THE CYBER DEFENSE REVIEW
2026, VOL. 11, NO. 2, 169-174
https://doi.org/10.55682/cdr/g930-7z5b

SENIOR LEADER PERSPECTIVE

# Defense in a Denser Fog of War

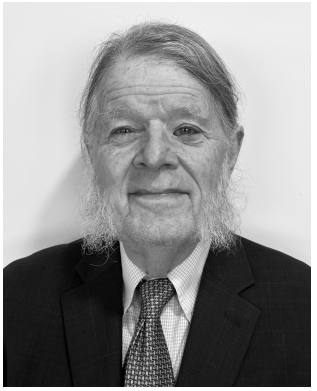
Daniel E. Geer, Sc.D.*

Senior Fellow, In-Q-Tel, Arlington, VA, USA

*Technical innovation throws off side effects. This essay argues that as machine learning overlays a densely interconnected cyberspace, the most telling side effect is indeterminism, the conversion of cyberspace from a deterministic setting (outcomes can be traced to demonstrable causes) to an indeterministic one (outcomes cannot be traced to demonstrable causes). Just as offense rewards creativity, defense rewards discipline, but discipline is based on that deterministic traceability of outcome back to cause. We have long suspected that cyber offense had an inherent strategic advantage over cyber defense; perhaps now it will be proven, not merely suspected. Nevertheless and heretofore, the thorniest problems have fallen to infosec people to fix, and there is no reason that will not be the case here: An indeterministic cyberspace is a dangerous one; can infosec people step up to the challenge?*

**Keywords**: cybersecurity; indeterminism; complexity; artificial intelligence; software vulnerabilities; trust; autonomous systems; cyber defense; information security

* Corresponding author: dan@geer.org



**Daniel E. Geer, Sc.D.**, Senior Fellow, In-Q-Tel. Milestones: The X Window System and Kerberos (1988), the first information security consulting firm on Wall Street (1992), convenor of the first academic conference on mobile computing (1993), convenor of the first academic conference on electronic commerce (1995), the "Risk Management is Where the Money Is" speech that changed the focus of security (1998), the Presidency of USENIX Association (2000), the first call for the eclipse of authentication by accountability (2002), principal author of "Cyberinsecurity: The Cost of Monopoly" (2003), co-founder of SecurityMetrics.Org (2004), convener of MetriCon (2006-present), author of "Economics & Strategies of Data Security" (2008), and author of "Cybersecurity & National Policy" (2010). Creator of the Index of Cyber Security (2011) and the Cyber Security Decision Market (2012). Lifetime Achievement Award, USENIX Association, (2011). Expert for NSA Science of Security award (2013-present). Cybersecurity Hall of Fame (2016) and ISSA Hall of Fame (2019). Testified five times before Congress.

Perhaps we should start with definitions, or to be more precise, a choice of definitions since most important words have alternatives. As always, there may be some bias in my choices.

The word "defense" comes, of course, from defend, which itself appeared in English circa 1300, having come from Latin by way of French.[1]

Defense is naturally paired with "offense", but not merely as one or the other of a duo of opposites, but rather that the two terms serve as selectors for strategies; defense implies a reactive strategy (where players anticipate the opponent's moves so as to react accordingly) while offense implies a proactive strategy (where players create opportunities, seizing openings in the opponent's defense). Where offense rewards creativity, defense rewards discipline.

None of that is news, including the central importance of the two terms in any discussion of cybersecurity. Here, we take the side of defense and defenders whose goal is to achieve and maintain a state of security. My definition of a state of security is the absence of unmitigable surprise, not no surprise, but no surprise that does not have a ready mitigation. In turn, this means that the goal of defensive security engineering is no silent failure, not no failure but no silent failure—you cannot mitigate (react to) a failure you do not know is happening. No silent failure is a straightforward problem in principle, but not in practice.

The wellspring of risk is dependence. You are not at risk from something unless you're dependent on it, a dependence that can be physical (locks) or metaphorical (freedom). Dependence is transitive, thus so is risk; that you may not yourself depend on something directly does not mean that you do not depend on it indirectly and perhaps invisibly. We call the transitive reach of dependence interdependence, which is to say correlated risk. In society today, the biggest locus of correlated risk is that of the Internet (except perhaps that of climate, which we ignore for this discussion). Our concern then is unacknowledged correlated risk. The unacknowledged correlated risk of cyberspace is why it is capable of black swan behavior—unacknowledged correlations contribute, by definition, to heavy tails in the probability distributions of possible events. Complexity invariably hides interdependence(s),

1. Origin and history of defense: https://www.etymonline.com/word/defense

and it is by that path that complexity becomes the enemy of security, the generator of surprises that do not have a ready mitigation.

The growing complexity of cyberspace is no surprise. That growing complexity is the subject of many worthy papers; here, we simply take it as a given. With the broadening deployment of machine intelligence, a new vector of complexity generation appears, that of indeterminism in and around cyberspace. A deterministic world is one where every effect has a demonstrable cause. An indeterministic world is one where some effects have no demonstrable cause. On the one hand, software designed to meet a spec will be deterministic and purposefully so. On the other hand, the stochastic character of machine intelligence makes its operation, as well as its outputs, indeterminate.

Software systems are always modular, often to a remarkable degree. This is not news. That modules get imported into systems from external pools is not news either. Nor is it news that not every module is a perfect fit in all its details, and some of those details will be unfortunate. In that wealth of composability those unfortunate details include what we ordinarily call vulnerabilities, and we know that vulnerabilities cluster at interfaces. Composing modules involves the propagation of trust by assumption, yet we know that security as a property is not composable. What if we were to characterize a vulnerability as that which proves assumed trust to have been misplaced? That's consistent with the zero-trust paradigm, but zero-trust is about logical subsystems in operation, not individual software modules in construction.

When Bratus and Shubina (2017) began to formalize the "weird machine" concept, they wrote:

> "Exploits [are] understood both as proofs-by-construction of the unexpected computation's existence (a.k.a. vulnerability) and as programs in their own right. But if exploits are programs, and generalize into programming models, what machines do they run on? Tautologically, exploits violate the expectations of the original programmer or designer of the vulnerable software, and hence their models of the target machine. These programmers or designers failed to see a richer machine embedded or emergent in the target. Exploits run on these richer machines and are proof-by-construction that these richer machines exist."

What Bratus and Shubina describe is a surprise transition from a thought-to-be deterministic system that was trustable to an indeterminate one that no longer is trustable, except that the system was indeterminate in the beginning, only invisibly so—it is our understanding that transitioned, not the code. Meanwhile, Dave Aitel reminds us that offense rewards creativity, writing that "The limiting factor in offensive capability is not finding vulnerabilities, it is having the talent to turn them into dependable tools". Is turning vulnerabilities into dependable tools (emphasis on dependable) what Bratus and Shubina were implying with their concept of "a richer machine embedded or emergent in the target"? Does finding that richer, weirder machine mean turning something indeterminate into something determinate? Sounil Yu [2] offers a table summarizing the idea:

2. S. Yu, personal communication

Table 1. From Perceived Determinism to Exploitation Mastery

| Stage | What happens | What changes |
|---|---|---|
| **Before Discovery** | The system appears deterministic, safe, and trustworthy. | Nothing changes in the system itself; rather, our understanding remains naïve and incomplete. |
| **Discovery of Vulnerability** | The system reveals unintended or unstable states that behave unpredictably. | Our understanding changes: the code was always indeterministic, but we mistakenly believed it to be deterministic. |
| **Exploitation Mastery** | Skilled attackers learn to manipulate and reliably reproduce the unintended behavior. | The indeterministic becomes operationally deterministic again, but now in the attacker's favor. |

To repeat, indeterminism is a component of, and a contributor to, complexity. Indeterminism is thus the enemy of security, a source of silent failure, a creator of mitigation needs you didn't know you had. What, then, are some sources of indeterminism?

- Per George Dyson (1997), "Emergent behavior is that which cannot be predicted through analysis at any level simpler than that of the system as a whole. Emergent behavior, by definition, is what's left after everything else has been explained." Emergent behavior, in other words, lacks a discernible cause and is categorically indeterminate.
- Larger systems are more prone to emergent behavior than smaller ones as their potential state-table, reflecting the number of possible component interactions, grows faster than linearly, so scale is another source of indeterminism.
- With Glenn Gaffney, we contended that autonomy inevitably produces emergent behaviors, which implies that trust-but-verify escapes our grasp as complexity rises (Geer and Gaffney 2023). Later with Aitel, we wrote that "We're now entering territory where even the threat models themselves become opaque. It's not just 'Why did the AI fix that bug?' It's "Why does the AI think that's a bug in the first place?' In short, autonomy means losing the plot. Humans are about to find themselves on the outside looking in, facing security decisions made by AIs whose reasoning isn't just complicated—it's fundamentally inaccessible." (Aitel and Geer 2025)
- The interaction of system components with unaligned goals contributes to indeterminism, yet code reuse is the goal of the open-source world (which is now dominant: some estimates run to 90+% of all deployed software uses open-source components - Nagle et al. (2022)). So score ensemble authorship as another contributor to indeterminism.
- Stochastic goal seeking leads to non-reproducibility of results, and machine learning is exactly that in operation. Side effects of how machine learning models work in practice, such as order of computation, batching, context size, retrieval techniques, and so forth contribute to machine learning's non-reproducibility and thus contribute to indeterminism.
- Repeating from above, vulnerabilities cluster at interfaces and defense rewards discipline. Collectively, we have only just managed to formulate a regime for controlling the effects of untrusted inputs under a discipline consistent with the classic computer science theories of formal languages and computational complexity, yet we now must deal with a new concept of computing that is driven by the natural languages of prompts that are essentially ambiguous, acting on execution agents with semantics that are unexplainable. Indeterminism arrives in yet other forms.

- As it now appears, an AI "programmer" can write a clean version of a system on demand and in a flash. As such, buying an expensive software license or access to an expensive software-as-a-service is less likely to occur if seemingly the same functionality can be had by generating the computational software when desired. While it is early to draw conclusions, we already have a huge technical debt due to abandonware (Geer 2013), yet a drive to zero the skills needed to generate and deploy systems seems fated to increase the fraction of software running on the Internet that has no meaningful provenance (Geer 2016). This surely overlaps with the indeterministic impact of AI entities that have both autonomy and the capacity for self-reproduction.
- If the explainability problem in artificial intelligence is as unsolvable as it seems, it won't matter to any human that is in the loop—an effective indeterminism will be a constant.

There are more; this is but a flavor. [3]

And so we come to the point this essay; over and over risks have been created that, in the end, security people had to be the ones to solve. Take using a credit card over the Internet. What industry made this happen and become a norm? The porn industry. Early online porn sites went to hackers (infosec people) to make it "work" and the infosec people met the challenge (did they ever). Take cloud computing, once considered insane by hardened infosec professionals, but eventually it was the infosec people who stepped up to make it work (again, did they ever). Take something as pervasive as the web-programming tool PHP; it wasn't designed, it just grew. It was one of the many examples of a software artifact that began as a simple prototype and escaped into common use before it even had a spec. Who had to deal with (and arguably fix) PHP's myriad defects? Infosec people, that's who.[4]

You get the idea. Ugly aspects of things that just grew ended up being "fixed" by infosec people. The ugly aspects of moving fast and breaking things have to get unbroken somewhere by somebody. AI is pushing indeterminism into damned near everything. Does this have to be fixed? If it does, it is going to fall to infosec people to do so, if they can. There's nobody else.

And what if infosec people take up the burden? Perhaps indeterminism is fixable; then, like other security risks, there will have to be mitigations somehow crafted plus reliable mechanisms that, somehow, preclude silent failure.

Or maybe it is too hard. Maybe it is not fixable. If we act as if it is unfixable, then we are into political theory swamps. As indeterminism piles up, society will, somehow, learn to ignore that indeterminism to notable effect. Quoting an article by commentator C.J. Riley (2025),

> "As theorist Karl Popper once noted, humans prefer 'intentional explanations of otherwise random events' because they preserve a sense of agency – someone is steering, even if steering toward harm... Political theorists from James Madison to Hannah Arendt warned that confusion erodes the public's ability to hold leaders accountable. Arendt argued that modern authoritarian tendencies thrive not on total lies, but on 'a mixture of truth and falsehood so scrambled the public ceases to trust itself'. A disoriented citizenry is not necessarily oppressed – just manageable."

3. In the spirit of "It is an ill wind that blows nobody any good", a long-time colleague suggests, "A salutary benefit of AI-generated code may be that it will slow down and maybe even freeze APIs and language evolution." A worthy conjecture to keep in mind.

4. M. Loveless, personal communication

In a beautiful one-liner, Skip Gates explained the propensity to believe in conspiracy theories: "Conspiracy theories are an irresistible labor-saving device in the face of complexity."

That's the challenge, the cliff. It is what cultural critic Neil Postman (1985) was getting at in his book, *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*, that a particular medium can only sustain a particular level of ideas. Postman calls the challenge the "information-action ratio", which he defined as "the relationship between a piece of information and what action, if any, a consumer of that information might reasonably be expected to take once learning it." Though Postman was focusing on the effect of television, he captured the essence thus: "The tie between information and action has been severed. Information is now a commodity that can be bought and sold, or used as a form of entertainment, or worn like a garment to enhance one's status. It comes indiscriminately, directed at no one in particular, disconnected from usefulness; we are glutted with information, drowning in information, have no control over it, don't know what to do with it." An Internet full of indeterminism, fertilizing an adaptive tolerance of conspiracy theorizing, is only more so, the analysis of which Shoshana Zuboff (2018) extended to the commercial Internet.

That is our challenge as infosec people, can we turn down the indeterminism meter and make it stick? Can we defend against its effects? Where do we start?

## REFERENCES


Aitel, Dave, and Daniel Geer. 2025. "AI and Secure Code Generation." *Lawfare* (June 27, 2025). https://www.lawfaremedia.org/article/ai-and-secure-code-generation.

Bratus, Sergey, and Anna Shubina. 2017. "Exploitation as Code Reuse: On the Need of Formalization." *Information Technology* 2:93–100. https://doi.org/10.1515/itit-2016-0038.

Dyson, George. 1997. *Darwin Among the Machines: The Evolution of Global Intelligence.* Basic Books.

Geer, Daniel. 2013. "On Abandonment." July/August 2013, *IEEE Security & Privacy* 11:96. https://doi.org/10.1109/MSP.2013.92. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6573309.

Geer, Daniel. 2016. "Provenance." March/April 2016, *IEEE Security & Privacy* 14 (2): 88. https://doi.org/10.1109/MSP.2016.34. http://geer.tinho.net/ieee/ieee.sp.geer.1603.pdf.

Geer, Daniel, and Glenn Gaffney. 2023. "Establishing the Conditions of Engagement with Machines." *The Cyber Defense Review* 8 (1): 15–28. https://www.jstor.org/stable/48730569.

Nagle, Frank, James Dana, Jennifer Hoffman, Steven Randazzo, and Yanuo Zhou. 2022. *Census II of Free and Open Source Software Application Libraries.* Linux Foundation. https://www.linuxfoundation.org/hubfs/Research%20Reports/lfr_harvard_censusII_mar2022_042824b.pdf.

Postman, Neil. 1985. *Amusing Ourselves to Death: Public Discourse in the Age of Show Business.* Viking.

Riley, C. J. 2025. "The Force-Multiplier Hypothesis: Why Government Actions Make Certain Conspiracies Plausible," November 25, 2025. https://cjrileybooks.com/2025/11/25/the-force-multiplier-hypothesis-why-government-actions-making-certain-conspiracies-plausible/.

Zuboff, Shoshana. 2018. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* Profile Books.