone lines, and safes. Crimes against children, terrorism, and organized fraud are planned in encrypted apps, and services that scan unencrypted content already report millions of abuse images a year. Engineers who can build global payment systems, they say, can build access with safeguards, such as keys held in escrow or scanning on the device before a message is encrypted.

The case against. Most cryptographers answer that there is no way to build a door that only the right people can open. Any access mechanism becomes a target, and the wiretap systems that Salt Typhoon reached (Module 9) show what happens to one; after that breach, in December 2024, officials of CISA and the FBI urged Americans to use encrypted messaging. Weakened encryption would expose everyone to criminals and foreign states, while serious criminals would move to tools beyond any one country's reach. Apple's withdrawal shows the likelier result of a demand: ordinary users simply lose protection.

The evidence and the course's reading. No country has shown an access system that works at scale without creating new risk. Police have other tools, among them device searches with warrants, records of who contacted whom, informants, and lawful hacking of a particular suspect's phone, which are slower and costlier but targeted. The course's reading is that governments should not require built-in access to encrypted services, because the loss of security falls on everyone while the gain for investigators can be partly met in narrower ways; scanning on the device moves the same trade to a different place.

Recommended Reading / Reference

Book: Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World* (2015), on how companies and governments gather personal data and what that costs

Reference: California Privacy Protection Agency, *Delete Request and Opt-out Platform (DROP)*, the state's one-request deletion service for its residents

Module 15 Key Points

Privacy and Data Brokers

• Identity data such as Social Security and license numbers cannot be changed once stolen, so the defense is to make it useless.

• Data brokers sell profiles of people they have never dealt with; California's registry lists more than 600 of them.

• California's DROP lets residents send one deletion request that every registered broker must process at least every 45 days.

• Credit freezes are free, take minutes online, and block new credit; parents can freeze a child's file until 16.

• Britain's demand for access led Apple to withdraw its strongest iCloud encryption for British users in February 2025.

Module 15 Quiz

Check Your Understanding

1. What does California's Delete Request and Opt-out Platform (DROP) let residents do?

A) Freeze their credit at all three bureaus
B) Sue data brokers in small-claims court
C) Send one deletion request that every registered data broker must check at least every 45 days
D) Block marketing calls

2. How does a credit freeze protect someone whose Social Security number was stolen?

A) It deletes the stolen data
B) It issues a new number
C) It reports the thief to the FBI automatically and repays any money taken from existing accounts
D) Lenders cannot see the frozen report, so new credit fails

3. What did Apple do on February 21, 2025, after Britain ordered access to encrypted iCloud data?

A) Complied in secret
B) Left the British market
C) Withdrew its strongest iCloud encryption from users in Britain
D) Built a way in for every government that asked, then published the design for anyone to inspect

4. Since when have credit freezes been free by federal law?

A) September 2018
B) Since the first credit bureaus opened
C) Since 2003, when Congress first allowed fraud alerts
D) Only since the breaches of driver's license scans in 2026

5. What does the Justice Department's bulk data rule of April 2025 restrict?

A) All online ads
B) Bulk transfers of sensitive personal data to six countries of concern
C) Every sale of data by American brokers to American marketers, which it banned outright from 2025
D) Any American's use of a foreign email or cloud service, even for personal messages to relatives abroad

Module 16 — Where Things Stand, and What to Watch

The Argument, the Contradictions, and the Calendar

Recommended Reading / Reference

Reference: The Known Exploited Vulnerabilities catalog and the advisories of the Cybersecurity and Infrastructure Security Agency, at cisa.gov, where new flaws under attack appear first

Reference: The annual reports named in the watch list below: the FBI's Internet Crime Report, Verizon's Data Breach Investigations Report, Mandiant's M-Trends, and IBM's Cost of a Data Breach study

16a. The Argument on One Page

Cybercrime has become an industry with a division of labor: some criminals sell access, some rent out ransomware, some launder the money, and some run scam compounds staffed by trafficked workers (Modules 2, 5, and 6). States work alongside it, spying, preparing to disrupt, and in North Korea's case stealing to fund the regime (Modules 8 and 9). Reported losses keep rising, and most of them come from fraud that needs no malware at all (Module 1).

Almost all of it enters through four doors: stolen sign-ins and sessions, flaws in internet-facing devices, people persuaded to act, and trusted suppliers (Modules 2 to 5 and 7). The defenses are known and mostly cheap. They are phishing-resistant sign-in with hardened recovery, fast patching of edge devices, a rule that money moves only after a call back on a known number, offline backups with rehearsed rebuilds, and someone watching the alerts at night (Modules 3, 4, 5, 11, and 12).

Recovery is the real measure. The victims in this course who were back in days had isolated backups, paper procedures, and a plan they had practiced; those who took months had to improvise (Modules 6 and 12).

Policy has split. Europe has turned guidance into duties with fines, while America in 2025 and 2026 cut its civilian cyber agency, dropped software makers' attestations, and leaned on offense, insurance, and contracts (Modules 4, 9, and 14). AI has raised the pace on both sides: the same models find flaws for maintainers and run intrusions for attackers, so the advantage goes to whichever side applies them faster (Module 10). The course's answer to its central question, whether defense can keep up, is that it can wherever the basics are in place and rehearsed, and that the remaining gaps are more organizational and political than technical.

16b. The Contradictions

More flaws are exploited, while the public bodies that track them have less capacity. The number of published flaws keeps rising and attackers turn the worst into working attacks ever faster, yet the national vulnerability database has stopped fully analyzing most new entries, the CVE program nearly lost its funding, and CISA has lost much of its staff (Modules 4 and 14). Defenders are being asked to move faster with less public help in deciding what to fix first.

Fewer victims pay ransoms, yet ransomware keeps spreading. The share of victims who pay is at a record low and the total paid fell in 2025, while ransomware's share of breaches rose and leak sites keep naming thousands of victims a year (Module 6). Gangs make up in volume what they lose in payment rates,

with more attacks on smaller firms and more extortion by theft alone, so falling payments measure victims' resilience rather than attackers' retreat.

The strongest sign-in is spreading, and attackers have moved next to it. Passkeys cannot be phished, yet the most damaging intrusions of 2025 and 2026 went through help desks, account recovery, and stolen sessions, which passkeys do not protect (Modules 3, 7, and 12). Each improvement in the front door raises the value of the side doors.

America is building up offense while shrinking civilian defense. The administration's cyber strategy and its presidential memorandum expand operations against attackers and enlist companies in them, while the agency that helps hospitals, water utilities, and towns defend themselves faces further cuts (Modules 9 and 14). The targets most exposed to retaliation are the ones with the least defense of their own (Module 8).

Insurance spreads good controls but excludes the worst case. Insurers' demands helped spread MFA and tested backups among mid-size firms, while the war exclusions leave a state's attack on a whole economy uninsured by design (Modules 9 and 13). The market works for the losses it can price and steps back from the ones that matter most.

Privacy rights are growing as identity data becomes public. More states give residents rights over their data every year, and California can now delete broker profiles with one request, yet Social Security numbers and license scans already circulate in bulk (Module 15). The law protects what can still be hidden; the safeguards that work, such as credit freezes and strong sign-in, assume nothing can be.

16c. What to Watch

When	What	Why it matters
November 2, 2026	Third reading of Britain's Cyber Security and Resilience Bill in the House of Lords	Once law, it brings managed service providers and data centers under regulation, with fines, for the first time (Module 14)
November 3, 2026	American midterm elections	The results will shape CISA's budget for fiscal 2027 and whether the 2015 information-sharing law gets a long extension (Module 14)
By December 11, 2026	Expiry of the stopgap budget law and with it the information-sharing law's extension	A lapse would again remove legal protection from companies that share threat data with the government and each other (Module 14)
Late 2026, expected	Publication of CIRCIA's final rule	Its scope and start date decide which critical-infrastructure operators must report attacks and ransom payments, and from when (Modules 11 and 14)
No date set	A nominee to lead CISA	The Homeland Security Secretary made a confirmed director the condition for rebuilding the agency's staff (Module 14)
No date set	Procedures under the memorandum on private offensive operations	They will show how far vetted companies may go against criminal groups' servers, and who answers for mistakes (Module 9)
No date set	The SEC's review of its cyber disclosure rules	Repealing the incident filing would remove the main public record of material incidents at listed companies (Module 14)
No date set	The first insurer's decision or court ruling under the war exclusions of 2023	If insurers pay, the exclusions reach fewer attacks than the market feared; if they refuse and win, companies exposed to a state's conflict must insure themselves (Module 13)

When	What	Why it matters
While the Iran war lasts	Iranian operations against American companies and utilities	Wipers and attacks on controllers are the likeliest route to physical and economic harm inside America (Modules 8 and 9)
No date set	The forfeiture case over the Prince Group's bitcoin	Whether victims of the scam compounds are repaid from the coins the government seized (Module 5)
February 2027	Chainalysis's count of ransom payments in 2026	Whether payments kept falling while attacks rose (Module 6)
Spring 2027	Mandiant's M-Trends, the FBI's Internet Crime Report, and Verizon's Data Breach Investigations Report	Whether exploits stay the leading way in, reported losses keep rising, and suppliers' share of breaches keeps growing (Modules 1, 2, and 11)
Summer 2027	IBM's Cost of a Data Breach study	Whether attacks that use AI keep raising the cost of a breach (Module 10)
October 12, 2027	End of extended consumer updates for Windows 10	Machines still running it stop getting security fixes (Module 4)
October 2027	The FIDO Alliance's next passkey index	Whether passkey sign-ins pull further ahead of passwords at the services measured (Module 3)
December 11, 2027	The Cyber Resilience Act applies in full	Every product with digital elements sold in the EU must meet security duties, backed by fines (Module 14)
2030 and 2035	NIST's dates for retiring the public-key methods that protect most traffic today, set in its transition plan first published in draft in November 2024	Systems must move to post-quantum algorithms, designed to resist quantum computers, before those computers can break today's keys

Dates as announced or reported by October 31, 2026.

16d. The Last Word

Most of what went wrong in this course went wrong at a seam that someone had stopped looking at: a portal without MFA, a help desk that took a caller's word, a backup console on the same network as everything else, a supplier whose access no one had reviewed. None needed new technology to fix. The useful question for a household or an organization is therefore not whether it is secure, which no one can answer, but three narrower ones, drawn from the board's questions in Module 14. Which door would an attacker use? How fast would we notice? How long would it take to rebuild? Asked every year and answered honestly, those questions explain most of the difference between the victims here who recovered in days and those who took months.

Module 16 Key Points

Where Things Stand, and What to Watch

• Most damage enters through four doors: stolen sign-ins and sessions, unpatched edge devices, persuaded people, and trusted suppliers.

• The defenses that work are known and mostly cheap; failures cluster at seams such as help desks, recovery flows, and reachable backups.

• Rehearsed recovery separates victims who return in days from those who take months.

• Europe is adding duties with fines while America shrinks its civilian cyber agency and leans on offense, insurance, and contracts.

• Watch CIRCIA's final rule, a CISA nominee, the first ruling on a war exclusion, and the reports of 2027 that will test these trends.

Module 16 Quiz

Check Your Understanding

1. Which contradiction does the course draw about ransomware?

A) Payments have risen every year since 2020
B) Gangs have stopped stealing data
C) Attacks have moved from companies to households
D) Fewer victims pay and total payments fell in 2025, yet ransomware's share of breaches rose and leak sites keep naming thousands of victims

2. When does the EU's Cyber Resilience Act apply in full?

A) It applied in full when it was published in 2024
B) Once every member state passes its own law
C) In 2030, after a review of its costs
D) December 11, 2027

3. Why, in the course's argument, does insurance leave the largest cyber risk uncovered?

A) War exclusions remove state-scale attacks
B) Premiums are too low
C) Insurers cover only households
D) Because insurers stopped paying all cyber claims after the NotPetya cases were settled in 2022 and 2024

4. Which three questions does the course's last word suggest every organization ask each year?

A) Who is to blame, and how much
B) Which vendor, which product, and which price
C) Which door an attacker would use, how fast it would be noticed, and how long rebuilding would take
D) Whether to buy insurance, and from whom

5. Which report, due in spring 2027, will show whether reported cybercrime losses kept rising?

A) Patch Tuesday
B) The FBI's Internet Crime Report
C) The FIDO Alliance's passkey index for October 2027
D) The list of flaws whose analysis the National Vulnerability Database has deferred

Answer Sheet

Answers to every module's Check Your Understanding quiz.

Module	Answers
Module 1 Quiz	1. C 2. A 3. B 4. D 5. A
Module 2 Quiz	1. B 2. D 3. C 4. C 5. A
Module 3 Quiz	1. D 2. B 3. B 4. A 5. C
Module 4 Quiz	1. D 2. A 3. C 4. D 5. B
Module 5 Quiz	1. C 2. D 3. A 4. B 5. D
Module 6 Quiz	1. B 2. A 3. C 4. C 5. A
Module 7 Quiz	1. B 2. D 3. C 4. B 5. D
Module 8 Quiz	1. A 2. D 3. C 4. B 5. A
Module 9 Quiz	1. D 2. C 3. B 4. A 5. D
Module 10 Quiz	1. B 2. A 3. C 4. D 5. B
Module 11 Quiz	1. C 2. A 3. B 4. D 5. C
Module 12 Quiz	1. A 2. D 3. C 4. A 5. B
Module 13 Quiz	1. D 2. B 3. C 4. A 5. C
Module 14 Quiz	1. A 2. B 3. D 4. A 5. B
Module 15 Quiz	1. C 2. D 3. C 4. A 5. B
Module 16 Quiz	1. D 2. D 3. A 4. C 5. B