



MedDevice AI Security Monthly

October 2026

Postmarket threats, premarket regulation, and the model inside the device. Three sections, the last Tuesday of every month.

Jason Sinchak • ELTON • October 6, 2026



Three sections, every month.

What happened, what the regulators moved, and the thing most programs have not tested yet.

SECTION 01 · POSTMARKET

AI threats to medical devices

What September did to fielded devices and the companies that make them. A threat report, five incidents, and a model anyone can download.

About 12 minutes

SECTION 02 · PREMARKET

AI regulatory updates

What moved on the standards board, what CDRH put on its FY2027 list, the GenAI paper clock, and the EU timetable since the September session.

About 11 minutes

SECTION 03 · PREMARKET

Securing AI inside the device

How a model that already shipped gets tested. Two methods, six bounded feasibility assessments, and what the scope needs from your data science team.

About 11 minutes, then Q&A

Monthly, last Tuesday, 1:00 PM ET

Next session: October 27

Every recording: eltoncyber.com/webinar-recordings

Same three sections every month. Fresh content in each.

01

POSTMARKET • SECTION ONE

AI threats to medical devices

What September's reports and incidents mean for fielded devices and their makers.

The labor gap closed in September.

Nine months of misuse in one report. Three of the cases are tooling you now have to assume exists.

ANTHROPIC THREAT REPORT • SEPTEMBER 2026

"Sophisticated attacks no longer require sophisticated attackers."

Seven harm domains, December 2025 through August 2026. Three of the cases are about the tooling a device maker already fields.

anthropic.com/threat-intelligence-report-september-2026

GTG-20006 • the evasion loop

Russian espionage, 130 days, 24 of 27 Ukrainian targets. When a security product caught the malware, the agent rewrote and rebuilt it until the detector went quiet.

GTG-10007 • the exploit foundry

Firmware images loaded into decompilers, thousands of decompile calls, pattern hunting. More than a dozen zero-day candidates in a single month, about 50 targets.

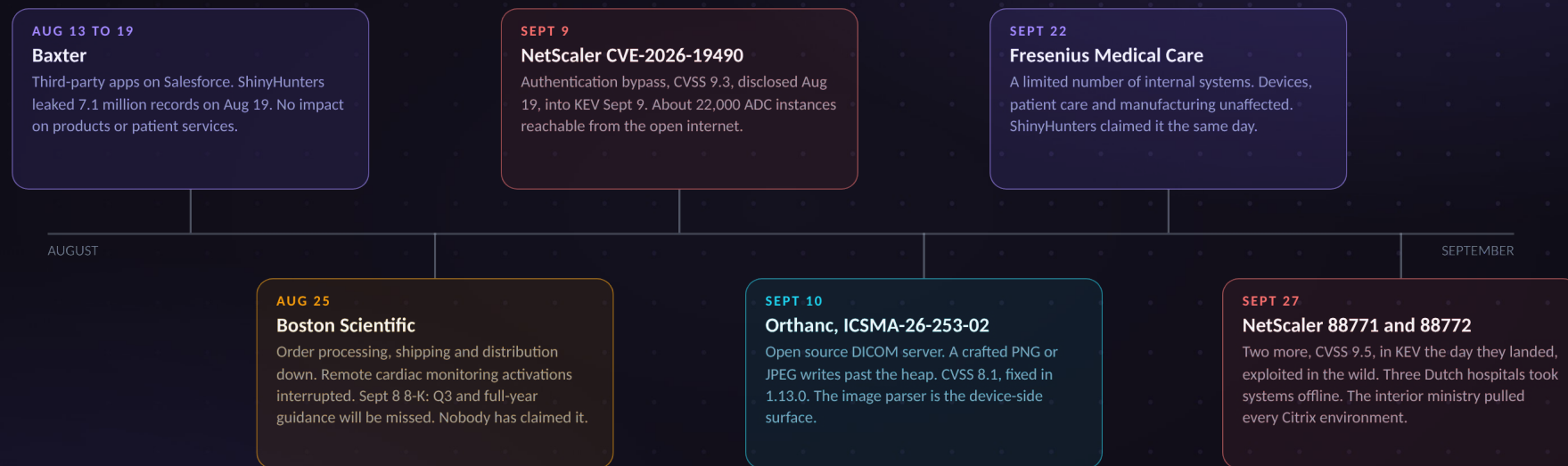
GTG-50014 • ShinyHunters

One stolen developer token to full admin in about three hours. 1.8 million Android APKs decompiled for keys. 2,100 Azure AD token sets, 40 tenants, 34 hours. Stolen AI API keys ran the next round.

The labor gap between a state team and one operator is closed. **Firmware is on the menu.**

Six weeks, six entries, three device makers.

Three device makers hit in six weeks, with two NetScaler waves and a DICOM parser bug in between.



The model you cannot gate.

Downloadable, four months behind the frontier, safeguards off for the price of a GPU weekend.

OPEN WEIGHTS - Z.AI GLM-5.3

\$20.40

A working exploit chain for two disclosed flaws on a PAC-hardened ARM64 target. 20 minutes of human attention, eight hours of GLM-5.3-Flash.

Download the weights, run them on your own box, and the usage policy is whatever you say it is.

Anthropic Frontier Red Team, Sept 29 - NIST CAISI, Sept 17

Safeguards are a setting

Bare attack request: refused. With a cover story: 64% comply. Prefilled reasoning: 92%. Abliterated weights: 100%, for about \$4,400 of compute.

Four months behind the frontier

CAISI: the most cyber-capable open-weight model released to date. ExploitBench 50 of 410 end-to-end exploits, against 56 of 410 for Claude Mythos Preview.

THE GATES ON THE FRONTIER SIDE

- Anthropic keeps Mythos behind the Cyber Verification Program. Defenders get patches and alerts, not the model.
 - OpenAI ships GPT-5.6-Cyber only in the Daybreak Red tier: 95% on advanced cyber tasks, against 57.3% a generation back and about 2% for the public model.
 - Hardware security keys required for that tier since Sept 1.
- Every one of these gates sits on the lab's own servers.

The attacker who matters to a fielded device no longer needs an account anywhere.

The published fix is the exploit spec.

Same-day exploitation was the exception. In September it was the pattern. Your clock did not move.

HOURLY

The advisory posts

Patch diff, CVE text, affected versions. The exploit spec is now public, for free.

HOUR 8

The exploit works

GLM-5.3-Flash, \$20.40, 20 minutes of a human. Two disclosed flaws chained on hardened ARM64.

DAY 0 TO 1

Exploited in the wild

NetScaler 88771 and 88772 entered KEV the day they were disclosed, already under attack.

DAY 30

Interim controls

FDA postmarket guidance: communicate compensating controls within 30 days for uncontrolled risk.

DAY 60

The fix ships

Fix within 60 days. Then the hospital's own change window, which the advisory does not control.

ATTACKER CLOCK: HOURS, AND IT HAS NO USAGE POLICY

YOUR CLOCK: 30 AND 60 DAYS, THEN THE FIELD

For most of the installed base the hospital change window behind day 60 is quarterly.

Same-day KEV used to be the exception. In September it was the pattern.

The build step that ran someone else's code.

GemStuffer: 2,000 packages, one documentation builder, and four months before anyone wrote it up.



WHERE THE SAME SURFACE LIVES IN A DEVICE PROGRAM

CI runners

Build firmware from third-party packages whose install hooks run code.

SBOM and SCA tools

Parse manifests and lockfiles written by whoever published the package.

Model registries

Unpickle and load artifacts. A model file is a program.

Doc and report generators

Read config from the thing they are documenting.

An agent swarm found the execution surface that was not on anyone's list. Ask what runs code in your build.

Three postmarket moves for October.

None of them need a budget cycle. All of them got more urgent in the last six weeks.

1

Assume N-day means same day

Patch windows for perimeter and device-adjacent appliances are now measured in hours. If a NetScaler sits in front of your remote service or your monitoring portal, its KEV entry is your incident.

Evidence: 88771 and 88772, KEV on day zero

2

Treat keys and tokens as the product

ShinyHunters did not break crypto. They decompiled 1.8 million apps for the keys inside and ran the next wave on stolen AI API keys. Scope them, rotate them, watch them, and keep them out of firmware and mobile companions.

Evidence: GTG-50014, Baxter, Fresenius

3

List every build step that runs code

Doc builders, SBOM tools, package install hooks, model loaders. Each one executes input written by someone outside the company. Put them in the threat model with the device itself.

Evidence: GemStuffer, four months to notice

None of these need AI to do, and all of them got urgent because of it.

02

PREMARKET • SECTION TWO

AI regulatory updates

The standards board, CDRH's FY2027 list, the GenAI paper clock and the EU timetable.

Standards board: one move since September.

One went from ballot to print, five stayed put, none are harmonised, and all of them still reach you.



SINCE SEPTEMBER

- ISO/IEC 27090 reached stage 60.00 on Aug 19. Ballot to print.
- IEC 63450 is still waiting on the CDV result from July.
- Harmonised in the OJ: still none. Decision (EU) 2026/1231 added no cyber and no AI standard.
- FDA recognises CR515. Nothing else on this board has a recognition number yet.

GSPPR 17.2 has said state of the art since 2021

One standard moved a lane. **Adoption is still not the trigger.** The notified body can cite any of these today.

CDRH put the AI guidance on the A-list.

Posted Oct 1. The AI-enabled device guidance goes final this fiscal year. Postmarket cyber stays a B.

A-LIST • CDRH INTENDS TO PUBLISH IN FY2027

- **Final:** AI-enabled device software functions, lifecycle management and marketing submissions. **Moved up from the B-list.**
- **Draft:** GenAI-enabled conversational devices for mental disorders.
- **Final:** Predetermined change control plans for AI-enabled devices, carried over.

Posted Oct 1. Section XII and its seven threats are inside the first item.

B-LIST • AS RESOURCES PERMIT

- **Draft:** Postmarket management of cybersecurity in medical devices. Still a B.
- **Draft:** Content of premarket submissions for device software functions.
- A GenAI risk assessment approach is mentioned as under construction, with no document listed.

Comments on the agenda itself: by Nov 30, 2026.

The AI document and the postmarket cyber document are on different lists, in different years, read by different reviewers.

26 questions, 13 days, zero security words.

FDA's GenAI discussion paper closes for comment on October 19. It still has no attacker in it.

26

questions

in five themes

2

risk axes

plus competency-based evaluation

0

security words

cybersecurity, poisoning, threat,
jailbreak, red team

CLOCK

13 days

Comments close October 19. Docket FDA-2026-N-7874.

Issued Aug 18. Two mentions each of adversarial and prompt injection, both in method and appendix text. No attacker model anywhere.

WHAT A DEVICE MAKER CAN PUT ON THE DOCKET IN A PAGE

Name the attacker. A GenAI function on a cyber device under 524B has a threat model, and the paper's two risk axes should carry it: who can reach the prompt, what the output drives, and which of the seven Section XII threats apply when the model talks back. Ask for one line that ties the GenAI framework to the AI-enabled device guidance now heading to final.

If the attacker model is going to be in this paper, someone in this room has to write it in.

Cyber and performance: one curve, two files.

Drift, poisoning and bias make one curve through one monitor. Two guidance documents claim it.

THE PERFORMANCE FILE

Read by the clinical and software reviewer

Performance drift

Model bias

Overfitting

PCCP retraining

Data quality

Monitoring thresholds, drift detection, retraining triggers, representativeness of the training set.

THE SAME EVIDENCE, READ TWICE

- Section XII lists performance drift as a cybersecurity threat. The performance section lists it as a lifecycle risk.
- IMDRF N93 draft, 4.4: poisoned data is described by its effect on performance.
- IEC PAS 63621, 6.5.7: lifecycle data quality analysis covers drift and hardware or software change in one clause.
- A false negative from a tampered model and one from drift look identical at the bedside.

THE CYBERSECURITY FILE

Read by the 524B reviewer

Data poisoning

Model inversion

Model stealing

Model evasion

Data leakage

Write access to the training set, the registry, the inference API. Pipeline integrity, provenance, access logs.

One monitor, one data pipeline, one set of thresholds. Two guidance documents, two reviewers, and a deficiency letter from either.

Three European clocks, none synchronised.

JTC 21 still writing, the Omnibus moved the AI Act date, CRA reporting live for your suppliers.

AI ACT STANDARDS • CEN-CENELEC JTC 21

- EN 18286 quality management: formal vote.
- prEN 18282 cybersecurity, 18228 risk, 18229-1 logging: enquiry stage.
- 18284 data quality: still drafting.
- Cited in the OJ: none. Target Q4 2026. Request M/613 expires Feb 28, 2027.

DIGITAL OMNIBUS • REG (EU) 2026/1744

Aug 2, 2028

Annex I high-risk, which is where an MDR or IVDR device with an AI function sits. Unconditional date.

Annex III moves to Dec 2, 2027. In force since Jul 27, 2026.

The deadline moved. The technical file did not get smaller.

CYBER RESILIENCE ACT • REPORTING LIVE

24h • 72h • 14d

Early warning, notification, final report, through the single reporting platform. Live since Sept 11, 2026.

MDR and IVDR devices are carved out under Article 2(2). The components and software your suppliers sell you are not.

Three clocks, none synchronised, all of them land in the same technical file.

Three premarket moves for October.

One of them has a date on it.

1

Comment by October 19

One page on the GenAI docket that names the attacker and ties the framework to the AI-enabled device guidance. Thirteen days. After that the paper's silence on security becomes the baseline.

FDA-2026-N-7874

2

Treat the draft as final now

It is on the A-list for this fiscal year. Section XII's seven threats are the AI test plan for any submission you are writing today, and the reviewer will read them as such.

CDRH FY2027 agenda, Oct 1

3

One evidence set for both files

Map every AI test to the FDA threat it answers and the performance claim it protects, and file it against the IEC PAS 63621 clause. The same record then serves the US file, the MDR file and the AI Act file.

Section 03 shows the mapping

Three moves, none of which wait for a document to go final.

03

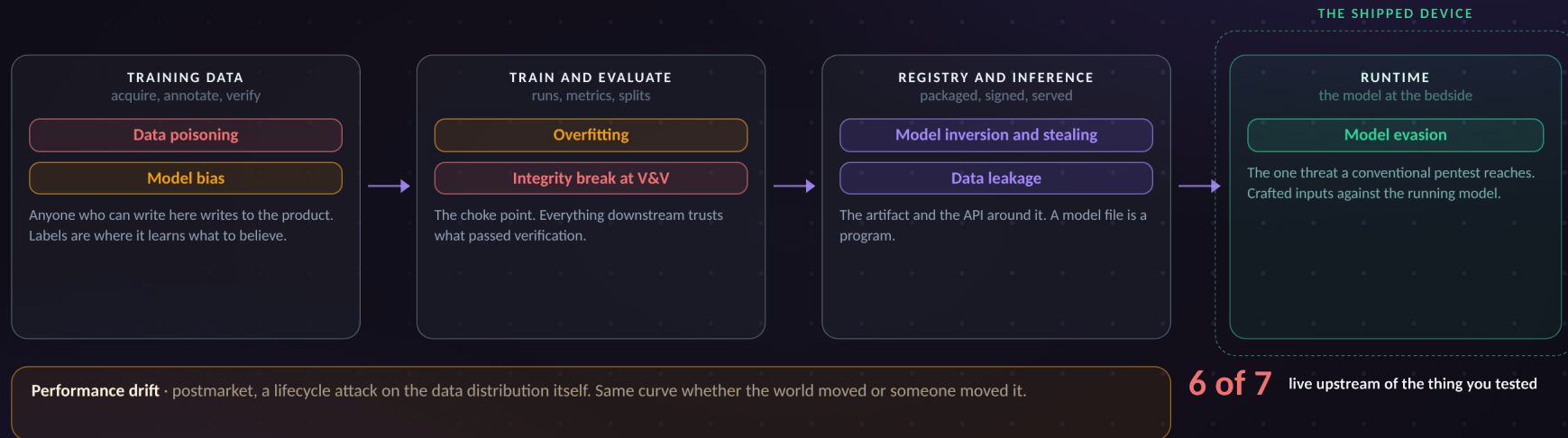
PREMARKET • SECTION THREE

Securing AI inside the device

How a shipped model gets tested, and what the scope needs from your data science team.

The model is the last frame.

Six of FDA's seven AI threats land before the device exists, in a pipeline that does not ship.



The model is the last frame. The pipeline that made it does not ship, and by the time cyber testing runs, parts of it no longer exist.

FDA named the threats, IEC the controls.

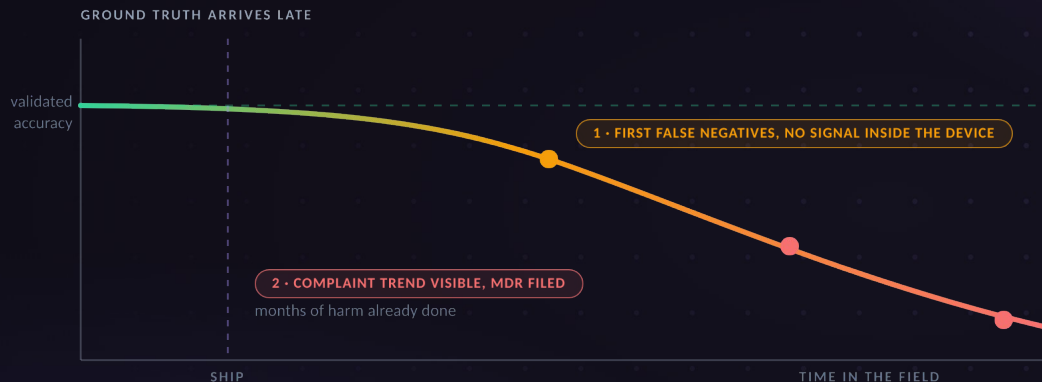
Section XII and IEC PAS 63621 describe one infrastructure from two sides. That is the test plan.

IEC PAS 63621 CLAUSE	WHAT IT REQUIRES	FDA THREAT IT ANSWERS
6.2, 6.3 Data requirements and planning	Intended use, population, representativeness and bias mitigation written down before any data is collected.	Model bias Overfitting
6.4 Data acquisition	Documented sources, sites, equipment, settings, operators, and every storage location with its access path and permissions.	Data poisoning
6.5.3 Dataset composition and revision control	Unique identifiers, approved items immutable, every change a new version, integrity protection on the set.	Data poisoning Data leakage
6.5.4 Data annotation	Named roles, who labeled what and when, every modification under revision control.	Model bias
6.5.6 Data quality verification	The set is checked against the documented characteristics before use: accuracy, completeness, authenticity, traceability.	Data poisoning Integrity
6.5.7, 6.6, 6.7 Lifecycle analysis, provisioning, decommissioning	Ongoing analysis with root cause and corrective action, explicitly covering drift and hardware or software change. Datasets delivered with a conformance statement, removed irretrievably.	Performance drift Data leakage

FDA wrote the threat list. IEC wrote the control list. Read together they describe one piece of infrastructure.

Drift and tampering draw the same curve.

Ground truth arrives months late. Pipeline integrity is the only signal that arrives first.



Drift: the world moved

New instrumentation, aging sensors, changing biology. The model does what it learned, on data it never saw. FDA calls it performance drift. IEC PAS 63621, 6.5.7, requires lifecycle analysis for exactly this.

Tampering: someone moved it

The data, the pipeline or the artifact was changed and the model does what an attacker taught it. From the output alone it is the same curve, so a performance complaint can be an unrecognised security incident.

Pipeline integrity is the one signal that arrives before the harm. Nothing inside the device reports the decline.

Same curve whether the world moved or someone moved it. Telling them apart takes evidence about the pipeline.

Prove it, or bound it.

Technical testing needs nothing from you and proves. Feasibility assessment needs access and rates.

METHOD ONE - TECHNICAL TESTING

Hands-on, reproducible, with evidence

Penetration testing across the pipeline: training data stores, the model registry, the inference API, runtime inputs, pipeline traffic, and the postmarket monitoring path.

Needs no model access from you

Every finding reproducible

CVSSv3 and v4 through the MDDT rubric

Answers: can an attacker reach the pipeline, and what can they write.

METHOD TWO - FEASIBILITY ASSESSMENT

Bounded, scored, rated

An advanced adversarial technique run once, at scale one: one target, one batch, one feed. Scored, rated for risk, with a recommended fix. Where full proof is a research project, bound it instead of skipping it.

Needs weights, data, thresholds, labels

Prerequisites listed up front

Fixed window, fixed fee

Answers: is the attack practical against this architecture, and what is it worth.

Every activity ties to one of the seven named threats. The report reads against the FDA list, with evidence per threat.

Six measurements, each one experiment.

One target, one batch, one feed. Is the attack practical here, and what is it worth.

DATA POISONING

Poison budget

How much poisoned data shifts the model toward an attacker's outcome. Fraction anchored to published results, converted to a record count, checked against the write access the technical tests found.

Needs: training set size, ingest path

MODEL STEALING

Model extraction

The query budget a copycat needs to reach useful fidelity, against your rate limits and output detail. Bare decisions without probabilities: estimate reported, no surrogate built.

Needs: API access, output format

MODEL INVERSION

One record, one gradient ascent

Whether training records can be reconstructed from outputs. With white-box access, one scored inversion against a single agreed target record. GAN-assisted inversion stays out of base scope.

Needs: weights or gradients, one target record

MODEL EVASION

Adversarial inputs

PGD, FGSM, C&W with white-box access, transfer and query methods without. Sweep the perturbation budget across epsilon to find the smallest change that flips the output, matched to the input modality.

Needs: weights for white-box, sample inputs

DATA LEAKAGE

Membership inference

A threshold attack on loss or confidence that separates records known to be in the training set from records known to be out, reporting the separation.

Needs: labeled in and out samples, the train and holdout split

PERFORMANCE DRIFT

Controlled drift

A data feed shifted just under the detector's thresholds, measuring how far performance can be walked before an alert fires. No drift detection: that missing control is the finding, the test does not run.

Needs: detector thresholds, a feed you control

Six measurements, each the size of one experiment. The answer is a number and a rating, not a demo.

Bounded by design.

Hard edges are what make an adversarial ML engagement priced, scheduled and deliverable.

PREREQUISITES UP FRONT

Each assessment lists what it needs: weights or gradients, training data, thresholds, labeled samples. You provide them with working access before testing starts, confirmed at kickoff.

Due ten business days before the window opens

NO ACCESS, STILL DELIVERED

A prerequisite missing, or access that does not work in the window: the threat is assessed from the architecture, the technical findings and published results. The gap is recorded as a scope limitation and the activity is deemed delivered.

The window does not slip. The fee does not change.

NO CONTROL, THAT IS THE FINDING

No ingest monitoring, no drift detection: the missing control is the finding and the dependent test does not run. Nobody is charged to test something that is not there.

The gap goes in the report with its rating

OUT OF THE BASE SCOPE, QUOTED SEPARATELY

Full-scale demonstrations. Poisoning or bias injection run through the training harness. Retraining. GAN-assisted inversion. Hardware side-channel timing work. The base engagement proves what is reachable and rates what is not.

Hard edges are what keep an adversarial ML engagement priced and deliverable.

Three questions for your data science team.

The answers are the scope. Most teams can get them in a week.

1

Who can write to the training set?

Not who is allowed to. Who can. Every storage location with its access path and permissions is a 63621 line item, and the first thing a technical test goes looking for.

Answers the poison budget prerequisite

2

What does the API give back?

Bare decisions, probabilities, or logits. That one design choice sets the query budget for extraction and the reach of inversion, and it is usually decided for a UI reason.

Answers the extraction and inversion prerequisites

3

At what threshold does the monitor fire?

If there is a number, controlled drift walks up to it. If there is no number, there is no monitor, and that is the finding for both the cyber file and the performance file.

Answers the controlled drift prerequisite

Three questions for your data science team this week. The answers are the scope.

See you October 27.

Every recording in one place, and the newsletter in between.

NEXT SESSION

October 27

Last Tuesday of the month, 1:00 PM ET. Same three sections, new content.

Registration link in the follow-up email, and on eltoncyber.com/webinars.

EVERY RECORDING

[eltoncyber.com/ webinar-recordings](https://eltoncyber.com/webinar-recordings)

Today's session, the slides as PDF, and September's session with Health-ISAC, Illumina and Triple Ring.

BETWEEN SESSIONS

The AI newsletter, weekly

Section 01 is built from it. Issue 18 covered GLM-5.3. eltoncyber.com/newsletters.

QUESTIONS

jsinchak@eltoncyber.com · the AI-Enabled Device Testing page is at eltoncyber.com/ai-enabled-device-testing

Thank you.