Cyber Case Study Program

The 72-hour ghost: AI vs the ticking clock

Case Number ACI-08-2026

Dr. Khushi Gupta

Editor in Chief: Karen Guttierri, PhD
Lead Editor: Volker Franke, PhD
Managing Editor: Anne Chance, PhD

About Case Studies

The ACI-TRENDS Global Cybersecurity Case Program is administered by the Army Cyber Institute at the United States Military Academy at West Point and publishes cybersecurity-related teaching cases, simulations and interactive exercises for use in academic and professional classrooms.

What are case studies?

Case studies are high-impact interactive learning tools structured around real or realistically simulated events placing learners in the role of decisionmakers confronting complex problems, tradeoffs, and uncertainty. Case studies immerse participants in the problem and its dilemmas.

Why use case studies?

Immersing participants cognitively and emotionally in this way requires them to grapple with ambiguity, incomplete information, and competing priorities while developing plausible courses of action.

Case studies are particularly important because cyber incidents unfold across technical, organizational, legal, and human domains simultaneously.

Case studies make invisible systems and cascading consequences visible, demonstrate how theory and policy apply under pressure, and build the analytical judgment required to anticipate, assess, and respond to real-world cyber threats.

Where to find ACI case study series.

More information on the case method can be found:

1. At the TRENDS Website: https://trendsglobal.org/whats-trending/
2. Decision-Making under Uncertainty: Using Case Studies for Teaching Strategy in Complex Environments by Volker Franke, PhD. https://jmss.org/article/view/57957
3. **[If there is a link to an ACI or WPP publication]**

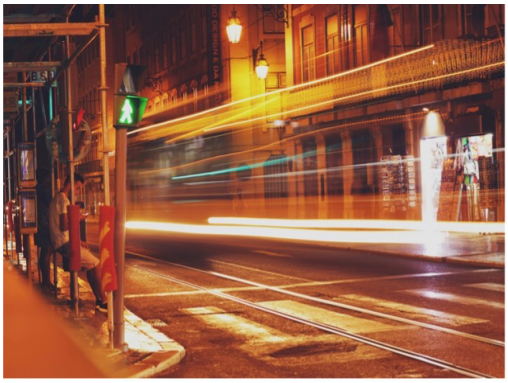
About the Author

Dr. Khushi Gupta is an Assistant Professor of Cybersecurity at the University of North Georgia. She holds a Ph.D. in Digital and Cyber Forensic Science and has authored more than 30 publications, contributing impactful research in digital forensics, malware analysis, and secure learning systems. She is a speaker, delivering talks at BSides Atlanta, the Official Cybersecurity Summit, and the Ransomware Cybersecurity Summit. Dr. Gupta also brings hands-on investigative experience from law enforcement and corporate forensics, supporting complex digital evidence analysis across criminal and civil matters. Committed to community advancement, she mentors students through organizations such as WiCyS, Women4Cyber, and BBWIC, helping aspiring professionals break into cybersecurity. She also serves on an IEEE committee, as Secretary for the P2834.1 Digital Forensics Standardization Group.

The 72-hour ghost: AI vs the ticking clock

Section 1: The Crisis

Background



It is Tuesday, March 14, 2028. The Philadelphia Police Department (PPD) faces a crisis that feels like a countdown. At 7:03 a.m., every high-ranking city official receives the same encrypted email:

"A bomb will detonate at the City Hall Centennial Gala in 72 hours. This is not a request. It is a judgment."

The leak of officials' private schedules and security codes suggests a massive internal data breach. Forensic analysts confirm that the message originated from a rotating sequence of burner accounts routed through Tor, and every email is signed by the same pseudonym: "The Ghost."

The PPD is now in a race against time to prevent a mass-casualty event.

The Investigation: The Digital Dead End & The Corporate "Haystack"

Detective Maria Karras's task force is the city's only hope.

1. **The Digital Dead End:** The source is forensically impossible to find. Every threat email was sent from different encrypted burner email accounts and routed through Tor exit nodes.[1] The "who" and "where" are dead ends.
2. **The "Ghost" Author:** The "what" is all they have. The plaintext of the 50+ threat emails shows a consistent, unique linguistic fingerprint, clearly the work of a single "Ghostwriter."

This "haystack" data is a 1.5-terabyte dataset...

The investigation focuses on the source of the data breach. Suspicion falls on the City Planning Department, through which all 4,000 city employees with access to the leaked data pass. Karras's team legally acquires the entire 5-year email archive for this department. This "haystack" data is a 1.5-terabyte dataset containing over 20 million "sent" emails from 4,000+ current and former employees.

Karras stares at the wall of monitors in the cyber lab: lines of decrypted text, metadata traces, and timestamps. Every lead ends the same way, an empty relay server somewhere in another country. Her team has less than three days, and the

1 A Tor exit node is the final relay in the Tor network through which traffic passes before reaching its destination. See: whatismyip.com/tor-exit-node

only thing bigger than the dataset is the silence inside it.

The pressure is mounting from every direction. The Police Commissioner calls Karras directly: “The Mayor wants answers by morning. The press already has the story. What do we have?” She has nothing she can give him. The District Attorney’s office sends a liaison who warns that any evidence obtained through unproven methods could collapse in court. Meanwhile, the Department of Homeland Security is pressing for a joint operation, insisting the PPD share all raw data with federal analysts, a move that would effectively hand over control of the investigation.

“Ma’am,” an analyst says, “we could run it through PatternFinder.”

Karras knows what that means: turning over 1.5 terabytes of sensitive government emails to a proprietary AI-powered corpus comparison tool.

The clock on the wall reads 47:22:03. Manual analysis of 20 million emails is a guaranteed failure. AI analysis might be an answer, or it might be a new security breach and a courtroom disaster. The Commissioner is waiting. The D.A. is warning. The public does not yet know there is a bomb.

AI analysis might be an answer, or it might be a new security breach...

What should Detective Karras decide?

While Karras weighs her decision, the physical response intensifies. Bomb squads sweep the City Hall complex daily but find nothing. The Fire Department establishes a 10-block evacuation plan, and the city’s Emergency Management Unit coordinates with Homeland Security to secure the gala site. Surveillance drones monitor the perimeter, and local hospitals activate mass-casualty standby protocols. Every negative bomb sweep heightens tension. If the digital investigation fails, there is no fallback.

Section 2: The bomb and the black box

Detective Karras’s team feeds the digital forensic tool the target documents, the plaintext of the 50+ “Ghostwriter” threat emails, and the known corpus: the 1.5 TB “haystack” of over 20 million “sent” emails from the 4,000+ employees.

After 54 hours of processing, with now less than 24 hours to go, the tool generates a single, high-confidence report: “The linguistic fingerprint of the target documents shows a 94.7% probability match with the writings of Employee ID: 3491.” The employee ID corresponds to Arthur Vance.

“The linguistic fingerprint of the target documents shows a 94.7% probability match...”

Karras calls an emergency briefing. The reactions around the table expose every fault line in the investigation:

The **Police Commissioner**
The **District Attorney's liaison**
The **AI vendor's technical representative**
The department's **Chief Information Security Officer**

Karras feels a jolt of adrenaline mixed with dread. Arthur Vance. She remembers his name from a protest case, a loud, abrasive, but brilliant activist. He is a plausible suspect. But is he a 94.7% plausible one? The number came from a system she cannot inspect, trained on data she cannot review, using methods its own creators will not disclose. The clock on her wall now reads 23:14:09.

She has been told, "It's your call." The Commissioner wants a raid. The D.A. wants corroboration that does not exist. The vendor cannot, or will not, explain the result. And somewhere in the city, a bomb is waiting.

What should Detective Karras decide? Should she seek a warrant based solely on the AI's output, or demand manual corroboration she may not have time to complete?

The Aftermath

Six weeks after the "72-Hour Ghost" crisis, the city is safe. The bomb was neutralized with minutes to spare.

Based on the AI-powered digital forensics tool's output, a judge signed the warrant. The raid on Arthur Vance's apartment uncovered all the components for the bomb, a detailed manifesto matching the threats, and digital evidence of him planning the attack. The arrest was made. The gala proceeded without incident.

The city's relief is short-lived. Vance's defense attorney files a "Motion to Suppress," arguing that all evidence found in the apartment should be thrown out of court.

The Defense's Argument: Fruit of the Poisonous Tree

The defense argues the warrant was invalid. The only probable cause was the "opinion" of a "black box" AI.[2]

2 https://www.law.cornell.edu/wex/fruit_of_the_poisonous_tree

The defense contends that this violates the defendant's Sixth Amendment right to "confront his accuser. "

The Prosecution's Counterargument

The prosecution responds that Detective Karras is the expert witness, the AI was her instrument, no different from a forensic database or a ballistics match. They argue that the "exigent circumstances" of an imminent mass-casualty event justified the investigative method and that suppressing the evidence would mean releasing a confessed bomber on a procedural technicality. The physical evidence, the bomb itself, the manifesto, and the digital plans independently confirm Vance's guilt.

The Judge's Dilemma

The Trial Judge must now rule on the Motion to Suppress. If the evidence is admitted, this case becomes the first precedent establishing that opaque AI output can serve as the sole basis for probable cause, opening the door for every law enforcement agency in the country to use black-box tools without disclosure. If the evidence is suppressed, a confessed bomber with a seized weapon walks free, and investigators nationwide lose the ability to use the most powerful forensic tools available in time-critical cases.

If the evidence is admitted, this case becomes the first precedent...

The courtroom is silent. The judge turns to the prosecution: "Counsel, help me understand. You are asking this court to accept the opinion of a system that neither you, nor the detective, nor anyone in this room can explain. On what basis should I distinguish that from asking me to accept a guess?"

You are asking this court to accept the opinion of a system that neither you, nor the detective, nor anyone in this room can explain.

How should the judge rule? What precedent should this case set for the future of AI-generated evidence in criminal proceedings?