

ISA STUDY OF AI DEPLOYMENT IN CRITICAL INFRASTRUCTURE IDENTIFIES GAPS

By Larry Clinton, President Internet Security Alliance

This is the third blog summarizing the initial results of the Internet Security Alliance (ISA) study entitled “Defining an Effective and Sustainable Model for AI Deployment in Critical Infrastructure.”

The study is based on reports of CISOs from five critical sectors – defense, energy, health, financial services and IT. The first phase of the study identified security challenges and gaps identified by the CISOs related to AI adoption in these five critical sectors.

These posts will outline the results of the first two phases in ISA’s program. Full results of phases one and two in this study will be presented in a congressional briefing on September second.

In the first phase of the study CISOs identified the challenges and gaps they were already experiencing within their sectors. In the second phase the individual sector specific reports were fed into AI tools which identified eight specific areas wherein all five sectors noted virtually identical gaps and challenges.

The cross sectoral challenges suggest the need to fundamentally rethink and reform the policy making process which currently is based on largely industrial age distinctions between sectors which AI may obliterate. This third report attempts to pinpoint policy gaps identified in the AI analysis

Summarizing the Convergent security challenges AI presents to critical infrastructure

1. Adoption is outrunning governance in every sector, including at the most capable organizations.
2. Agentic AI presents a different governance problem from the AI these sectors have managed before.
3. The risk sits primarily in the operating envelope around the model rather than in the model itself.
4. Periodic compliance no longer describes the systems being governed.

5. The timing of defense has inverted. Attackers now operate at a speed human review cycles cannot match.
6. The threat is industrialized rather than exotic.
7. Smaller entities in every sector cannot carry this burden alone, and several perform functions of national consequence
8. All five sectors prefer risk-tiered, incentive-based approaches to uniform prescriptive mandate.

Policy Gaps that must be filled

The convergent findings above describe what the sectors are experiencing. The gaps below describe what does not yet exist and must be built before an effective and economically sustainable model is possible.

1. There is no common vocabulary or risk taxonomy for AI systems. Sectors, agencies, and vendors classify the same system differently, which prevents consistent requirements, comparable evidence, and meaningful information sharing.
2. There is no lifecycle standard for AI systems and autonomous agents. The weakest points are discovery, since AI increasingly arrives embedded in purchased products rather than as a deliberate deployment decision, and retirement, since agents commonly retain credentials, access to tools, and data connections after their purpose has ended.
3. No assurance model operates continuously. Beyond that, there is a genuine technical limit: reconstructing what a system knew and how it was configured at the time of a past decision is not reliably achievable today, which matters for audit, forensics, and liability.
4. There is no legal footing or reciprocity for sharing AI-specific threat information. Existing protections were written before prompt manipulation, model poisoning, malicious agents, and synthetic media indicators existed as categories. Organizations will not contribute information they cannot protect, and they will not continue contributing to a channel that yields no operational value.
5. There is no provenance standard for AI components and no visibility into concentration. Organizations cannot establish the origin, testing history, or modification history of the models and components in their stack, and no

mechanism exists to observe how heavily sectors now depend on the same small set of models, cloud platforms, and data sources, or what a correlated failure would look like.

6. There is no authority or liability framework for machine-speed defensive action. No government has defined rules of engagement, permissible autonomy, or accountability boundaries for automated defense, which leaves operators exposed either way, whether they act or wait.
7. There is no economic mechanism to bring under-resourced essential entities to a defensible baseline. Requirements written for well-resourced enterprises can push essential smaller providers out of a market entirely, which worsens both national resilience and concentration.
8. There is no verification capacity. Nearly every proposed incentive relies on independent assessment rather than self-attestation, and an assessor market of sufficient scale and competence does not currently exist.

These eight items share an important property. None of them can be closed by any individual company, however capable or well-resourced, because each is a deficiency in shared infrastructure such as standards, law, threat exchange, and economics. That property is what makes them public policy problems rather than management problems, and it is the reason the Social Contract framework applies.

In our next post we will outline how a social contract models, as opposed to the traditional model based on industrial age assumptions and governance techniques, is a superior fit for the challenges of the AI Age.