

Joint Intelligence

The Science of Resilient, Human-Centered Security

This talk presents a psychology-grounded framework for building resilient, adaptable security and IT-audit systems that integrate human and artificial intelligence, aligned with how people think, decide, and work. When organizations encode judgment into rules, automate practitioners out of the loop, and reduce risk to metrics, the resulting controls hold under expected conditions but break under the novel ones that inevitably arise over time. This approach often produces unintended consequences — workarounds, blind spots, and new risks the controls themselves create. Drawing on my research into how people and AI work together, the talk explains why resilience depends on how well a system is suited to human cognition and shows how to redesign security so people and technology reinforce each other. Central to that design is a division of labor: algorithms supply scale, pattern detection, and tireless vigilance, while people supply what no algorithm can — deciding what is relevant, framing the right question, judging meaning, and remaining accountable.

The talk also covers surrogation, where a measure quietly replaces the goal it serves; the difficulty of quantifying qualitative risk; and how to audit AI that behaves as an emergent whole, where its stated reasoning may not reflect its real reasoning. Grounded in resilience engineering, it shows participants how to keep human judgment central as systems automate, measure what genuinely matters, and assure systems that cannot be fully inspected. It closes on a principle at its heart: designed with intent, AI can amplify human individuality rather than flatten it.

Major Subjects

The Brittleness Trap

Why rule-based controls and full automation create fragility in complex environments, and the category error of treating complex problems as merely complicated.

The Human Aperture

Relevance, problem framing, and the judgment that must precede any algorithmic analysis — including the recognition of opportunity, not just threat.

Complementarity and Calibrated Trust

Where machines outperform humans, where humans remain essential, and the psychology of algorithm aversion and over-reliance that derails the partnership.

Meaningful Human Control

Designing override, modifiability, and human-on-the-loop authority so automation strengthens judgment instead of eroding it.

The Measurement Problem

Surrogation, Goodhart's and Campbell's laws, and the McNamara fallacy applied to security and risk metrics — and how to decide when results justify action.

Auditing the Black Box

Assurance over emergent, opaque systems: the faithfulness problem, testing reasoning in steps, white-box versus black-box access, continuous monitoring, and the NIST AI RMF and ISO/IEC 42001 standards.

Designing for Resilience and Individuality

The four cornerstones of resilient systems — anticipate, monitor, respond, learn — joint cognitive system design, and building AI that honors human individuality rather than flattening it.

Learning Objectives

By the end of this talk, participants will be able to:

1. Distinguish complex from merely complicated security and risk problems, and recognize when rule-based controls or automation introduce brittleness rather than resilience.
2. Identify the cognitive contributions that remain irreducibly human — framing the question, realizing relevance, and judging meaning — and allocate them deliberately alongside algorithmic strengths.
3. Recognize algorithm aversion and over-reliance in audit and security decisions, and apply meaningful-human-control design principles, including override and modifiability, to calibrate trust.
4. Evaluate metrics for surrogation and the gap between a measure and the construct it represents, and determine when results justify action.
5. Apply a process-level approach to auditing opaque and emergent AI systems — testing reasoning in steps, assessing explanation faithfulness, and using continuous, behavioral assurance methods aligned with the NIST AI RMF and ISO/IEC 42001.
6. Design joint human–AI systems that build resilience through anticipation, monitoring, response, and learning while amplifying individual differences rather than averaging them away.

Program Details

Program Level: Intermediate

Prerequisites: Familiarity with IT audit, information security, or risk management; no data-science background required.

Advanced Preparation: None

Available Lengths: 1, 2, 3, or 4 hours

Instructional Delivery Method: Group Live / Group Internet-Based

Designed For: IT auditors, information security professionals, CISA and CISM holders, GRC and compliance officers, security and risk managers, information systems auditors, and CISOs.

Field of Study: Information Technology

© Toby Groves, PhD