

Recognition Before Authority

What AI Should and Should Not Do in the First Minutes of a Working Incident

Dan Deppert

Co-Founder & CTO, HazNetix

October 27, 2026 · Version 1.0

Audience: fire chiefs, command officers, engineers, and public-safety technology leaders. Not a procurement brochure.

Author disclosure: Dan Deppert is Co-Founder & CTO of HazNetix, which is developing AI-assisted, human-supervised hazard-intelligence workflows. This paper does not represent a claim of field-validated operational performance.

In one minute

Idea	Meaning
Recognition before authority	Improve the picture before giving a system the call.
Known / inferred / unknown / unsupported	Keep those four states separate. Unknown is a valid state. Unknown must not silently become inference, and inference must not silently become fact.
Display → Recommend → Act	Present information. Offer a bounded recommendation. Execute a consequential action. Near-term case is Display, then bounded Recommend. We do not recommend letting AI act on strategy now.
Human judgment	Strategy, commitment, and withdrawal stay with the commander.
Earned authority	A capability does not get more authority because it can produce an action. Each step requires evidence.

The important question is not whether artificial intelligence (AI) will appear in hazardous-incident response. The question is what authority it should be given when information is incomplete, time is limited, and the consequences of being wrong are high.

Hazardous incidents do not suffer from a lack of information in the abstract. They suffer from incomplete information arriving from different sources at different times while someone is responsible for a decision.

AI may help organize those fragments and mark what is missing. That is not the same thing as owning the call.

Feed recognition. Do not replace judgment.

This paper does not claim that AI has been proven to save lives, complete a 360 faster, or improve strategy on the fireground. No published field trial shows this yet.

The problem is recognition under incomplete information

The first minutes of a working incident are an information problem.

The officer has dispatch remarks, one visible side, occupancy guesswork, and whatever preplan actually loaded. Size-up is supposed to keep running: what command can see, reconnaissance command cannot see, and knowledge collected before the tones. The 360 is supposed to catch the basement, the void, the side that did not match the front.

National Institute for Occupational Safety and Health (NIOSH) guidance drawn from its line-of-duty-death investigations keeps naming the same hole: incomplete size-up, missed basement or hidden fire, and risk assessments that did not drive a strategy change.¹ Those reports do not prove software would have changed the outcome. They show that facts already on the scene failed to become a command picture in time.

Structure fires remain common. Traumatic fireground deaths are a minority of firefighter fatalities, but they are the subset where size-up, communications, changing conditions, and strategy can become consequential.²

The U.S. Fire Administration's answer is training: its National Fire Academy command course uses simulations to teach officers to recognize critical cues at complex, multi-alarm incidents.³ National Institute of Standards and Technology (NIST)–hosted fire-service guidance makes the same point in engineering language: crews must absorb a great deal of information in a hazardous environment, and volume itself can hinder assessment.⁴

A system that only restates what it already has is solving the wrong problem.

Missing information is itself operational information.

Why the usual answers fall short

The service already bought part of this problem.

Computer-aided dispatch (CAD), mobile data terminals, digital preplans, hydrant layers, and thermal imagers are in the cab. The National Fire Protection Association (NFPA) already has a pre-incident planning standard (NFPA 1620, now part of NFPA 1660), and NIOSH tells departments to put preplan information into the dispatch computer so it reaches every responding unit.⁵ The failure is not the absence of a standard. The failure is delivery: the plan does not always reach the first-due officer in the minute it would change the first report.

If you cannot find the preplan, you cannot use it. Finding and sharing the plan can deliver most of the gain before anyone adds a model. Not all useful software is AI.

A generative paragraph after arrival is late and usually wrong for the job. Experienced commanders do not compare option lists on the bumper. Klein, Calderwood, and Clinton-Cirocco probed 156 decision points from 26 fireground commanders: fewer than 12 percent showed comparison of two or more options; more than 80 percent were recognition of a typical situation and a typical action for that prototype.⁶

Size-up is trying to reach projection — what happens next — from incomplete cues. A tool that dumps unranked facts attacks perception. A tool that surfaces “lightweight attic, no hydrant on this block, prior fire in the B exposure” can feed recognition without forcing a decision tree the commander will not use.

An unfiltered model output is another radio channel. Task saturation already degrades situation awareness. Adding noise is not support.

One operating model: Display → Recommend → Act

Job	What the system does	Who decides
Display	Collects and presents information, including what is missing	Human interprets
Recommend	Analyzes inside a validated envelope (the conditions the tool was tested in) and offers a bounded recommendation	Human decides
Act	Executes a consequential operational action	Human supervises or overrides

Figure 1. Three jobs an AI aid can do, and who decides.

The near-term case is Display, then Recommend only inside a validated envelope. We do not recommend letting AI act on incident strategy now.

NIST Special Publication 1500-29 (NIST SP 1500-29) already lands here, in fire-service language: AI will not replace firefighters or take control of emergency responses; it may help crews make faster, more informed decisions; final decisions remain with the firefighter or incident command.⁴ That document also publishes a 0–4 autonomy ladder; the Appendix maps it to the three jobs above.

Two other lines from the same guidance matter more than the ladder. The guidance expects the first AI safety products to be locked — same input, same output, testable with pass/fail. Adaptive models keep learning after sale and can drift, and the guidance warns that traditional pass/fail tests may not fit them. On flashover prediction the document is blunt: a miss can be fatal; a chronic false alarm trains disbelief.⁴

Format matters as much as domain. In a 2022 study built around mental-health emergency calls, prescriptive advice from a biased model moved emergency resource choices; descriptive flags did not.⁷ Hazardous-incident tools should describe and flag. They should not issue the tactic.

The 2024 Urban Fire Forum position is the chief-level version of the same rule: override, no solutionism, do not replace vital labor, do not rely solely on generative output for decisions that affect people.⁸ Federal Emergency Management Agency (FEMA) listing AI on an authorized equipment list is a buyer fact. It is not a safety validation.⁹

What command officers will accept

This is survey plus named judgment. It is not a field outcome study.

The Center for Public Safety Excellence (CPSE) Center for Innovation scan (July 2025, 156 fire chiefs and key personnel) is the cleanest adoption picture available: most departments already use some AI in the office; almost none use it to run incident command; most still have no AI policy.¹⁰

The operational rule is one rule. Fire Chief Brian Schaeffer, paraphrased: if an AI recommendation leads to a poor outcome, the vendor is not the answer. “There’s only one right answer: The human in charge.” His test, also paraphrased: justify the decision as if the computer were not there.¹¹

Fire Chief Gregory A. McComb drew the office/scene line. AI platforms, he wrote, are best used in “low-stakes, administrative environments. The fireground isn’t one of them.” And: “There’s no algorithm for a smoke condition that doesn’t match a size-up.”¹²

Other command voices land in the same place. In a 2026 CPSE Catalyst podcast episode on the AI scan, Chiefs Brad Brown and Don Lombardi discuss the limits of AI against an incident commander’s “gut feeling.”¹³ This paper takes a narrower position: AI should never be what replaces that judgment. The design consequence: speak when something is wrong; if conditions are normal, do not add a channel.

The tool may flag a missing side, a sensor limit, or a preplan fact. It may not write the strategy, and it may not train the next officer to wait for the strategy to appear.

Known, inferred, unknown, and unsupported

A reasonable engineering hypothesis is not an operational result.

Unknown is a valid state. Unknown must not silently become inference, and inference must not silently become fact.

State	Meaning	First-due example
Known / observed	An inspectable input with a source	Smoke showing, Side Alpha. Officer declared offensive / Alpha attack.
Inferred	A conclusion supported by available evidence, clearly labeled as inference, with confidence and limitations visible	Heat-detector rise in Division 1, confidence bounded, closed-compartment conditions outside the training set.
Unknown	Required information that has not been established	Charlie and Delta not reported. Basement unconfirmed. Occupancy unconfirmed. Preplan age unknown.
Unsupported conclusion	A finding the available evidence does not justify. Prohibited. Must not be presented as fact.	“No basement, because the preplan is silent.” “Occupants are present, because dispatch said possible.” “Fire is confined to Division 2 Alpha,” from the front seat only.

Figure 2. Four information states on first due. Unknown must not silently become inference, and inference must not silently become fact.

Known from doctrine. Size-up starts at dispatch. A 360, or a delegated 360, is part of the initial risk assessment.

Known from harm records. Incomplete size-up and weak risk assessment recur in investigated line-of-duty deaths. Pattern, not a rate of “preventable by software.”¹

Known from how experts decide. Recognition-primed action is the measured habit in Klein’s sample.⁶ Serving chiefs describe the same habit as gut, or show-your-work-as-if-the-computer-were-not-there.¹¹

Known from adoption. Office AI is common. Formal AI policy is uncommon. Command AI is rarer still.¹⁰

Known from measured AI work. The strongest published fireground-adjacent models are sensor-side, mostly simulation. P-Flash predicted flashover about 86 percent of the time at one minute in simulation. FlashNet later reported up to 92.1 percent at 30 seconds across 17 residential floor plans in simulation. In 13 Underwriters Laboratories (UL) live burns, P-Flash worked in open rooms and almost never when the fire started in a bedroom behind a closed door.¹⁴ That gap is the point.

Known from human-factors research. Prescriptive recommendations move people.⁷ Automation bias produces omission and commission errors in experts. Simple training does not wash it out.¹⁵

Inferred — labeled as inference. Putting current preplan facts, 360 completeness, and bounded sensor trends in front of the first-due officer should shorten the time to a complete picture. Testable. Not a result.

Unknown. No published field trial shows that an AI size-up or command aid reduced injury or death. Minute savings are unknown. Whether an explainable display is readable under full personal protective equipment (PPE) and radio load is unknown. Who is at fault when an aid is wrong is, for now, a question for department policy.

Worked example: first-due residential fire

Hypothetical. Not a case study. Not a product demonstration.

0230 hours. Two-story residential. Smoke showing from Side Alpha. Dispatch: possible occupants. The officer can see Alpha and a sliver of Bravo. Charlie is a fence and a grade. Strategy already declared from the front seat: offensive, Alpha attack.

Display on the screen the officer already uses:

- Preplan, if one exists, with date: basement yes or unstated, lightweight attic, no hydrant on this block.
- Completeness: Charlie and Delta not reported. Basement not confirmed. 360 not complete.
- Trend, only if a sensor exists and the model states its limit: heat rise in Division 1, closed-compartment conditions outside the training set.

Not: “Go defensive.” Not: “Make the interior.” Not a flashover countdown the live-burn record says the model misses behind a closed door.¹⁴

Recommend, inside a validated envelope: “360 not complete. Treat Charlie and below-grade as unknown before interior commitment.” That is a recommendation about information completeness. It does not invent a tactic.

If Engine 3 later reports a walk-out and fire on Charlie, provenance updates. Recommend may say the existing Alpha-offensive plan does not account for the new report. Changing strategy is still a human act.

Still human: strategy, interior commitment or withdrawal, search priority, offensive versus defensive.

Logged: each input and source; what was marked unknown; displays; recommendations, bounds, and abstentions; declared strategy; overrides; the Charlie report and its time; any later strategy change.

The officer still walks or delegates the 360. The aid's job is to make the missing side expensive to ignore.

When evidence is insufficient, a visible unknown is safer than a confident unsupported finding.

Abstention must be a supported system state. When the model is outside its validated envelope, it should expose that limitation rather than fill the gap.

Design principles

These are constraints. They are not a second architecture.

Principle	Forced by	Requirement
Evidence before inference	Automation bias; confident miss	Show the observation or record behind a conclusion. If the basis cannot be inspected, do not speak in the voice of a finding.
Completeness awareness	Missed sides; incomplete size-up	Identify what is missing. "Side Charlie not reported" is operational information.
Human authority	Cognitive offloading; office-to-scene bleed	Consequential decisions stay with command. Approval is explicit, not implied by silence.
Visible uncertainty	Stale preplans; task saturation	Confidence, freshness, contradictions, and limits must be visible. Unknown is a valid state.
Bounded recommendation	Closed-door misses; out-of-distribution scenes	Recommend only inside a validated envelope. If the scene does not match, abstain.
Reconstructable history	Adaptive drift; after-action and public record	Preserve inputs, recommendations, abstentions, decisions, overrides, timestamps, and model version. Locked models are the honest first class.
Degraded operation	Failed radios, cameras, cloud, preplans	Command must continue if the AI dies. The system supports command. It is not a prerequisite for command.

"The officer can always override" is a policy sentence. It is not a measured control when the officer is task-saturated and the aid sounds sure. Schaeffer's test, paraphrased: justify the decision as if the computer were not there.¹¹

If a pilot shows the system making command slower, contributing to missed information, or driving a bad prescription against visible conditions, shut it off. A demo video is not evidence.

Write the policy before the tool rides. Assume an open-records request. Occupancy guesses and "likely vacant" labels will not get a free pass.

Earned authority

Capability does not gain authority merely because it can produce an action. Each increase in authority requires evidence.

Stage	Gate before the next stage
Observe	Can we name the inputs and their sources?
Display	Does the officer see provenance and unknowns? Does this beat the current non-AI workflow on missing-side identification?
Analyze	Is the envelope defined? Locked or adaptive? What happens outside it?
Recommend	Overrides, abstention quality, and envelope misses are logged and reviewable.
Validate	Scoreboard below. Kill criteria named in advance.
Bounded Act	Only after the gates above. Act means a consequential operational action — strategy change, crew commitment or withdrawal, an autonomous device into a hazard on its own authority. Fetching a preplan is Display plumbing. Do not hide Act inside a convenience feature.

Figure 3. Authority moves only through an evidence gate.

Pilot scoreboard

Do not start with lives saved. That endpoint is too far from the design and too easy to claim.

Where applicable, these measures should be compared with the current non-AI workflow or established operational baseline.

- 360 completed or explicitly delegated before interior crews commit
- Basement or void status known, or marked unconfirmed
- Time from arrival to a stated strategy
- Time from a material condition change to a revised strategy
- Critical preplan facts surfaced before they mattered, with staleness visible
- Override rate and stated reason
- Abstention rate, and whether abstention was correct later
- Wrong recommendations inside the claimed envelope
- Cognitive load and missed benchmarks — up or down

Those are pilot measures, not product claims.

Grant eligibility is not acceptance testing. Train the aid the way a department trains a thermal imager. If the crew cannot explain when it lies, it does not ride.

Claim boundary

No published field trial shows lives saved, faster 360s, or better strategy. Flashover numbers above are simulation and live-burn evidence, not command trials. Office adoption is not fireground readiness.

Try the tool. Measure the scoreboard. Do not narrate a safety outcome the pilot did not measure.

Close

The near-term opportunity is not autonomous incident command. It is better information delivered earlier, with its source, uncertainty, and limitations intact.

Build that first. Measure whether commanders see missing information earlier, complete the operational picture faster, and reassess when conditions change.

If the evidence supports greater capability, expand deliberately. If it does not, stop.

AI should earn authority through validation. It should not receive authority simply because the technology exists.

Pause here before the next tool decision

This paper is a position, not a product pitch. Before the next AI conversation in your department, stop on these points. Capability without clear limits is how ordinary tools do damage: a system with broad access and a vague goal will invent the rest. Hazardous-incident work cannot afford that pattern.

Reflective prompts:

1. In your current first-due workflow, where would Display help, and where would it get in the way?
2. What can a system actually observe on your incidents, and what would it only be inferring?
3. Where are the hard limits: what must stay human even if the screen looks confident?
4. If the tool is wrong under stress, how would you know, and how would you recover?

If you want a tighter staff discussion after that, use these five:

1. What can the system actually observe, and what is it inferring?
2. What does it do when important information is missing?
3. Which decisions remain explicitly human?
4. What evidence is required before moving from Display to Recommend?
5. Can we reconstruct what the system knew, recommended, and got wrong?

About the author

Dan Deppert is Co-Founder and Chief Technology Officer of HazNetix, where he leads the development of AI-assisted, human-supervised hazard intelligence and 360 size-up workflows for public safety. His professional background spans enterprise technology, technical operations, and customer-facing systems, with an emphasis on reliability, practical implementation, and human oversight. His work focuses on applying emerging technologies to high-consequence environments while preserving operational judgment, accountability, and safety. He also writes about practical applications of artificial intelligence at dandeppert.com. The views expressed in this paper are his own.

Contact and ask

Command this kind of incident? Tell me where this is wrong. Email dan.deppert@haznetix.ai. I am collecting officer feedback for version 1.1 and the pilot scoreboard.

Version and how to cite

Version 1.0, October 27, 2026.

Suggested citation: Dan Deppert, “Recognition Before Authority: What AI Should and Should Not Do in the First Minutes of a Working Incident,” position paper, Version 1.0 (HazNetix, October 27, 2026), <https://dandeppert.com/insights/recognition-before-authority/>.

Appendix. Published ladders mapped to this paper

This paper	NIST SP 1500-29 (simplified)	Product-class language
Display	Level 1 Insight	Class 1 — collect and display
Recommend	Level 2 Collaboration — human vets and acts	Class 2 — sense, analyze, recommend
Act	Level 3 Operation and Level 4 Full autonomy	Class 3 — acts with no human in the loop

NIST Level 0 — AI not used. No AI or autonomy integration in systems or processes. Levels 3–4 and Class 3 stay out of scope for incident strategy.⁴

Endnotes

1. National Institute for Occupational Safety and Health (NIOSH), NIOSH Alert: Preventing Deaths and Injuries of Fire Fighters Using Risk Management Principles at Structure Fires, DHHS (NIOSH) Publication No. 2010-153 (July 2010), <https://doi.org/10.26616/nioshpub2010153>; NIOSH, Workplace Solutions: Preventing Deaths and Injuries of Fire Fighters Working at Basement and Other Below-Grade Fires, DHHS (NIOSH) Publication No. 2018-154 (July 2018), <https://doi.org/10.26616/nioshpub2018154>. Both accessed Oct 2026. Recurring themes in NIOSH Fire Fighter Fatality Investigation and Prevention Program (FACE) reports include incomplete size-up, below-grade or hidden fire, and risk assessments that did not drive strategy change. Individual investigations vary; cited here as a recurring pattern, not as a national preventable-by-software rate.
2. Scale, not causation. National Fire Protection Association Research, Fire Loss in the United States (2024 data), <https://www.nfpa.org/education-and-research/research/nfpa-research/fire-statistical-reports/fire-loss-in-the-united-states>: an estimated 470,500 structure fires, a structure fire every 67 seconds, and 3,170 civilian structure-fire deaths. United States Fire Administration on-duty firefighter fatalities, <https://www.usfa.fema.gov/statistics/reports/firefighters-departments/firefighter-fatalities.html>: 93 (2023), 72 (2024), 81 (2025; provisional, USFA, as of October 16, 2026). Most on-duty deaths are medical. USFA's 2025 provisional summary attributes 27 of 81 to activities related to operating at a fire scene and 48 to cardiovascular events. Traumatic fireground deaths are the minority subset in which size-up and strategy can become consequential. These figures do not measure AI effect. Accessed Oct 2026.
3. United States Fire Administration, National Fire Academy, Command and Control Decision-Making at Multiple Alarm Incidents (R0297), course catalog, <https://apps.usfa.fema.gov/nfacourses/catalog/details/614>. Accessed Oct 2026. A simulation-intensive course for command officers; students are taught to recognize critical cues specific to complex emergency incidents. Training, not a product recommendation.
4. Christopher Sampl et al., Artificial Intelligence in the Fire Service: Considerations for Implementing Artificial Intelligence into Electronic Safety Equipment, NIST Special Publication 1500-29 (June 2025), <https://doi.org/10.6028/NIST.SP.1500-29>. Accessed Oct 2026. Fire-service-informed guidance hosted in the NIST SP 1500 series; not a binding NIST standard and not an official NIST regulatory position. Load-bearing lines used here: AI will not replace firefighters or take control of emergency responses; tools may support faster, more informed decisions; final decisions remain with the firefighter or incident command; locked versus adaptive models; flashover false-negative / false-positive warning; 0–4 autonomy ladder.
5. NFPA 1620, Standard for Pre-Incident Planning, now consolidated into NFPA 1660 (2024 edition). See Nicole Cassels, “What Is the New NFPA 1660?,” NFPA blog, April 25, 2024, <https://www.nfpa.org/news-blogs-and-articles/blogs/2024/04/25/what-is-the-new-nfpa-1660>. Accessed Oct 2026. The NIOSH Alert in endnote 1 (Publication No. 2010-153) tells departments to enter preplan information into the dispatch computer so that “the critical information is provided to all responding units.”
6. Gary Klein, Roberta Calderwood, and Anne Clinton-Cirocco, “Rapid Decision Making on the Fire Ground: The Original Study Plus a Postscript,” *Journal of Cognitive Engineering and Decision Making* 4, no. 3 (2010): 186–209, <https://doi.org/10.1518/155534310X12844000801203>. Accessed Oct 2026. 156 decision points from 26 experienced

- fireground commanders; fewer than 12 percent showed comparison of two or more options; more than 80 percent were recognition of a typical situation and typical action.
7. Hammaad Adam, Aparna Balagopalan, Emily Alsentzer, Fotini Christia, and Marzyeh Ghassemi, “Mitigating the Impact of Biased Artificial Intelligence in Emergency Decision-Making,” *Communications Medicine* 2 (2022), <https://doi.org/10.1038/s43856-022-00214-4>. Accessed Oct 2026. Study setting was mental-health emergency calls (choosing police or medical help), not the fireground. Prescriptive advice from a biased model moved resource choices; descriptive flags did not. Adjacent-domain experimental evidence, not a fireground trial.
 8. Urban Fire Forum / Metropolitan Fire Chiefs, Safe and Ethical Use of Artificial Intelligence (September 20, 2024), <https://ctif.org/sites/default/files/2024-09/2024%20UFF%20Position%20Statement%20-Safe%20and%20Ethical%20Use%20of%20AI-%2020%20Sept%202024.pdf>. Accessed Oct 2026.
 9. Federal Emergency Management Agency, Authorized Equipment List item 04AI-00-AI, <https://www.fema.gov/grants/tools/authorized-equipment-list/04ai-00-ai>. Accessed Oct 2026. Buyer-eligibility fact only.
 10. Center for Public Safety Excellence, Center for Innovation, Strategic Scan 1: Use of and Attitudes towards Artificial Intelligence in the Fire and Emergency Service (surveyed July 2025; published September 2025), <https://www.innovationcpse.org/wp-content/uploads/2025/09/CPSE-Strategic-Scan-1-Artificial-Intelligence-2025.09.03.pdf>. Accessed Oct 2026. n = 156 responses (19.5 percent of 799 distributed) from Commission on Fire Accreditation International (CFAI)–accredited and Chief Fire Officer (CFO)–designated leaders. Approximately 79 percent reported some departmental AI use; 69 percent administrative use; about 18 percent active-incident management or after-action; incident command / fireground accountability / working-incident situational awareness: three to six respondents; formal AI policy: 14 percent yes. A pulse of that leader set, not a census of all U.S. departments.
 11. Brian Schaeffer, “When the Machine Thinks for Us,” *FireRescue1*, December 1, 2025, <https://www.firerescue1.com/artificial-intelligence/when-the-machine-thinks-for-us>. Accessed Oct 2026. Quoted: “There’s only one right answer: The human in charge.” The “justify the decision” test in the body is a paraphrase; his words are “Make the officer justify the decision as if there were no computer involved.”
 12. Gregory A. McComb, “Are We Training Firefighters or AI Dependents?,” *Firehouse*, October 3, 2025, <https://www.firehouse.com/technology/artificial-intelligence/article/55317276/are-we-training-firefighters-or-ai-dependents-a-fire-chiefs-warning-on-the-risks-of-overrelying-on-chatgpt>. Accessed Oct 2026. Quoted: “The fireground isn’t one of them.” and “There’s no algorithm for a smoke condition that doesn’t match a size-up.”
 13. CPSE Center for Innovation, Catalyst podcast, Episode 2, “Fire Service and AI: ‘Trust but Verify,’” with Fire Chief Brad Brown (Grand Rapids, Mich., Fire Department) and Fire Chief (ret.) Don Lombardi (West Metro Fire Rescue, Colo.), May 12, 2026; episode page, *FireRescue1*, May 15, 2026, <https://www.firerescue1.com/artificial-intelligence/cpse-catalyst-podcast-detailing-the-ai-focused-strategic-scan-ep-2>. Accessed Oct 2026. The episode description states that the chiefs discuss how AI “will never be able to replace incident command’s ‘gut feeling.’” The “speak when something is wrong” rule in the body is the author’s design conclusion.
 14. Simulation and live-burn evidence, not command trials. NIST reporting on P-Flash, “How AI Could Alert Firefighters of Imminent Danger” (June 2021), <https://www.nist.gov/news-events/news/2021/06/how-ai-could-alert-firefighters-imminent-danger>: the model correctly predicted flashovers one minute beforehand for about 86 percent of simulated fires; in 13 UL full-scale house-fire experiments it performed well on open-area starts and almost never when the fire started in a bedroom behind a closed door. Wai Cheong Tam et al., “A Spatial Temporal Graph Neural Network Model for Predicting Flashover in Arbitrary Building Floorplans,” *Engineering Applications of Artificial Intelligence* 115 (2022), <https://doi.org/10.1016/j.engappai.2022.105258>: FlashNet predicted flashover 30 seconds preceding occurrence with overall accuracy of about 92.1 percent across 17 residential floor plans in simulation. NIST SP 1500-29 (endnote 4) discusses AI flashover prediction in general terms; it does not evaluate these models. Accessed Oct 2026.
 15. Raja Parasuraman and Dietrich H. Manzey, “Complacency and Bias in Human Use of Automation: An Attentional Integration,” *Human Factors* 52, no. 3 (2010): 381–410, <https://doi.org/10.1177/0018720810376055>. Accessed Oct 2026. Automation bias as omission and commission errors; simple training does not reliably eliminate the effect.