

WHITE PAPER · DECISION SUPPORT SYSTEMS IN PUBLIC SAFETY

The record was always there. Nothing carried the decision.

Twenty years from a thumb drive carried across a detective bureau to a decision support system built on a department's own directives, and how to put a language model near police work without letting it decide anything.

AUTHOR

Justin Frank

Philadelphia Police Department, 2003 to 2023: detective, sergeant in the 39th District and Chief Deputy at the Sheriff's Office. Chief Executive Officer, GY6 Solutions.

EDITION

Version 3.8

November 2026

Every screen referenced in this paper is synthetic.

THE ARGUMENT

Police technology stores records and analyzes them. Nothing carries the decision.

FIRST

Systems that store

Records management and dispatch. They made the record exist and made it retrievable.

SECOND

Systems that analyze

Fusion and analytics on top of the record. They made patterns visible, then stopped at a screen.

THIRD

Systems that decide

This generation does not exist. Nothing carries a decision to the person with authority, records the answer and releases the work it was blocking.

A detective bureau loses cases when a video retention clock runs out while a preservation request waits on an approval nobody knew was pending. It loses them when a canvass is assigned out loud and recorded nowhere, or when two divisions work the same name for three weeks.

Better search touches none of that. It is a routing and accountability problem, and no system has solved it because the rules for what must happen next, and which rank may approve it, were never machine readable. They live in a binder and in a supervisor's memory. CompStat is a meeting. What the market sells as CompStat is a dashboard, and the work stalls in the distance between the two.

> **The claim.** A department's own directives are the missing piece. Load them, and a system can say what happened, what is supposed to happen next and who may decide it. It is also the only defensible place for a language model in police work. Grounded in agency data alone, a model writes a confident paragraph with no authority behind it. Grounded in written policy with the section cited on screen, it gives a supervisor something to check, a defense attorney something to open, and the system a reason to refuse when policy requires a rank that has not signed.

What follows

- The model proposes, with the governing section shown. A named person answers, and the record carries who and why.
- Nothing is replaced. This is a layer over the systems the department already owns.
- Everything runs in the department's own tenant, under its own retention, labels and identity.
- The corpus belongs to the department. Change a directive and the check changes with it.

What this paper declines to do

- **Put a number on your savings.** Time-saved figures in vendor papers come from another department or are invented. Section 2 shows how a department measures its own.
- **Call a design a finished product.** Sections marked designed are specifications, and the line is held on every page.

PART 1 · LINEAGE

A thumb drive, then a system, then a way to read the rules.

> *"When I made detective, the way information moved between the front of the detective bureau and the back was a thumb drive. Somebody carried it. I built a system to fix it. It worked, it grew into the workflow behind CompStat, and parts of it still run in that department today. I am not asking anyone to trust a system. I am asking them to look at one they can inspect."*

Every system involved worked. The records system held records and the statistical process produced statistics. What was missing was a carrier: something to move a piece of work, and the obligation attached to it, across the building without a person walking it there.

DBIS, the Detective Bureau Information System, was built in the Philadelphia Police Department's Northwest Detective Division, adopted citywide and extended into the special investigative units. A structure-only extraction of the surviving system counts:

27 files
24 readable Access databases and three workgroup-security files

~2.4M rows
across the whole estate

2007 to 2024
roughly, through stats-year 2024

Counts come from an mdbtools extraction recorded in the product knowledge base. They describe shape, not contents. No record from it appears in any GY6 product.

WHAT THE BUREAU TAUGHT	WHERE IT SURVIVES IN THE PRODUCT
The assertion needs checking, not the record. A database that returns four addresses for one subject holds four records linking them, and says nothing about where he lives.	Verification status sits on the relationship between two entities, never on either one alone.
Bedrock and leads are different things. The system of record is what you verify against. Everything arriving from outside is a claim until a person evaluates it.	A bedrock layer that is ground truth, and an investigative layer that enters with a default status and waits for a human.
What stops a case is usually an obligation nobody carried: a preservation request, a camera pull, an approval.	A step routes to the rank that can act, and the record carries who was asked, when, and what they answered.

Why it could not be built sooner

Three things arrived together. Agencies now run their own cloud tenants, so a decision layer can live inside one. A directive corpus is a few hundred unstructured PDFs, and until language models could return an answer with a section number attached, the rules could not take part in a system at all. And retrieval over a controlled corpus turned a confident paragraph into a citation a supervisor can open. PRISM carries the DBIS architecture forward as a vendor-neutral decision support system. PRISM IDSS loads the department's directives into it, which is the first machine-readable statement of what must happen next and who may approve it.

> **The shift.** Models have been sold into this market for a decade, mostly as prediction, with uneven enough results to produce moratoria. The change here is that the governing authority becomes part of the system. The machine says: this is the step, this is the section that requires it, this is the rank that must sign, and I will not proceed until they do. That answer can be refused, audited and acted on.

PART 2 · WHAT DECISION SUPPORT RETURNS

The saving is work that stops waiting. The research says where to stop.

WHERE TIME IS LOST	WHAT CHANGES
Rebuilt answers. Priors, cases, subjects, cameras and warrants assembled by hand.	Assembled once, each line sourced. Detective hours move from collection to investigation.
Approvals nobody knew were pending. The retention clock runs while the form waits on a desk that was never told.	The request routes to the named rank, can be answered from a phone, and evidence is preserved inside the window.
Duplicated work. Two divisions on one subject, neither aware.	Deconfliction goes to the supervisor whose authority covers both.
Weekly packets and oversight requests. Numbers rebuilt and argued over, and access reconstructed by hand across systems.	One method across divisions, every figure opening to its records. Every access carries a stated purpose, so an oversight answer is a query.

Why there is no percentage here. A department should measure its own savings, before and after, on a scope it chose: hours from preservation request to approval for one division over one month, items that expired while pending (usually unknown, and the most persuasive number in the building once it exists), duplicate-effort incidents per quarter, and hours spent on the weekly packet and on one oversight request.

What sixty years of decision research forces

FINDING	WHAT IT FORCES HERE
Decisions sort into bands. Structured ones belong to automation, unstructured ones to people, and support earns its place in the semi-structured middle. (Gorry and Scott Morton, 1971)	Payroll is structured. Which door is worth knocking on is semi-structured, and that is the only band this product serves.
Experts do not compare options. Fireground commanders weighed one workable course first. (Klein, 1985)	No ranked menu. The product scores the gap, and a card shows the one fact that would change the answer.
Two failure modes. Automation bias causes errors in novices and experts. Algorithm aversion makes people drop an aid after seeing it err, unless they may adjust it. (Parasuraman and Manzey, 2010; Dietvorst, 2015)	Bounded adjustment, with a one-line reason on an override that is never a punishment. High acceptance is a warning sign.
Explanations raise acceptance, not accuracy. (Bansal, 2021; Zhang, 2020; Buçinca, 2021)	A fluent rationale is no safeguard. Cards show cited evidence and where models disagreed, and are tested on wrong recommendations.

These findings come from aviation and fireground command, which match police work in time pressure and irreversible consequences. In Endsley's terms a dashboard serves level one, perceiving the elements. The detective already does that. What is missing is level two: what the elements mean together, what is absent and what the governing rule says next. Level three, projecting what happens, stays with the person.

> **House rules that follow.** Evidence first, recommendation after, no pre-ticked default, agreement shown as a three-state status and never a percentage, human answers collected blind before any tally. Ratified October 7, 2026: conflicts escalate automatically to the next-level supervisor on critical decisions, and the recommending engine uses more than one model. The rest is proposed until ratified.

PART 3 · SOLVING YOUR OWN PROBLEM

Improve what the department owns, and write down what it already knows.

The usual instinct is to buy a platform and move onto it. The foundational study in Communications of the ACM found that enterprise packages carry an embedded business model, and that breakage follows when an organization's real procedures were shaped by a different setting. It named three kinds of misfit: data (incompatible formats), functional (checks and steps that do not fit the work) and output (a field the organization must report that the package has no place for). Output misfit was the most common. The study also advised examining processes two or three levels deeper than reference models, and giving key users an active role because they hold the tacit knowledge.

Policing is harder than business. A manufacturer accepts some disruption by adopting a vendor's procurement process. A police department cannot adopt a vendor's process at all, because its processes are directives, consent decree provisions, statute and case law. In a factory an output misfit is an annoying report. In a detective bureau it is a disclosure obligation the system has no field for.

That reverses the modernization question. The department's processes were never written anywhere a system could read them. They exist as a directive corpus nobody indexed, and as the tacit knowledge of supervisors who learned the rules by doing the job. Nonaka's work on knowledge creation names the hard step as externalization, making the unwritten writable. A directive corpus plus a decision record is externalization in its most literal form.

1 · KEEP THE ESTATE	2 · MAKE RULES READABLE	3 · CAPTURE THE TACIT LAYER	4 · AUTOMATE THE DECISION STEP
Records and dispatch stay. Nothing is migrated or re-entered, and accreditation stays intact.	Index the directives the department already publishes. This takes a week, not a program.	A sergeant's knowledge of what follows becomes a check with a citation, authored and owned by the department.	The judgment stays with the person. The software carries the rest.

Who holds the leverage

	RENTED MODEL	OWNED MODEL
Data	Copied into the vendor's environment under the vendor's retention. Leaving means an extraction project.	Never moves. Stays under the city's retention, labels, legal hold and identity.
Rules	Configured in the vendor's platform and maintained by them. A directive changes and the city files a ticket.	The directive corpus is the city's. Change the policy and the check changes with it.
Knowledge	Accumulates as vendor know-how about your department and leaves when they do.	Accumulates in a knowledge base the city owns and keeps whoever builds next.
Leaving	Migration, re-entry, retraining and a fight over the exports.	Stopping leaves every system as it was. Nothing to unwind.

The asset created is a knowledge base the department owns, not a license. It keeps its value if the vendor is replaced and survives a change of administration. It also serves more than one client. The investigation surface, the command dashboard and a public application can all read one knowledge base through the Model Context Protocol, instead of keeping separate copies.

Soh, Sia and Tay-Yap, Communications of the ACM; van Beijsterveld and van Groenendaal, Information Systems Journal, 2016; Nonaka, Toyama and Konno, Long Range Planning; Microsoft Learn, Model Context Protocol.

PART 4 · FROM DECISION TO EXECUTION

A decision that is not executed changes nothing. Follow one request from the desk to the ground.

The example is a leave request, because the mundane case is where the time goes.

STEP	WHAT EXISTS TODAY	WHAT THE MACHINE DOES	WHO DECIDES
Request	A day-off request sits in a schedule the supervisor opens separately.	Nothing. It stays a record.	Nobody yet
Consequence	The supervisor works out by hand what granting it leaves uncovered.	Counts availability against minimum staffing by assignment and rank, and names the post left open.	Nobody yet. Arithmetic.
Rule	The governing contract and directive terms live with whoever has been there longest.	Surfaces the clause that governs the move, cited, with its date and status.	Nobody yet. Retrieval.
Call	Approved eventually, recorded with no reasoning.	Proposes. Moves no one.	A named supervisor, in a sentence, from a phone.
Execution	Someone retypes it into the roster and people learn at roll call.	Writes the order, records who approved it under which rule, and notifies everyone affected.	Already decided. Clerical work that should never have been human.

The same chain under pressure, with location in it

Two call-outs land at 2200 and a sector goes thin. The question stops being who is on the roster and becomes who is available and where. A roster says a unit is assigned to the 22nd District. A wrist device says that unit is eleven blocks away on a foot assignment and has been stationary for nine minutes. Those are different facts, and only one is current. The count, the distance, the coverage cost and the governing clause are computable. Whether this sector can run thin tonight is not, and never will be. The exchange runs as a plain-language conversation from a phone, and the output is an order with a name on it, not a dashboard a commander is invited to interpret.

Satellite location is evidence, not knowledge. It degrades near buildings and tree cover, and a position shows where a device reported itself to be at a time. It does not say where a person is.

Where the line is drawn with field communications

This product owns the case, the record and the decision. Push-to-talk, channel audio and field messaging come from a separate field communications product. The working example is the interop between PRISM and MSX, the GY6 field communications product. What crosses is an event with a case identifier and a time: that a transmission happened, by whom and when. Audio crosses only where retention rules say so. Moving a reference instead of the media keeps one retention authority, and a meeting held beside the case is the thumb drive again in a newer format.

Constraints named before anyone asks

- Platform scheduling runs in the standard government cloud and not in the high-assurance or defense tiers. Agencies in those tiers need a different route.
- Shared channels do not support guests or platform-native push-to-talk, external participants cannot be placed on legal hold, and in government clouds platform file sharing is unsupported in shared channels. Partners join by directory-to-directory connection that both agencies configure.
- Vendor-managed connectors to external workforce systems were retired in December 2025. Integrating an existing scheduler is a build and should be priced as one.

Microsoft Learn: Shifts and Shifts connectors; Walkie Talkie in Teams; Shared channels. Samsung Galaxy Watch public-sector documentation.

PART 5 · THE OFFICER-SAFETY LAYER

A department can locate a vehicle, a phone and a case file. It cannot reliably locate its own officers.

Every other problem in this paper is too much information badly arranged. This one is the opposite: the fact exists nowhere, and no analysis recovers a fact nobody holds. An officer stops transmitting, a foot pursuit leaves the last assignment behind, a unit enters a building and the next radio call comes from somebody else. An assist call asks for help now. It gives no position, and an officer under fire may manage nothing.

DISPATCHED	ON SCENE	SHOTS	AFTER
The system knows the assignment and address. An assignment is a plan and not a measurement.	A disposition is given or it is not. Voice is the only channel, and voice needs an officer able to use it.	The picture ends. Whatever was not transmitted in the prior seconds exists nowhere.	A search by people moving toward the sound for a fact an instrument on the wrist would have reported continuously.

The general shape, not any one incident.

> **The claim this section refuses to make.** No incident is cited, and no claim is made that this capability would have changed any outcome. Those are after-action findings that belong to the department conducting them. A capability claim carried on a line-of-duty death would be the failure this paper refuses elsewhere. This is also the one place where a model contributes nothing. The answer is an instrument, and the decision layer's job is to carry what the instrument says to the person who can act, fast enough to matter.

One location source is a single point of failure for the fact you cannot lose

SOURCE	WHAT IT KNOWS	WHERE IT FAILS
Wrist device, satellite	The position of a device worn on the person, reported continuously. Dual-frequency positioning on current hardware.	Degrades near buildings and tree cover, worst indoors, in stairwells and below grade. Battery over a long tour.
Radio, satellite	The position of the unit the officer will not leave behind.	Needs line of sight. Voice preempts data, so reports can lag at peak traffic.
Cellular and wireless assist	Position where satellite cannot reach.	Needs a network. Accuracy varies with density.
Dead reckoning	Carries the last good fix through a loss of signal.	Error grows silently. Good for minutes, not an hour.
Vehicle location	Where a vehicle is, and only that.	Knows nothing about where a person is.
Terminal login	Nothing. It was never a location source.	The one that causes harm, because departments rely on it.

The identity problem inside vehicle location. A vehicle system reports where a car is. The claim that a named officer is in it comes from a login, an assertion made at the start of a tour and never revisited. It survives a foot pursuit, a partner swap, a car left running and an officer who never signed out, and it can be wrong without anyone lying. This is the paper's first rule turned on itself: the vehicle's position is bedrock, and "Officer Reyes is in that vehicle" is an assertion that needs its own status, source and clock. A wrist device collapses the problem, because the thing located and the person sought are the same object. The design takes every position the department already produces, labels each with what it measures, and stops letting a login pass for a measurement.

Samsung public-sector Galaxy Watch documentation: dual-frequency positioning, freestanding operation, device management and stated limits.

PART 5 · THE OFFICER-SAFETY LAYER

A single threshold is a false-alarm machine. A conjunction that never happens on an ordinary tour is not.

DESIGNED AND PROPOSED, NOT BUILT

This is the alert that must still be trusted at three in the morning on the four hundredth night. Heart rate above a number at that hour is a flight of stairs, a foot pursuit or a bad dream on a meal break, and an alert that fires on it is dead within a month. What is diagnostic is four conditions in one short window:

ASSIGNED	SILENT	STILL	BROKEN BASELINE
On scene, not in service or on a meal break.	No voice traffic or disposition longer than that assignment type normally runs.	Motion has stopped and not resumed.	A physiological break from this officer's own baseline, or loss of signal, which is information too.

Each alone is noise. All four together are rare, and rarity lets an alert survive a real platoon. The baseline is the officer's own and never a health record. The system watches for a break from a pattern and stores no condition, diagnosis or trend. The vendor states the device is not for medical use, so heart rate is a signal and never a diagnosis. The strongest single report is a one-touch duress alert, which is intentional and unambiguous and is the only one that needs the officer able to act.

1 · ASK THE OFFICER	2 · TELL THE SUPERVISOR	3 · THE BAND
A silent prompt on their own device with a one-tap clear. Most of what the conjunction catches ends here.	No answer inside the window, and the sergeant gets the position and the time last confirmed, nothing else.	A person escalates. The machine proposes and never makes the call, because a wrong broadcast costs everyone on that radio.

Clearing is human and recorded. An alert is never cleared by elapsed time.

Two systems, two audiences, one wall

	OPERATIONAL LAYER	WELLNESS LAYER
Question	Is something wrong right now, and where is this officer?	What is this officer's own trend, and what would reduce their risk?
Who sees it	The supervisor, and only a position, a time and an alert condition.	The safety office and the officer. Not the supervisor, district or commander.
Participation	A condition of the assignment, bargained.	Voluntary, withdrawable without consequence. Never an input to fitness for duty.

The wall is the precondition for any of this. If officers believe wellness data reaches the chain of command, they will not wear the device and the operational layer loses its sensor. The wall has to be demonstrable in the schema and not asserted in a memo.

A BMJ case distribution study of more than 4,500 U.S. police officer deaths from 1984 to 2010 found 441 sudden cardiac deaths concentrated in the duties the job is made of: restraints and altercations (25% of cases, 34 to 69 times routine risk), pursuits (12%, 32 to 51 times), physical training (20%, 20 to 23 times) and medical and rescue operations (8%, 6 to 9 times). The authors call for cardiovascular prevention among officers. That is the argument for giving a safety office instruments it has never had. A case distribution study shows where risk concentrates. It does not show that any intervention prevents an event.

What it will never do

- Measure performance or response time by officer.
- Support discipline. It is not a source for an internal investigation, and retention is short enough that it cannot become one.
- Infer health. It reports a break and not a meaning.
- Broadcast silently. A person makes the escalation.

Bargain before building

Officer location is a condition of employment and in many jurisdictions a mandatory subject of bargaining. That conversation is gate one. The pilot measures a single thing, the time from an assist call to a confirmed position against the department's own records, and names its failure modes: degraded indoor position, connectivity (it supplements the radio and never replaces it), battery at hour ten, and the three a.m. false positive, judged as seriously as the catches.

Varvarigou, Farioli, Korre, Sato, Dahabreh and Kales, BMJ 2014;349:g6534. The conjunction rule, escalation ladder and retention window are specifications. None is built.

PART 6 · THE CASE AS A PIPELINE

Closing a case works like closing a sale. Score the gap, not the case.

EMERGING, DESIGNED, NOT BUILT

An opportunity advances through gates as evidence of real commitment accumulates, and a case advances the same way. Both are worthless once a step can be asserted instead of evidenced. The measure is called the P-Solve Factor™.

Published solvability instruments produce one number that decides whether to work a case. That is triage, and it draws serious criticism: screening out, geographic disparity, and the objection that where a victim lives should not decide whether a case is investigated. The useful question is what the case is missing that would make it solvable, and whether it can still be had.

PUBLISHED INSTRUMENTS REPORT	THIS REPORTS
A total against a suspend threshold	What the case has, each element pointing at its record
A decision to work or suspend	What it lacks, ranked by weight, and what is still obtainable and for how long
	The next action that moves the case most

A measure that says what to do next cannot be used to justify doing nothing, which answers the triage critique in the design. Cases pass four gates, each evidenced and none asserted: the record is open, perishable evidence is secured (the preservation request answered, not merely sent), people are interviewed, and the offender is identified with evidence to support arrest and charge. Federal clearance reporting supplies the terminal condition. Its exact wording must come from the reporting specification before it reaches any product surface.

Every missing element carries a clock. "Missing: video. Retention window closes in six days" is a work order and not a score. Evidence age (offense to report) and case age (report to now) are tracked separately, which shows commanders what was obtainable at intake and lost to delay.

Divergence monitoring. Weights do not travel: a factor that predicts clearance in one force fails in another. The design ships published weights as the default and never fits weights to the agency's own outcomes, because that is how an instrument becomes a prediction engine. It reports the divergence between published weighting and local experience instead. Every element also carries a known-at timestamp. In one study the strongest predictor was "arrest made," an outcome that had leaked into the predictors. Without a known-at time, a solvability score measures hindsight and calls it solvability.

The vehicle sweep: matching one vehicle against a city's crime

DESIGNED, NOT BUILT

Correlate every plate read against every offense and the list is long, confident and useless, because a commuter's vehicle co-occurs with its corridor's crime every day. The signal is excess over what the vehicle's own pattern predicts. Four corrections make that an instrument: expectation from the vehicle's own reads by hour and weekday, coverage normalization (a block with no reader is no evidence of absence, and the output says so), a false-discovery cut with the expected number of false leads printed beside the results, and windows set by the offense. Ninety minutes and three miles is not co-occurrence. Four minutes and four hundred feet on the approach is.

A read places a vehicle and not a person, and the step to a person is an assertion with its own status, source and clock. A citywide sweep requires a criminal predicate recorded with the query, and a query without one does not run. The sweep returns candidate clusters on method features, each case shown as elements held, such as 7 of 22, plus the one obtainable element that would move the most cases. For example: recover exterior video at the third location, the only one of nine with a camera on the approach, before its retention window closes in six days. That either corroborates the vehicle at a second scene or collapses the cluster before anyone knocks on a door, and a cluster killed cheaply on day one saves nine cases a month of waste.

No person is scored, no suspect is produced, reference layers never volunteer names, and nothing is promoted by repetition. A claim seen in nine cases is one claim seen nine times until a person evaluates it. Element weights are uncalibrated, and most carry an explicit flag saying so.

Two agencies serve the same order and neither can see what the other did. Proving you looked is the real duty.

Protection orders

CAPTURE DEMONSTRATED ON SYNTHETIC RECORDS STATE DATABASE WRITE-BACK DESIGNED, NOT BUILT

This is a design gap inside a statute that sends every order to two agencies and never required anyone to share what happened next. It is not a compliance failure and should not be sold as one.

INSTRUMENT	WHAT IT COMPELS
23 Pa.C.S. § 6105(e)	The prothonotary transmits the order within 24 hours and State Police enter it within 8. The listed fields describe the order. There is no service field.
§ 6106(f) to (g), as amended 2025	The court orders the sheriff or appropriate agency to serve, and the order goes to both the police and the sheriff with jurisdiction.
Pa.R.C.P. 1930.4(h)	Proof of service stating date, time, place, manner and who was served, filed within ten days. The only recording duty with a clock.
Directive 3.9, PPD, revised 02-02-24	Officers assist service and complete the affidavit on a 75-48. The Operations Room Supervisor scans each one and emails it to the Sheriff's Office before the end of each tour.

The deconfliction channel between the agencies is a scanned email once per tour. It is not queryable, and at two in the morning on a violation call neither agency can say whether the order has been served. No statute, rule or directive we found requires service to be entered in the statewide database or sets a deadline for it, and no published source documents duplicate service in Philadelphia. The defensible claim is that the architecture permits it and cannot detect it.

The design captures service at the door, from a phone: the affidavit as structured fields or a photo of the signed form, a pin with its own accuracy and timestamp, body-worn video correlated by control number, any failed attempt, and the serving unit, visible to the other agency at once. Attempts are the evidence of diligence, and a system that records only success discards the half a court asks for. The supervisor's duty shrinks to approving the workflow on one screen. Open and owed: whether the statewide database carries a service-status entry, and who populates it in this county.

Warrants

When a defendant did not have reasonable notice, the Commonwealth must show due diligence in finding them or the time does not come off the speedy-trial clock. A Pennsylvania court put it plainly: "the mere issuance of a bench warrant does not establish due diligence." That account lives in a warrant log, an attempts table, radio records and service reports across units that never shared a database, so the department cannot say which warrants it could defend. The original system already knew attempts mattered: its warrant log held 5,601 warrants against an attempt table of 11,099 rows. What changes is that any authorized person can pull every attempt on demand.

A live warrant against a dead person gets acted on at traffic stops and in background checks. The system surfaces candidates, where a subject matches a death record and attempts have gone silent for years, for a person to verify, because a match is a claim until somebody confirms it. Administrative closes produce no arrest and register in no department measure. Make them countable and they start happening.

Affidavits

An affidavit is a list of facts, each pointing at a record, measured against a legal standard. That makes it the best thing a decision support system can prepare and the worst thing to let a model write unsupervised. Each sentence resolves to a record with a known-at time. A ballistic correlation is carried as a laboratory result with a reference number and never paraphrased into a conclusion the lab did not reach. The requested time window carries a justification, so an over-broad request shows before it is sworn. Facts considered and excluded are listed with reasons and shown to the affiant before signature, which a suppression motion is built from.

Built and demonstrated: a complete search warrant application runs in the demonstration today, drafted from the case, reviewed by a named affiant and approved by a named supervisor. Cell-site data is a new category on the same machinery. The Supreme Court held that a person keeps a reasonable expectation of privacy in the record of physical movements captured through cell-site location information, so more than about a week of it generally requires a warrant. The Court wrote narrowly, and a drafting system must not quietly settle the open questions on real-time location, tower dumps and exigency. It cites the standard and makes the detective and supervisor answer. The system drafts and never swears, and it never writes the sentence that says probable cause is met.

PART 8 · GUARDRAILS

Most of the safety is where the uncertainty is stored. Guardrails built into the data model cannot be bypassed.

DUAL-LAYER GRADING: RATIFIED SPECIFICATION, DESIGNED, NOT BUILT

A guardrail written as a rule gets bypassed under pressure. One built into the schema cannot be, because the schema has nowhere to put the unsafe thing.

Status lives on the relationship. An address is a valid location and a person is a valid person. What needs verification is the assertion connecting them. A commercial database returning four addresses for one subject holds four records linking them, and systems that attach confidence to entities flatten that into one lead, which is how a door gets knocked at the wrong house. Here the source says four records exist, the system stores four assertions each with its own status, source and date, the detective sees four unverified claims with the source on each face, and a person evaluates and signs. Bedrock is the records system data, which is ground truth. Everything arriving from elsewhere stays a claim until a person evaluates it, and nothing is promoted by time or repetition.

Confidence a detective can read at 2 a.m. and an analyst can defend in court. Field states are verified, unverified, inferred and disputed, adapted from the NATO and Admiralty tradition. Stored grades are separate, because one word cannot carry how reliable the source is and how credible this report is. A reliable source can report something improbable, and an assertion can be verified and no longer current. The columns for separately gradeable reliability and credibility do not yet exist, and no surface of the product describes them in the present tense.

REFUSAL	WHY	WHERE IT IS ENFORCED
No forecasting, scoring or prediction about a person	A risk number attached to a human is an assertion no evidence standard accepts, and once it exists it becomes an input and then a justification.	Items carry confidence. People carry no suspicion value because there is no field for one.
The engine proposes a step, never a suspect	A system that names people does the detective's job badly. One that names steps does its own job well.	Recommendations are built around actions. Declining to recommend is a valid output.
Reference layers never volunteer names	Registration data exists for compliance. Surfaced unprompted, it becomes a suspect list.	Answers questions an investigator poses. Never an input to a ranking of a person.
A fuzzy match always shows its score	A mistyped serial is a false lead at a named person.	Confidence-scored and displayed, never silent.
No draft policy is ever cited	Citing a draft directive is the same failure as citing the wrong one, and harder to spot.	Draft directives are excluded from retrieval by default.
No citation ships until retrieval is measured	A wrong citation in front of people who know their directives ends the engagement.	A gold-set retrieval threshold must pass first.
A human decides and the record says who	Nothing is confirmed because a machine was confident.	The model can be switched off and the labels, sources, approvals and audit trail still stand.

Decision cards: how the work gets decided

A card carries one question and explained directions, with the evidence first, the recommendation after and a named person on the answer. A sergeant's card reads: two off on the four-to-twelve, cover from the detail or run the sector thin tonight? A detective's reads: the sweep returns nine candidate cases, open the cluster or work this case alone? A captain's reads: nine hours on the detention clock with two checklist items outstanding, escalate now or hold? Nothing is preselected, because a displayed recommendation is a default and a default steers the answer. On contested cards the decider records a leaning before the recommendation appears. The card shows the one fact that would change the recommendation and never names the model.

No card names a suspect, scores a person, releases anybody, files anything or moves a unit. Going against a recommendation costs one line of reason, recorded as a record and never a punishment. Declining to decide is valid, and the snooze is recorded. Acceptance, override and reversal rates are tracked, and near-total acceptance is a warning sign because it is what over-reliance looks like from outside. A demonstration that only shows correct recommendations teaches the wrong lesson, so cards are exercised on cases where the recommendation is wrong.

The regulatory frame was designed against, and one model in one context window is the wrong architecture.

28 CFR Part 23 requires criminal intelligence information to be labeled for sensitivity and confidence with the submitting agency identified, held to a criminal predicate, disseminated on right-to-know and need-to-know, and reviewed and purged. The labeling model above meets the confidence requirement with room to spare. The FBI CJIS Security Policy governs systems and people, so personnel screening and the security addendum are part of deployment.

State statute is converging on the same shape. California's SB 524 requires disclosure of the program used, an audit trail to the source media, retention of the first draft, officer certification of accuracy, and bars vendor resale of agency data. Utah's SB 180 requires a disclaimer and a certification. A vendor whose product generates reports and holds agency data has to retrofit all of it. A system whose records were always labeled, retained in place and approved by a human has nothing to retrofit. Verify current statutory text and status against the primary source before relying on any of this in a procurement.

Why the single box fails, on the vendor's own account. The naive design puts everything known about a case into one prompt and asks for a conclusion. It demos well and is the worst available design for this domain. Microsoft's guidance says a retrieval system "must return highly relevant, concise results." Its example: a model accepts roughly 128,000 tokens, the corpus is ten thousand pages, and sending everything wastes tokens and degrades quality. Past a point more context means a worse answer, silently, and the result is a fluent synthesis where no claim traces to a source. In business that costs money. In an investigation it produces an unreviewable conclusion about a person.

RETRIEVAL	ASSEMBLY	PROPOSAL	DECISION	EXECUTION
A narrow question against a controlled corpus, returning passages with citations. Access is trimmed to the asker's permissions first.	Sources gathered, each line labeled with status and origin. Conflicts are surfaced and not reconciled.	A next step with the governing directive cited and evidence shown. Declining is valid.	A named person with authority answers. This level has no model in it.	The answer is carried out and recorded with who and why, releasing what it blocked.

No single inference spans raw material to consequence, each level is separately inspectable, and the one that matters most contains no inference at all. Access is enforced at retrieval, inheriting the organization's identity permissions instead of inventing a second set. **Token allotments** budget model capacity per seat, which makes consumption attributable to who asked and against which case. An unlimited assistant invites speculative querying about people, and a budgeted one makes each query a small act with an owner. Specified, not yet in the build. **The smallest sufficient model:** heavy synthesis gets a capable model, and staging and routine retrieval get a lean one. Where a judgment is consequential, two independent passes review the same material unseen by each other, and disagreement is escalated and not averaged.

Part 9 · Canon: which version the machine should believe

Give a retrieval system a 2019 directive, a 2024 revision and a draft never issued, and it does not adjudicate. It averages, returning a fluent answer built from all three with no sign they conflict. A human notices the dates and the draft watermark and asks a sergeant. Canon is the written form of that knowledge: the rules about which content wins. A well-organized knowledge base with no automation is underused. Automation on an unmaintained one is dangerous, because every answer inherits its staleness. A fabricated citation is embarrassing and usually caught. A real citation to a superseded directive verifies, looks right to everyone and surfaces as wrong only in a courtroom. Only a precedence rule addresses that, so keeping the corpus current is a named job with an owner.

PROPERTY	WHAT IT MEANS
Named and ordered	One place holds the current version and superseded ones move out. A written precedence decides conflicts: established beats synthesized beats raw, newer usually beats older, a recorded decision beats everything.
Owned and dated	Every article has an owner and a status: established, emerging or speculative. New intelligence is appended under a dated heading and an established article is never overwritten, so the department can say what it believed and acted on at the time.

When sources contradict, the system shows both sources and both dates and does not quietly pick the one it read last. A system that says "these two directives disagree" feels less capable than one that just answers, and it is the reverse. A surfaced conflict is a governance event. A silent one spreads into every downstream answer. **Authority is scoped by subject, not rank.** An executive may rule on how a capability is presented and have no standing on how a detective works a scene. The policy unit is canon on what a directive says, the division commander on how the division is organized, and the detective on how the work is done. A decision record carries the date, the ruling in one clause, who ruled, what it supersedes, a pointer to the reasoning, a stated condition under which it would be judged violated, and the owning domain. When something fails, the failure and the rule that would have prevented it are both recorded. A claims ledger separates what may be said in the present tense from what may not, and a claim with no entry does not ship.

PART 10 · WHAT IT MEANS, AND WHAT IS BUILT

Trust stops being the question. Work that shows itself takes its place.

Trust dominated a decade of procurement and was always unanswerable, because a department cannot verify it. A department can verify whether a system shows its work, cites the rule, routes to the right rank and keeps the record. Those are ordinary engineering requirements. The system runs inside the department's own cloud tenant, so the authorizations already held are the ones that apply. For government workloads that means physically isolated U.S. datacenters, contractual commitments on where data is stored, and access limited to screened U.S. persons. The directive corpus, the checks derived from it and the decision record are the department's assets, and they stay if the vendor is replaced. The judgment stays with the sergeant, the commander and the detective.

SEAT	WHAT IS DIFFERENT ON AN ORDINARY TUESDAY	WHAT IS NOT CLAIMED
Detective	The case is open before assignment, every line sourced, with a visible list of what has not been done. The approval nobody was told about routes and comes back.	It does not solve the case, identify a suspect or rank anyone. It removes the waiting, not the work.
Supervisor	Pending items are visible before the meeting and can be answered from a phone. Doing nothing is recorded as a decision and not as silence.	It does not supervise. It makes what is pending legible.
Commander	One method across divisions, figures that open to records, packets that reproduce months later.	It does not set priorities. It ends the argument over whose figure is right.
Oversight and counsel	Who looked at what and why, answerable without asking the vendor.	It is not a compliance certification. It is the record that makes compliance demonstrable.

> Twenty years ago the information moved on a thumb drive because nothing else could carry it. The information problem got solved. The carrying problem never did. What makes this moment different is that a department's own rules can finally be read by a machine, so one system can hold what is true and what is supposed to happen because of it. A system that labels its facts, cites its rules, refuses to score people and keeps the record is not a constrained version of a more powerful product. It is the only version that survives contact with the work.

BUILT, DEMONSTRABLE ON SYNTHETIC RECORDS

- Case view, reporting, warrants, deconfliction and link analysis.
- The approval chain: routed to a named rank, answerable remotely, releasing what it held.
- Provenance labeling on every fact, with its source.
- Role-based seats, plus inspection and oversight surfaces.
- Directive retrieval against a live corpus. Acquisition and extraction are proven, and no citation appears on screen until retrieval passes a measured threshold.

DESIGNED, NOT CLAIMED AS BUILT

- Dual-layer grading: reliability and credibility as separate dimensions.
- Token allotments for supervisors and officers.
- The officer-safety conjunction rule, escalation ladder and retention window.
- Vehicle sweep statistics, coverage normalization, false-discovery control and cross-case action ranking.
- P-Solve element weights, which are uncalibrated.
- A feed that writes service status to a statewide database.

Sources

Microsoft Learn: RAG in Azure AI Search; What is Azure Government; Shifts; Walkie Talkie; Shared channels; Model Context Protocol; Application design for AI workloads. Samsung public-sector Galaxy Watch documentation (one blog source treated as a lead only). 28 CFR Part 23; FBI CJIS Security Policy; California SB 524; Utah SB 180. Varvarigou and others, BMJ 2014;349:g6534. Soh, Sia and Tay-Yap, Communications of the ACM; van Beijsterveld and van Groenendaal, 2016. Nonaka, Toyama and Konno, Long Range Planning. Gorry and Scott Morton, 1971. Klein, Calderwood and Clinton-Cirocco, 1985. Endsley, 1988 and 1995. Parasuraman and Manzey, 2010; Mosier and others, 1997. Dietvorst, Simmons and Massey, 2015 and 2016. Bansal, 2021; Zhang, Liao and Bellamy, 2020; Buçinca, 2021.

The lineage, system inventory, verification architecture, refusals and canon model come from the author's build record and the governing decision log. No record from any agency system appears in this paper or in any GY6 product. Statutory text changes, so verify against the official source before relying on it in a procurement or filing. PRISM IDSS is GY6.AI property, developed by Justin Frank. Engineered by GY6.AI. Delivered and operated by GY6 Solutions Inc. Public release, November 2026.