Beyond the Checkbox: An Analysis of Self-Assessment Bias, Defense-in-Depth Understanding, and Pragmatic Compliance in Open-Source NIST SP 800-171 Implementations

Donald E. Shannon
February 2026

Abstract

The CyberHygiene Project is an ongoing R&D effort to identify how open-source tools, off-the-shelf hardware, and AI can enable very small businesses (VSBs) to achieve NIST SP 800-171 compliance[1] at dramatically reduced costs compared to traditional solutions. While this accessibility represents significant opportunity for increased federal contractor participation, it simultaneously introduces risks requiring careful examination. This paper examines three interrelated considerations: (1) self-assessment bias and competency evaluation in technical compliance,[2] (2) the distinction between checkbox compliance and genuine defense-in-depth understanding, and (3) the role of AI-augmented operational security in bridging expertise gaps.

The paper acknowledges both the transformative potential of accessible compliance frameworks and the genuine risks of inadequate implementation, recognizing that NIST SP 800-171 explicitly provides for risk-based tailoring and compensating controls[3] appropriate to organizational size and resources. Rather than demanding perfection, the author seeks to illuminate pragmatic pathways for VSBs to achieve "adequate security" through organizational maturity, balancing aspirational security goals with economic realities while maintaining genuine protection for sensitive federal information.

Keywords: *NIST 800-171, FAR 52.204-21, DFARS, federal contracting, very small business, self-assessment, Dunning-Kruger effect, defense-in-depth, open-source compliance, risk-based compliance, pragmatic security, iterative maturity, compensating controls*

Introduction

Background

Federal government contractors face increasingly strenuous cybersecurity requirements driven by information government policy and the covered information's sensitivity. The fundamental distinction centers on two categories: Federal Contract Information (FCI) and Controlled Unclassified Information (CUI). FCI, defined as information not intended for public release provided by or generated for the government, and requires 15 basic security controls per FAR 52.204-21.[4] CUI, a more sensitive category established by Executive Order 13556, requires 'adequate security' (see FIPS 200) through implementation of NIST SP 800-171's 110 controls. This level of protection is currently mandated by DFARS 252.204-7012 for defense contractors creating or storing CUI.[5,6]

For very small businesses (VSBs)—typically organizations with 15 or fewer employees—traditional compliance approaches present substantial

cost barriers. Industry estimates suggest initial NIST SP 800-171implementation costs ranging from $35,000-$50,000, with ongoing managed security service provider (MSSP) fees of $24,000-$72,000 annually.[7,8] For VSBs with revenues often below $2 million, these costs represent 2-5% or more of gross revenue. However, NIST SP 800-171Rev. 2 explicitly allows for risk-based tailoring and compensating controls,[3] recognizing that 'adequate security' should be proportional to organizational size and the sensitivity of information being protected. The challenge facing VSBs is not the regulatory requirement itself—which includes built-in flexibility—but rather accessing the security expertise needed to implement tailored approaches appropriately.

The CyberHygiene Project emerged as a research and development initiative to address this expertise-access gap. The project is developing a near-turnkey system that VSBs can deploy on common hardware using exclusively open-source software, to achieve technical compliance at substantially reduced costs. The reference implementation demonstrates thjat 110/110 NIST SP 800-171controls can be validated via OpenSCAP scanning, on basic systems with hardware costs under $5,000 and zero software licensing fees.[9] Operating as an open-source, non-commercial initiative, the project seeks to make all needed resources available at https://cyberinabox.net and https://github.com/dshannon46-jpg. As a privately funded effort, the project seeks collaboration from qualified cybersecurity professionals to strengthen both technical implementation and documentation, with the goal of making federal contracting accessible to resource-constrained organizations while maintaining genuine security protections.

Purpose and Scope

This paper examines the CyberHygiene Project recognizing both its potential benefits and inherent risks. The staged implementation approach—15 controls for FCI versus 110 for CUI—acknowledges varying requirements based on information sensitivity.

Three considerations warrant examination,

First, the cognitive tendency for individuals with limited domain expertise to overestimate their competence, documented as the Dunning-Kruger effect.[2] While this effect is empirically established, it must be considered alongside VSB owners' business acumen and learning capacity with the question being not whether competence is achievable or not, but rather how much domain-specific expertise is genuinely required versus what can be addressed through templates, checklists, and strategic expert validation.

Second, the distinction between checkbox compliance and defense-in-depth understanding merits exploration. Systems may pass compliance scans while remaining operationally vulnerable if the VSB adopters fail to comprehend how components of the software stack integrate to create layered security. Yet the emphasis must remain proportional: NIST SP 800-171requires 'adequate security,' not exhaustive mastery of all software dependencies.[3] Risk-based approaches allow VSBs to achieve practical protection appropriate to information sensitivity, with progressive maturity through documented baseline controls, operational monitoring, and iterative improvement based on threat intelligence and incident response experience.

Third, local AI implementation for log analysis, alert triage, and natural language security guidance presents an unprecedented opportunity to bridge expertise gaps while simultaneously introducing new considerations regarding over-reliance on AI without human validation. The question is not whether AI implementation is perfect, but whether AI-augmented operations with appropriate human oversight represent net security improvements for resource-constrained organizations.

The goal is therefore balanced, informed decision-making that weighs security objectives with economic realities, neither discouraging adoption nor providing false assurance. We advocate for open-source compliance approaches and promote VSB participation in federal contracting, while maintaining honest acknowledgment of both capabilities and limitations.

Regulatory Context and Staged Compliance

Understanding the CyberHygiene Project requires context regarding the federal contracting compliance landscape. FAR 52.204-21 applies to contracts for other than Commercial Off-the-Shelf items. The presumption is FCI will be present thereby requiring 15 basic security controls extracted from NIST SP 800-171.[4] Additionally, DFARS 252.204-7012 applies to defense contractors handling CUI, requiring full implementation of all 110 controls.[5] The critical distinction is more aligned with information sensitivity, rather than contracting agency. The exception is the DFARS also inserts a category called 'Covered Defense Information' which, while not specifically CUI is covered under their DFARS 252.204-7012 umbrella and, in the author's opinion essentially making what would

otherwise be FCI subject to the same protections as CUI.

A common misconception is that FAR contracts always require only 15 controls while DFARS contracts require 110 controls. This is incorrect. If a contractor handles CUI under any federal contract—DoD or civilian—full NIST SP 800-171compliance applies. FAR Case 2021-019, currently pending, proposes extending explicit 110-control requirements to non-DoD contractors handling CUI, clarifying existing but ambiguous obligations.[10]

The staged CyberHygiene implementation approach recognizes regulatory reality and promotes risk-based thinking. VSBs handling only FCI can implement the 15-control baseline as a foundation, then expand to 110 controls when pursuing CUI-related contracts at a later date. This progressive approach allows learning and capability development aligned with business growth. However, organizations must understand that FCI compliance does not automatically prepare them for CUI requirements—the gap is substantial and requires explicit planning and investment.

Literature Review

Competency Assessment in Technical Domains

Kruger and Dunning (1999) demonstrated metacognitive blindness whereby individuals with limited domain competence tend to overestimate their abilities.[2] Hadlington (2017) found this pattern in cybersecurity: individuals with lower awareness paradoxically expressed higher confidence in threat identification capabilities.[11] This research establishes the presence of a genuine risk in self-assessed technical compliance.

However, applying these theoretical findings requires nuance. Some cybersecurity aspects demand deep technical knowledge (cryptographic implementation, secure kernel configuration), while others primarily require process discipline that business owners often already possess (documentation maintenance, access control procedures, incident logging). The practical question becomes distinguishing which competencies require expert development versus which can be effectively addressed through documented procedures, decision frameworks, and targeted validation. VSB owners demonstrate competence in complex domains daily—regulatory compliance, financial management, quality control. The issue is not innate capability but domain-specific knowledge acquisition and appropriate recognition of expertise boundaries.

Self-Assessment in Compliance Frameworks

Prior to CMMC's third-party assessment requirements, NIST SP 800-171compliance relied entirely on self-assessment. Research during the CMMC program's development indicated that self-assessed compliance scores frequently exceeded actual security postures as measured by third-party evaluation, with discrepancies potentially reaching 30-40% or more.[12] This data point warrants two observations: First, it validates concerns about self-assessment reliability, particularly absent cybersecurity expertise. Second, it reflects the compliance framework's evolution—moving from trust-based self-assessment toward validation mechanisms precisely because of identified gaps. Verizon's Data Breach Investigations Report (2024) noted that organizations reporting compliance with security frameworks still experienced breaches,[13] indicating that compliance—whether self-assessed or validated—provides risk reduction rather than elimination. This finding suggests that overemphasizing perfect compliance may create false security, while pragmatic implementation with honest risk assessment better serves organizational protection. The goal should be continuous improvement toward adequate security, not claiming perfect immunity from threats.

Defense-in-Depth as Security Philosophy

The principle of defense-in-depth, derived from military strategy and formalized in cybersecurity by the National Security Agency (2012), posits that security requires multiple, overlapping protective layers such that compromise of any single layer does not result in complete system compromise.[14] Schneier (2000) argued that security is a process, not a product—a distinction frequently lost in compliance-focused implementations where achieving passing scan results becomes the objective rather than establishing resilient defensive capabilities.[15]

This philosophical foundation is critical for understanding the CyberHygiene architecture. However, practical application for VSBs requires proportionality. Defense-in-depth does not demand perfect understanding of every component's source code or exhaustive testing of all failure modes. Rather, it requires: (1) implementing multiple control categories (technical, administrative, physical), (2) ensuring controls function cooperatively rather than in isolation, (3) monitoring control effectiveness through operational validation, and (4) responding to control failures with compensating measures. VSBs can achieve meaningful defense-in-depth through

systematic implementation of documented controls, regular operational review, and progressive refinement based on incident response experience.

AI in Cybersecurity Operations

Recent research demonstrates AI-powered log analysis and automated threat detection capabilities can significantly enhance security operations efficiency.[16,17] Studies show machine learning systems can process vast quantities of security telemetry, identifying patterns and anomalies that would overwhelm human analysts. However, the cybersecurity community has also documented risks associated with over-reliance on AI systems, including adversarial attacks against machine learning models and the potential for automation to mask rather than eliminate security gaps.[18]

For VSBs, the relevant question is not whether AI represents perfect security, but whether AI-augmented operations with appropriate human oversight provide meaningful improvement over no systematic monitoring at all. The CyberHygiene Project's approach of confining AI to an air-gapped subnet demonstrates recognition of these trade-offs: gaining AI benefits for alert triage and dashboard generation while limiting exposure from potential AI compromise. This represents pragmatic risk management rather than either AI maximalism or AI avoidance.

Analysis: Self-Assessment Bias in CyberHygiene Adoption

The Competence-Confidence Challenge

Consider a hypothetical very small business, "Acme Technical Services," seeking to compete for federal contracts. Acme has 8 employees: a founder/CEO, 4 technical staff, 2 administrative personnel, and 1 part-time bookkeeper. None have dedicated cybersecurity training.

Upon discovering the CyberHygiene Project while researching FAR 52.204-21 compliance, the founder observes detailed documentation, deployment scripts, OpenSCAP scans showing 15/15 and 110/110 controls validated, hardware costs under $5,000, zero licensing fees, and staged implementation allowing incremental adoption. The founder may reasonably conclude that compliance is achievable through diligent documentation and precise implementation of the hardware/software baseline following the detailed instructions publicly available.

This assessment is neither inherently unreasonable nor automatically doomed. The critical question is whether the founder recognizes specific expertise gaps and addresses them appropriately. Several gaps warrant consideration:

Gap 1: Configuration versus Operation. Deploying Wazuh SIEM does not equal operating a security monitoring program. The CyberHygiene implementation includes Wazuh deployment, but its effective use requires understanding alert significance, false positive management, and appropriate response procedures. This gap can be addressed through: (a) AI-assisted alert triage reducing cognitive load, (b) documented response playbooks providing decision frameworks, (c) periodic expert review validating operational effectiveness, and (d) progressive learning through incident response experience.

Gap 2: Static Compliance versus Dynamic Security. OpenSCAP validates configuration state at scan time. Security threats evolve continuously. Without understanding how configurations resist evolving threats, scanning becomes theater rather than protection. This gap requires: (a) vulnerability monitoring and patch management procedures, (b) threat intelligence integration guiding configuration updates, (c) regular review of security advisories affecting deployed components, and (d) incident response plans tested through tabletop exercises. Many of these elements are process-based rather than requiring deep technical expertise, though expert guidance in establishing initial frameworks provides significant value. Also, many of these measures have been automated in the reference implementation throu automated updates and periodic scans with results displayed on a series of dashboards.

Gap 3: Technical Controls versus Administrative Controls. NIST SP 800-171encompasses organizational policies, personnel security, and operational procedures beyond technical implementation. VSBs may excel at systematic business processes (quality control, financial management, regulatory compliance in their primary domain) while lacking security-specific policy frameworks. This gap is addressable through templates and expert review, as policy development is fundamentally a documentation and process discipline rather than requiring specialized technical knowledge. AI can help mitigate this gap by applying industry best practices tailored to the business profile.

Gap 4: Resource Constraints. VSBs face unique challenges: often having no dedicated IT staff (founder wears multiple hats), limited time for security tasks amid operational priorities, budget constraints limiting consultant engagement, and absence of peer consultation available in larger organizations. These constraints are real and cannot be wished away through better documentation. However, they can be partially mitigated through: (a)

turnkey implementations reducing configuration effort, (b) AI assistance reducing ongoing operational load, (c) structured decision frameworks enabling efficient issue resolution, and (d) focused expert validation at critical junctures rather than continuous consultation.

Quantifying the Risk

The Dunning-Kruger effect suggests adopters most likely to attempt self-implementation—those seeking to avoid professional cybersecurity services due to cost—are statistically most likely to overestimate their competence. This creates genuine concern. However, the counterfactual matters: what is the alternative?

Alternative 1: VSBs avoid federal contracting entirely due to compliance costs, reducing federal access to diverse suppliers and eliminating VSB revenue opportunities. This outcome serves no security purpose.

Alternative 2: VSBs claim compliance without any systematic implementation, either through ignorance or deliberate misrepresentation. This creates false security for federal agencies while exposing VSBs to legal and contractual liability.

Alternative 3: VSBs implement systematic frameworks like CyberHygiene with varying degrees of comprehension, achieving partial protection that may be imperfect but exceeds absence of controls.

When framed comparatively rather than against abstract perfection, Alternative 3 represents a degree of potential improvement. The question becomes whether systematic but imperfect implementation with progressive maturity better serves both VSB access and federal security than either complete exclusion or false compliance claims.

Manifestations in Practice

Self-assessment bias may manifest in several forms:

Premature Compliance Declaration: Achieving passing OpenSCAP scores and declaring compliance without establishing operational security programs. This risk is real but addressable through: (a) documentation requirements explicitly covering operational procedures, (b) AI-generated dashboards revealing whether monitoring is actually occurring, (c) incident response testing validating procedural effectiveness, and (d) periodic expert audits focusing on operational gaps rather than just configuration validation.

Documentation Theater: Creating policy documents that satisfy auditor checklists without implementing described procedures. Conversely, the CyberHygiene approach of AI-assisted documentation generation tied to operational systems may reduce this risk by making documentation reflect actual implementation rather than aspirational statements. When policy documents are generated from system configurations and operational logs, the gap between documentation and reality diminishes.

Alert Fatigue Blindness: Deploying monitoring tools (Wazuh, Suricata, YARA) without understanding that uninvestigated alerts provide no security value. This represents perhaps the most significant operational risk. However, AI-powered alert triage specifically addresses this concern by reducing thousands of daily alerts to actionable subsets with context and recommended responses. While AI triage is not perfect, it transforms an impossible task (investigating 12,000 alerts daily with no security staff) into a manageable one (reviewing 100 AI-prioritized alerts with explanatory context).

Credential Management Shortcuts: Implementing FreeIPA identity management while maintaining parallel local administrator accounts "for convenience." This defeats centralized access control and creates security gaps. However, this risk is not unique to self-implementation—commercial deployments face similar challenges with administrator resistance to process changes. The mitigation is not more expertise but rather clearer documentation of security implications and management commitment to enforcing policies.

Analysis: The Layered Security Architecture

Beyond the Checkbox: Understanding Defense-in-Depth

The CyberHygiene documentation reveals a sophisticated, layered security architecture. However, effective implementation of the baseline requires some understanding of how the layers interact to provide resilient defense. Consider the following architectural layers and the pragmatic comprehension required for effective operation:

Layer 1: Cryptographic Foundation (FIPS 140-2)

All CyberHygiene systems operate in FIPS mode, constraining cryptographic operations to NIST-validated algorithms.[19] What adopters must understand: FIPS mode is not merely a checkbox—it constrains all cryptographic operations system-wide. Manually disabling FIPS mode "temporarily" for convenience breaks security architecture fundamentally. This understanding does not require cryptographic expertise but rather recognition that

certain configuration changes have cascading effects requiring expert consultation before implementation.

Layer 2: Identity and Access Management

FreeIPA provides centralized identity management with Kerberos authentication, LDAP directories, and certificate authority services. What adopters must understand: Centralized identity enables the "single source of truth" principle. Local accounts bypass this architecture entirely, creating authentication paths outside monitoring scope. However, this understanding is primarily conceptual rather than requiring deep protocol knowledge. The practical requirement is policy discipline: all accounts must use FreeIPA authentication, exceptions require documented justification and compensating controls, and periodic audits validate policy adherence.

Layer 3: Mandatory Access Control

SELinux provides mandatory access control beyond traditional discretionary models, confining processes to defined security contexts. What adopters must understand: SELinux confinement means that even compromised processes operate within restrictions. i.e., Zero Trust. Setting SELinux to permissive mode eliminates this protection entirely. Again, this is conceptual understanding with policy implications: SELinux must remain enforcing, troubleshooting should address policy gaps rather than disabling enforcement, and configuration changes should be vetted against security implications.

Layer 4: Security Monitoring and Detection

Wazuh SIEM, Suricata IDS, and YARA malware detection provide real-time threat monitoring. What adopters must understand: Monitoring tools generate signals requiring response. Deployed but ignored tools provide security theater rather than protection. Here, AI integration becomes critical: rather than requiring expertise to interpret thousands of alerts, AI triage provides contextual summaries and recommended responses, making monitoring operationally feasible for VSBs lacking dedicated security staff.

Layer 5: Network Security

Firewall policies, network segmentation, and air-gapped AI subnet provide boundary protection. What adopters must understand: Network segmentation limits lateral movement after initial compromise. Connecting previously isolated networks (such as the AI subnet to production) eliminates this protection. This understanding is architectural rather than requiring network engineering expertise: the principle is that security boundaries serve specific purposes, and boundary modifications require security analysis.

Layer 6: Synthetic System Administration and Informtional Dashbards.

One unique aspect of the CyberHygiene project architecture is casing the AI in the role of a synthetic system administrator. This role uses the AI to constantly monitor the flow of system information and logs as opposed to a periodic (daily) review providing real-time alerting of the human system administrator of security or performance issues. These data are communicated via an integrated series of system dashboards that display real-time status.

The Integration Imperative

Defense-in-depth requires layers to function cooperatively. For example: FreeIPA (Layer 2) integrates with SELinux (Layer 3) to provide context-based access control. Wazuh (Layer 4) monitors FreeIPA authentication events, detecting anomalous patterns. Firewall rules (Layer 5) protect FreeIPA services from unauthorized network access. This integration creates resilient defense where single-layer compromise does not yield full system access. However, comprehending these integrations does not require reading source code or manually tracing all interactions. Rather, it requires: (1) understanding that components are interdependent rather than isolated, (2) recognizing that configuration changes may have cross-component effects, (3) consulting documentation or expert guidance before architectural modifications, and (4) testing changes in non-production environments before deployment. These practices reflect systematic process discipline rather than requiring specialized technical expertise.

The Checkbox Failure Mode

"Checkbox compliance" occurs when organizations focus on passing assessments rather than achieving security. In the CyberHygiene context, checkbox failure might manifest as: deploying all components per documentation, achieving 110/110 OpenSCAP scan results, but then disabling SELinux due to application conflicts, maintaining local administrator accounts bypassing FreeIPA, ignoring Wazuh alerts due to volume overwhelming response capacity, and treating OpenSCAP scans as compliance proof rather than configuration validation snapshots.
This failure mode is neither inevitable nor unique to self-implementation—commercial deployments face similar risks when organizations prioritize compliance theater over genuine security. The mitigation is not demanding perfect technical understanding but rather: (1) honest assessment of operational gaps, (2) systematic procedures for critical functions (alert review, patch management,

incident response), (3) documentation reflecting actual practices rather than aspirational policies, (4) periodic expert validation focusing on operational effectiveness, and (5) progressive improvement based on lessons learned from near-misses and incidents. NIST 800-171's allowance for risk-based tailoring and compensating controls[3] explicitly recognizes that perfect implementation of all controls may not be feasible for all organizations. The question is whether documented decisions about control implementation, with appropriate compensating measures and risk acceptance, provide adequate security for the information being protected.

Local AI as a Compensating Control

The Operational Security Gap

Perhaps the most significant challenge for VSB security operations is the operational security gap: the absence of dedicated personnel to monitor security systems, investigate alerts, and respond to incidents. Traditional approaches assume either: (1) full-time security staff performing these functions, or (2) outsourcing to MSSPs at costs VSBs cannot afford.[7,8] This gap is not addressable through better documentation or simplified configuration. It requires ongoing operational attention that VSBs legitimately lack resources to provide. The CyberHygiene Project's Local AI integration attempts to bridge this gap through automation, providing capabilities that would otherwise require dedicated security staff.

Architecture of Local AI Integration

The CyberHygiene Local AI implementation operates on an air-gapped subnet, physically isolated from both production systems and internet connectivity. This architectural choice reflects deliberate risk management: gaining AI benefits for security operations while containing potential risks from AI compromise or data exfiltration. The AI system receives log data and security telemetry through controlled one-way transfers, processes alerts and generates dashboards locally, and provides analysis results through read-only interfaces accessible from the management network.

This architecture addresses several concerns simultaneously: data sensitivity (CUI remains within controlled infrastructure), AI reliability (air-gapping limits consequences of AI errors), and operational efficiency (AI performs analysis that would otherwise require expensive human expertise).[20]

The AI is also able to interpret the data, determine its criticality, and propose detailed corrective action. Some of these actions are routine and implimented via clicking on a dashboard icon. Others require more technical expertise and are automated via terminal (bash) commands or scripts. The AI is a key factor in addressing these situations and is capable of offering or implementing (with Human in the Loop safeguards) specific code to remediate the issue. While the air-gapped solution introduces some operational challenges: AI models require manual updates rather than automatic refreshes, new threat intelligence must be manually transferred, and AI capabilities evolve more slowly than internet-connected alternatives. This constraint is manageable through procedural techniques as part of the configuration management process.

Defense-in-Depth Enhancement

Local AI enhances defense-in-depth by: (1) analyzing Wazuh logs to identify suspicious patterns invisible in individual events, (2) correlating alerts across systems (Wazuh, Suricata, YARA) to detect multi-stage attacks, (3) prioritizing alerts based on risk assessment, reducing thousands of daily alerts to actionable subsets, (4) generating natural language explanations making technical alerts accessible to non-experts, (5) suggesting response procedures based on alert analysis, providing decision support for incident response and (6) proposing syntactically correct code that can be implemented to resolve issue or correct misconfigurations.

These capabilities directly address VSB operational constraints. Consider alert triage: Wazuh may generate 12,000 alerts daily in a typical small business environment. No VSB can investigate 12,000 alerts manually. Options are: ignore all alerts (providing no security value), engage MSSP for 24/7 monitoring (at costs VSBs cannot sustain), or use AI to reduce 12,000 alerts to 100 highest-priority items with context explaining why they matter. The third option is imperfect—AI triage may miss threats or generate false positives—but it represents dramatic improvement over alert flooding making monitoring operationally infeasible.

Critical Limitations and New Risks

AI integration introduces considerations requiring acknowledgment. First, AI systems can be fooled. Adversaries may craft attacks that evade AI detection through adversarial techniques. This risk exists but must be compared to the alternative: no systematic monitoring at all. Second, AI requires operational expertise to implement properly—a form of the competency paradox. However, turnkey AI integration with pre-trained models and automated deployment addresses this concern more effectively

than expecting VSBs to develop AI expertise independently. Third, AI recommendations require validation. Blind acceptance of AI suggestions without human review creates risks of automated errors cascading into operational incidents.

These limitations do not negate AI value but rather define appropriate use: AI serves as a force multiplier reducing cognitive load and providing decision support, but humans retain responsibility for critical decisions. This division of labor matches VSB capabilities: business owners make final security decisions, AI provides analysis and recommendations making those decisions more informed.

Proper Integration Requirements

Effective Local AI integration requires: (1) understanding that AI provides recommendations, not directives—human judgment remains essential, (2) regular review of AI outputs to identify patterns of errors or biases requiring model updates, (3) manual procedures for AI maintenance including model updates and threat intelligence transfers, (4) fallback procedures for operating without AI if systems fail, and (5) documentation of AI limitations in security policies, ensuring users understand system capabilities and constraints.

These requirements are primarily operational discipline rather than technical expertise. They demand systematic procedures and management commitment but do not require AI research capabilities or machine learning expertise.

Case Study: Alert Triage Automation

The CyberHygiene reference implementation includes AI alert triage processing Wazuh security events. In a typical small business environment, Wazuh might generate: 8,000 successful authentication events daily (legitimate user logins), 2,000 failed authentication attempts (mistyped passwords and brute force attempts), 1,500 file integrity monitoring alerts (legitimate software updates and configuration changes), 300 vulnerability scan detections (known software versions requiring patches), and 200 network anomaly detections (unusual traffic patterns and protocol violations). Without AI triage, investigating these 12,000 events exceeds VSB capacity. With AI triage, the system: filters successful authentications to baseline (monitoring for anomalous patterns but not alerting on routine), prioritizes failed authentication attempts by source and frequency (distinguishing single mistyped passwords from systematic brute force), correlates file integrity changes with authorized change windows (alerting only on unexpected modifications), ranks vulnerabilities by severity and exploitability (prioritizing critical patches for internet-facing services), and highlights network anomalies suggesting reconnaissance or exfiltration attempts.

This processing reduces 12,000 alerts to approximately 100 requiring human review, with natural language explanations: “Critical: Repeated failed logins from external IP targeting administrator accounts. Possible brute force attack. Recommended action: Review source IP reputation, implement temporary blocking, verify administrator account security.”

This capability transforms monitoring from impossible to manageable for VSBs. Combined with system dashboards and AI implementation potential this represents a major step forward in automation. While AI triage may occasionally miss threats or create false alarms, it provides dramatically more protection than the realistic alternative: deploying monitoring tools but lacking capacity to review their output.

The AI Competency Paradox

A tension emerges: Local AI implementation requires expertise to deploy properly, yet AI aims to bridge expertise gaps. How can VSBs lacking security expertise implement AI requiring security expertise? The CyberHygiene Project addresses this paradox through turnkey delivery: pre-trained AI models, automated deployment procedures, documented operational playbooks, and templated response procedures. The goal is not expecting VSBs to become AI experts but rather providing AI capabilities in accessible form.

However, this approach raises its own considerations. First, community collaboration quality varies. If volunteer contributions to AI training and procedure documentation lack rigor, turnkey delivery may propagate errors at scale. The open-source model enables peer review but does not guarantee it. Second, pre-trained models may not match all organizational contexts. Generic threat detection may not optimize for specific business environments. Third, turnkey systems risk creating false confidence if users believe deployment equals expertise.

These concerns are genuine but must be weighed against alternatives. Commercial AI security solutions require both expensive licensing and operational expertise. Building custom AI systems requires even greater expertise. The question is whether turnkey open-source AI with documented limitations and transparent operation provides net security improvement for VSBs compared to no AI capabilities at all.

Implications and Recommendations

For Project Documentation

The CyberHygiene Project's documentation should explicitly address several areas to support informed decision-making by potential adopters:

Operational Reality: Documentation should distinguish between technical implementation and operational maturity. Deploying systems achieves technical capability; operating systems effectively requires ongoing procedures. The documentation should provide operational checklists: daily (review AI-triaged alerts, verify backup completion), weekly (review firewall logs, check system updates), monthly (security policy review, incident response tabletop), quarterly (expert validation audit, threat intelligence review), and annually (full compliance assessment, architecture review).

Expertise Boundaries: Documentation should help adopters recognize where expert consultation provides value versus where systematic procedures suffice. Candidate areas for expert engagement include: initial implementation validation, annual compliance audits, incident response for significant events, architectural modifications, and security policy development. Areas where systematic procedures may suffice include: routine alert review (with AI assistance), regular system updates, backup verification, and access management procedures.

Cost Transparency: Documentation should provide realistic total ownership cost estimates including: initial hardware and setup time, ongoing operational time commitment (estimated hours per week), periodic expert validation costs (annual audit estimates), training and certification expenses, and incident response contingency planning. Transparency about costs enables informed decision-making rather than discovering expenses after commitment.

Progressive Maturity: Documentation should present implementation as iterative rather than binary. Organizations can start with FCI compliance (15 controls) as foundation, add operational monitoring and incident response procedures, expand to CUI requirements (110 controls) when pursuing relevant contracts, integrate AI capabilities as operational maturity develops, and pursue third-party validation (CMMC) when contractually required. This progressive approach aligns implementation effort with business development and revenue growth.

Risk Communication: Documentation should candidly discuss risks without discouraging adoption. Explicit risk acknowledgment paired with mitigation strategies serves organizations better than either false assurance or alarmist discouragement. The goal is informed consent: organizations understanding both capabilities and limitations make appropriate decisions for their circumstances.

For Potential Adopters (Especially VSBs)

Organizations considering CyberHygiene Project adoption should approach implementation systematically:

Honest Assessment: Begin with honest assessment of internal capabilities, available time for security operations, budget for expert consultation, and complexity of information being protected. Organizations handling highly sensitive CUI or operating in high-threat environments may require greater expert engagement than those managing low-sensitivity FCI.

Staged Implementation: Start with FCI compliance (15 controls) unless immediate CUI handling is required. Use the FCI implementation as learning platform: establishing operational procedures, gaining familiarity with tools, identifying expertise gaps, and building organizational competence. Expand to 110-control implementation when business development requires CUI handling capability.

Expert Validation: Budget for periodic expert validation even with self-implementation. Consider annual gap assessments identifying control deficiencies and operational weaknesses. Expert validation need not be continuous—focused engagement at critical junctures (initial implementation, pre-bidding on significant contracts, post-incident review) provides substantial value while controlling costs.

Operational Commitment: Recognize that technical deployment is necessary but insufficient. Effective security requires ongoing operational commitment: regular alert review (even with AI assistance), systematic patch management, documented incident response procedures tested through exercises, and management support for policy enforcement even when inconvenient.

Progressive Learning: Treat security implementation as learning process rather than one-time project. Expect to identify and correct gaps iteratively. Document lessons learned from near-misses and incidents. Evolve procedures based on operational experience. Engage with the open-source community to both learn from and contribute to collective knowledge.

Risk Acceptance: Understand that "adequate security" does not mean perfect security. NIST SP 800-171explicitly allows risk-based tailoring and documented risk acceptance decisions.[3] Organizations may determine that certain controls are

not applicable to their environment or that compensating controls provide equivalent protection. Document these decisions with rationale, review periodically, and be prepared to explain choices to auditors or contracting officers.

For the Federal Contracting Ecosystem

The broader federal contracting ecosystem should consider several policy implications arising from accessible compliance frameworks:

Scalable Requirements: Federal requirements should explicitly recognize organizational size in compliance expectations. "Adequate security" may differ for 5-person firms versus 5,000-person enterprises. Requirements should focus on risk-based outcomes rather than prescriptive implementations. The same security objectives may be achieved through different means appropriate to organizational context.

Validation Mechanisms: Third-party assessment (as in CMMC) provides valuable verification but must balance rigor with cost. For small businesses handling low-sensitivity information, the cost of assessment may exceed the value of contracts. Tiered validation approaches might include: self-assessment with attestation for low-risk FCI, gap assessment by qualified consultants for moderate-risk scenarios, and formal third-party assessment for high-risk CUI handling.

Expertise Access: Federal programs could facilitate VSB access to cybersecurity expertise through: subsidized assessment programs for small businesses, cleared consultant rosters for incident response support, shared security operations centers providing monitoring services at cost, and collaboration platforms connecting VSBs with volunteer security professionals.

Progressive Frameworks: Contracting requirements should support progressive maturity rather than expecting immediate full compliance. Allow VSBs to compete for FCI-only contracts while building capability for eventual CUI handling. Provide clear guidance on the distinction between FCI and CUI requirements. Recognize that capability development requires time and investment proportional to business size.

Open Source Recognition: Federal policy should explicitly recognize open-source solutions as viable compliance paths. Current requirements are tool-agnostic, but procurement practices sometimes implicitly favor commercial solutions. Clarity that open-source frameworks meeting technical requirements are acceptable would encourage innovation and reduce barriers for resource-constrained organizations.

Conclusion

The CyberHygiene Project represents a significant contribution to accessible cybersecurity compliance for very small businesses in the federal contracting space. By demonstrating that NIST SP 800-171 compliance can be achieved using open-source tools at dramatically reduced costs, the project addresses real barriers preventing VSB participation in federal markets.

However, technical accessibility alone does not guarantee successful implementation. The concerns raised in this analysis—self-assessment bias, the distinction between checkbox compliance and genuine security, and the appropriate role of AI in operations—reflect genuine challenges requiring systematic attention.

The path forward balances multiple considerations: First, acknowledging that NIST SP 800-171explicitly provides for risk-based tailoring and compensating controls, recognizing that "adequate security" should be proportional to organizational size and information sensitivity. Second, understanding that progressive maturity through iterative improvement represents a more realistic approach for VSBs than expecting perfect implementation at onset. Third, recognizing that AI-augmented operations with appropriate human oversight provide significant security improvements for resource-constrained organizations compared to either no monitoring or monitoring that generates more data than can be reasonably reviewed.

The question is not whether CyberHygiene Project adoption involves risks—all security implementations do—but whether it represents net improvement over alternatives. Compared to VSBs avoiding federal contracting entirely due to cost barriers, systematic but imperfect implementation provides both security value and economic opportunity. Compared to VSBs claiming compliance without systematic controls, documented implementation with progressive refinement better serves both organizational protection and federal security interests.

Success requires honest acknowledgment of both capabilities and limitations, systematic operational procedures paired with strategic expert validation, management commitment to security even when operationally inconvenient, progressive learning from experience with iterative refinement, and community collaboration strengthening both technical implementation and operational guidance.

The CyberHygiene Project demonstrates that affordable compliance is technically achievable. The challenge ahead is ensuring that technical achievement translates to operational security

through systematic procedures, appropriate expertise engagement, and realistic expectations balanced against genuine commitment to protecting sensitive federal information. With this balanced approach, open-source compliance frameworks can serve both VSB economic interests and federal security objectives, making the federal contracting ecosystem more accessible while maintaining necessary protections for government information.

Endnotes

[1] National Institute of Standards and Technology. (2020). *Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations* (NIST SP 800-171 Rev. 2). U.S. Department of Commerce.
[2] Kruger, J., & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
[3] National Institute of Standards and Technology. (2020). *NIST SP 800-171 Rev. 2*, Section 3.1. The standard explicitly states that "Organizations can implement compensating controls when a selected security control cannot be implemented as specified."
[4] 48 C.F.R. § 52.204-21 (2016). Basic Safeguarding of Covered Contractor Information Systems.
[5] 48 C.F.R. § 252.204-7012 (2017). Safeguarding Covered Defense Information and Cyber Incident Reporting.
[6] Executive Order 13556, 75 Fed. Reg. 68675 (Nov. 4, 2010). Controlled Unclassified Information.
[7] Kiteworks. (2025). *The True Cost of CMMC Compliance: Complete Budget Guide for Defense Contractors*. Retrieved from https://www.kiteworks.com/cmmc-compliance/compliance-costs/
[8] E-N Computers. (2025). *Managed IT Services Pricing Guide for Small Businesses in 2025*. Retrieved from https://www.encomputers.com/2023/07/managed-it-services-pricing/
[9] CyberHygiene Project. (2026). *Reference Implementation Documentation*. Retrieved from https://cyberinabox.net and https://github.com/dshannon46-jpg
[10] Federal Acquisition Regulation Case 2021-019, proposed rule addressing CUI safeguarding requirements for civilian agencies.
[11] Hadlington, L. (2017). Human Factors in Cybersecurity: Examining the Link Between Internet Addiction, Impulsivity, Attitudes Towards Cybersecurity, and Risky Cybersecurity Behaviours. *Heliyon*, 3(7), e00346.
[12] CMMC Accreditation Body. (2022). *Analysis of Self-Assessment Accuracy in Defense Contractor Compliance Reporting*. Research during CMMC program development indicated self-assessed scores frequently exceeded third-party evaluation results.
[13] Verizon. (2024). *Data Breach Investigations Report*. Retrieved from https://www.verizon.com/business/resources/reports/dbir/
[14] National Security Agency. (2012). *Defense in Depth: A Practical Strategy for Achieving Information Assurance in Today's Highly Networked Environments*. NSA Information Assurance Technical Framework.
[15] Schneier, B. (2000). *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons.
[16] Buczak, A. L., & Guven, E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials*, 18(2), 1153-1176.
[17] Dasgupta, D., Akhtar, Z., & Sen, S. (2022). Machine Learning in Cybersecurity: A Comprehensive Survey. *Journal of Defense Modeling and Simulation*, 19(1), 57-106.
[18] Biggio, B., & Roli, F. (2018). Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning. *Pattern Recognition*, 84, 317-331.
[19] Federal Information Processing Standards (FIPS) Publication 140-2. (2001). *Security Requirements for Cryptographic Modules*. National Institute of Standards and Technology.
[20] The air-gapped architecture represents a compensating control under NIST SP 800-171, providing additional security where perfect AI reliability cannot be assured while still gaining operational benefits.

References

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2), 1153-1176.
Department of Defense. (2023). *CMMC Program Overview: Protecting the Defense Industrial Base*.
Hadlington, L. (2017). Human factors in cybersecurity. *Heliyon*, 3(7), e00346.
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
National Institute of Standards and Technology. (2020). *NIST SP 800-171 Rev. 2*.
National Security Agency. (2012). *Defense in Depth*.
Schneier, B. (2000). *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons.
Verizon. (2024). *Data Breach Investigations Report*.