COMPANION PAPER 2

Comparison Scoring of Existing Frameworks

Scored against the IQ4hire 4D Cybersecurity Framework



IQ4hire Research · November 2026
Proposed and not yet validated

COMPANION PAPER 2

Comparison Scoring of Existing Frameworks

IQ4hire Research · November 2026 · Proposed and not yet validated

This paper supports the framework paper, "The IQ4hire 4D Cybersecurity Framework." It shows how the comparison was scored, so that anyone can repeat it or disagree with it.

1. Purpose and scope

The framework paper says that flat frameworks leave out the questions that decide outcomes: how fast a control acts, who owns each gap, whether you have tested yourself, and whether you are ready for quantum computing. This paper tests that statement against the text of the frameworks, line by line, and supplies a scoring sheet for comparing providers.

What was read

Source	What we read	How it is used
NIST Cybersecurity Framework (CSF) 2.0, NIST CSWP 29, Feb 26, 2024	The document in full, all 32 pages	**Scored**
NIST CSF 2.0 Reference Tool download (Implementation Examples and Informative References, 185 subcategories, about 6,300 words of examples)	Read in full from NIST's site	**Scored** (the "measured" reading)
CIS Critical Security Controls v8.1	The public list of the 18 Controls with their descriptions	Control-level floor scores only (section 6). The detailed Safeguards were not read
MITRE ATT&CK and D3FEND	ATT&CK home and resources pages; D3FEND home and about pages	Content libraries, not scored (section 7)
MITRE Engage	The public v1.0 data files (goals, approaches and 31 activities) from MITRE's GitHub repository	Content library, not scored (section 7)

Where a source was not read, this paper says nothing about it.

2. Method

What is scored. Not how well an organization performs, but **how much the framework asks of the reader** on each part of the 4D Framework. A framework that never raises the question scores 0 on it.

The coverage ladder

Score	Meaning
0	Not mentioned in the text
1	The topic is touched, but the framework does not ask the question the part asks
2	An outcome is stated: what should be true, with no way to measure it
3	An outcome with implementation guidance (examples, references)
4	A measurable criterion: a metric, a target or a defined level

Score	Meaning
5	Requires timed, tested evidence

Rules

1. Score the specific question the 4D part asks. Adjacent coverage that does not answer it earns 1.
2. Cite the outcome ID or the search record for every score.
3. **Document reading (strict):** score only the CSF 2.0 document. **Measured reading:** add what NIST's own Implementation Examples supply, which are guidance and not requirements.
4. Part scores are averaged within a dimension (A has five parts, S has two, K has two). The **4D Coverage Index** is the fourth root of the product of the four dimension scores, as in the framework paper.

Two scorers. Scorer 1 is the paper's author. Scorer 2 is an independent AI model instance that was given only the CSF 2.0 document text and the ladder, was not shown Scorer 1's scores, and was told to be strict. It is not a human, and it may share blind spots with the model that helped write this paper. Human second scorers are listed under next steps (section 9).

Limits. The index measures coverage of the 4D picture, not the quality of NIST's work.

3. Results: NIST CSF 2.0

Part	The question the 4D Framework asks	Scorer 1 (document)	Scorer 2 (document)	Measured (with online examples)	Basis
T	Do controls cover everything you run, in every direction?	2	2	3	ID.AM-01 to -08, DE.CM-01 to -09. Examples cover IT, IoT, OT, mobile, containers, APIs and cloud services, and "constantly monitor networks to detect new hardware." No coverage measure
A1	Who owns the handoff between edge protection and internal monitoring?	1	1	1	DE.CM-01, DE.AE-03 (correlation across sources). No edge or WAF telemetry; no owner for the handoff
A2	Who watches outbound traffic; is egress default-deny?	1	1	2	ID.AM-03 baselines of data flows; examples add comparing actual flows to baselines (DE.CM-01 Ex4). No egress control or default-deny
A3	How fast does a person look at behavior outside the playbook?	2	1	2	DE.CM-03, DE.AE-02 (examples add behavior analytics and "regular manual reviews of log events"). No time measure; "playbook" never appears
A4	Short-lived, per-run credentials for machines and AI agents?	2	2	2	PR.AA-01 (identities and credentials for users, services and hardware), PR.AA-05 (least privilege). Examples add issuing and revoking certificates and tokens. No credential lifetimes; AI agents appear once, as a trigger for updating policy (GV.PO-02)

Part	The question the 4D Framework asks	Scorer 1 (document)	Scorer 2 (document)	Measured (with online examples)	Basis
A5	Who owns handoffs with vendors and connected apps?	2	2	3	GV.SC-01 to -10, ID.AM-04, DE.CM-06. Examples add responsibility matrices, security requirements in service-level agreements and supplier participation in exercises (GV.SC-02, -05, -08). No scoring of technical handoffs; connected-app permissions not addressed
S1	How fast is protection in place after a new flaw?	2	2	3	ID.RA-01, -07, -08, PR.PS-02, RS.MI-01, -02. Examples add "routine and emergency patching within the timeframes specified in the vulnerability management plan." No timeframe is given; no expiry dates on temporary controls; no kill-switch authority
S2	Do you test yourself like an attacker, rehearse and hunt?	1	2	3	ID.IM-02 (improvements from tests and exercises). Examples add penetration testing "of selected high-risk systems as approved by leadership," tabletop exercises, threat hunting (ID.RA-03) and monitoring of deception technology (DE.CM-03, RS.AN-03). Offered as options, never required or scored
K1	Do you adapt now and clean up fully after a hit?	2	2	2	RS.MI-02, RS.AN-03, RC.RP, ID.IM-03. Nothing on credential rotation, removing persistence or expecting repeat attacks
K2	Is there a public-key inventory and a post-quantum plan?	0	1	1	PR.DS-01 and -02 protect data at rest and in transit. The document never mentions quantum, cryptography or encryption. The examples mention encryption and cryptographic key management only as protections. **Quantum appears nowhere**

Dimension	Scorer 1	Scorer 2	Mean of two	Measured
T	2.0	2.0	2.0	3.0
A (average of 5)	1.6	1.4	1.5	2.0
S (average of 2)	1.5	2.0	1.75	3.0
K (average of 2)	1.0	1.5	1.25	1.5
4D Coverage Index	**1.48**	**1.70**	**1.60**	**2.28**
Weakest	K	A	K	K

Where the scorers differ. They agree on seven of ten parts and differ by one point on three: A3 (Scorer 2 reads "no time measure" as a 1), S2 (Scorer 2 credits the outcome about improving from tests as a 2) and K2 (Scorer 2 gives the adjacent data-protection outcomes a 1, where Scorer 1 gave 0 because cryptography is never named). Each reading is defensible. We report both and their mean.

Sensitivity of the measured reading. The measured index depends on two judgment calls: S1 and S2 at 3. If both are scored 2 the index is 2.06. If K2 is scored 0 it is 2.06. The measured reading therefore runs from about 2.1 to 2.3.

What this shows

- On every reading, CSF 2.0 stays below 3.0, the "partial" anchor of the 4D rubric. Index values run from 1.48 (document, strict) to 2.28 (with NIST's online examples).
- Reading the online examples helps CSF 2.0 in one important place. Offense is not absent from the examples: they suggest penetration testing, tabletop exercises, threat hunting and deception monitoring. The document itself mentions none of them.
- They are options, not requirements. Nothing is scored, timed or required, and no example sets a response time.
- Strongest ground: coverage (T) and supply chain (A5, ten outcomes with concrete examples).
- Weakest ground: K2. Quantum computing is mentioned nowhere in the document or the examples, and cryptography appears only as a protection to apply. After that comes the handoff between edge and internal monitoring (A1) and short-lived machine and AI credentials (A4).
- Missing under every reading: a score, a response time, and a requirement that anything has been tested.

Cautions. NIST says the CSF "does not embrace a one-size-fits-all approach" and that its outcomes "are not a checklist of actions to perform." This comparison does not claim NIST failed at what it set out to do. It shows what the document and its examples leave for the reader to supply. It also does not show how organizations that use the CSF perform.

4. The search record

Anyone can repeat these searches.

Term (case-insensitive)	CSF 2.0 document (about 9,400 words)	Online Implementation Examples (about 6,300 words)
quantum	0	0
cryptograph*	0	4 (certificates, keys, hashes)
encrypt*	0	6 (protect data at rest and in transit; encrypt or block outbound email)
penetration	0	1 (ID.IM-02)
red team	0	0
threat hunting	0	1 (ID.RA-03)
deception	0	2 (DE.CM-03, RS.AN-03)
honeypot	0	0
tabletop	0	1 (ID.IM-02)
exercise	2	5
playbook	0	0
hours, minutes	0	0
short-lived, rotate, rotation	0	0
egress	0	0
artificial intelligence	3 (scope statements)	1 (GV.PO-02, as a reason to update policy)
timely	4	1
supplier	16	not counted

Term (case-insensitive)	CSF 2.0 document (about 9,400 words)	Online Implementation Examples (about 6,300 words)
checklist	1	not counted

The one use of "checklist" in the document: "These outcomes are not a checklist of actions to perform; specific actions taken to achieve an outcome will vary by organization and use case."

5. Tiers

NIST's Implementation Tiers (Partial, Risk Informed, Repeatable, Adaptive) describe "a progression from informal, ad hoc responses to approaches that are agile, risk-informed, and continuously improving" (CSWP 29, Appendix B and Fig. 4), for cybersecurity risk governance and management. They describe how an organization manages risk. They are not defined for the 4D parts, so neither scorer used them as a measurable criterion. A looser reading would lift A5 and K1 by a point or so and would not change the conclusion.

6. CIS Controls v8.1: control-level floor

The public CIS page lists 18 Controls with a short description of each. The 153 Safeguards that carry the detail were not read, so these are **floor** scores from the descriptions alone. They are not an index and are not comparable to the CSF scores above. Reading the Safeguards would likely raise several.

4D part	CIS Control (public description)	Floor
T	Control 1, Inventory and Control of Enterprise Assets: "to accurately know the totality of assets that need to be monitored and protected"; Control 2 covers software	2
A1, A2, A3	Control 13, Network Monitoring and Defense: "comprehensive network monitoring and defense"; Control 8, Audit Log Management	1
A4	Control 5, Account Management ("service accounts") and Control 6, Access Control Management ("create, assign, manage, and revoke access credentials and privileges for user, administrator, and service accounts")	2
A5	Control 15, Service Provider Management: evaluate providers that hold sensitive data or run critical platforms	2
S1	Control 7, Continuous Vulnerability Management: "remediate, and minimize, the window of opportunity for attackers"	2
S2	**Control 18, Penetration Testing:** "Test the effectiveness and resiliency of enterprise assets through identifying and exploiting weaknesses in controls… and simulating the objectives and actions of an attacker." Control 17, Incident Response Management, adds a capability "to prepare, detect, and quickly respond"	2
K1	Control 17 and Control 11, Data Recovery	2
K2	Control 3, Data Protection: "identify, classify, securely handle, retain, and dispose of data." Encryption is not named in the description	0

Unlike the CSF, the CIS Controls make penetration testing its own Control, so a reader who already uses CIS has an offense element. Like the CSF, the descriptions set no response times, ask for no evidence, and do not mention quantum computing.

7. MITRE libraries: content, not scored

These are not requirement sets. They tell a reader what adversaries do and what countermeasures or engagement techniques exist. Scoring them on "how much they ask of the reader" would be unfair, so this paper records what content they supply to each part.

4D part	Content supplied	Source read
S1 (shield)	D3FEND is "a knowledge graph of cybersecurity countermeasures." ATT&CK lists mitigations by technique	D3FEND home; ATT&CK menu
S2 (offense: test)	"ATT&CK provides a common language and framework that red teams can use to emulate specific threats and plan their operations," with guidance on adversary emulation and on running your own ATT&CK evaluations	ATT&CK resources page
S2 (offense: traps)	**Engage v1.0** is MITRE's adversary engagement framework: three goals (Expose, Affect, Elicit) plus two strategic goals (Prepare, Understand), seven approaches (Collect, Detect, Prevent, Direct, Disrupt, Reassure, Motivate) and 31 activities, including Lures, Network Manipulation, Isolation, Introduced Vulnerabilities and Malware Detonation. Its data describe deception, decoys and honeypots at length	Engage data files on GitHub (v1.0)
T, A, K1, K2	None found in the MITRE pages and files we read	Not assessed further

Engage is the deepest public source for the "traps" half of S2. Some of its activities aim to affect or disrupt an adversary. Those need legal review before use, and the framework paper keeps all active measures on systems you own.

8. The 4D comparison sheet for providers and organizations

Use this sheet to compare managed security providers, software vendors or your own team. Ask each question in the "show me" column. Score each row as the **lower of the level claimed and the evidence rung** (framework paper, section 4). The how-to-score paper gives the full method.

Part	The question	Show me (evidence to request)	Claimed (1 to 5)	Evidence rung (1 to 5)	Score (the lower)
T	Does it cover everything you run?	The current inventory of internet-facing apps, APIs and test copies, each with an owner, and the monitoring coverage report			
A1	Who owns edge-to-network?	The named owner for what edge protection sees and what internal monitoring sees, and the last incident where both were used			
A2	Who watches outbound traffic?	The outbound rule set (default-deny) and who reviews its alerts			
A3	Who looks at what is outside the playbook?	The last time such activity was found, and the time until a person looked			
A4	How are machine and AI identities handled?	The list of AI agents and machine identities, credential lifetimes, and what each can reach. For agents that write controls, how each control is signed and verified			

Part	The question	Show me (evidence to request)	Claimed (1 to 5)	Evidence rung (1 to 5)	Score (the lower)
A5	Who owns vendor and app handoffs?	The list of connected third-party applications with permissions, and the written handoff terms with each provider			
S1	How fast is protection in place?	The time to protection for the last critical flaw, whether a person or a pre-approved rule produced it, the temporary control used with its owner and expiry date, and who holds the kill switch			
S2	Do they test themselves first?	The last outside or inside test (scope, date, findings, fixes), the last rehearsal of the 3 a.m. call, and any traps in place			
K1	Does a hit stay contained and cleaned up?	The last credential rotation, the last containment exercise, and the cleanup record after the last incident			
K2	Is there a quantum plan?	The encryption inventory ranked by how long each secret must last, and the migration plan with dates			

Scoring. T, A, S and K are the averages of their parts. The **4D Index** is the fourth root of T × A × S × K. Report the weakest dimension next to it.

A hypothetical example (not real providers). Provider A claims (T, A, S, K) of (5, 2, 3, 3). Provider B claims (3, 3, 3, 3).

	Claimed	Index	After "show me"	Index	Weakest
Provider A	(5, 2, 3, 3)	3.08	(3, 2, 3, 3), because it could show a working inventory but not full coverage	2.71	A
Provider B	(3, 3, 3, 3)	3.00	(3, 3, 3, 3), all backed by evidence	3.00	none

On claims alone, Provider A looks slightly better. After "show me," Provider B ranks higher. Do not rank on the index alone. Treat any dimension below 3 as a condition to fix in the contract.

9. Limits and next steps

1. **Human second scorers.** The second scorer so far is an AI model instance. Two or more certified security professionals should score the same ten parts from the document and the search record without seeing these scores.
2. **CIS Control Safeguards.** Read the 153 Safeguards and replace the floor with a measured score.
3. **MITRE Engage live site.** v1.0 data was read from GitHub. Check the live site for newer versions.
4. **Re-run when NIST revises the CSF or its online examples,** and publish the changes.

References

1. NIST. (2024). The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29, February 26, 2024. https://doi.org/10.6028/NIST.CSWP.29
2. NIST. CSF 2.0 Reference Tool: Implementation Examples and Informative References (download). https://csrc.nist.gov/projects/cybersecurity-framework/filters

3. Center for Internet Security. CIS Critical Security Controls v8.1. https://www.cisecurity.org/controls/cis-controls-list
4. MITRE. ATT&CK. https://attack.mitre.org · ATT&CK resources. https://attack.mitre.org/resources/
5. MITRE. D3FEND. https://d3fend.mitre.org
6. MITRE. Engage v1.0 data. https://github.com/mitre/engage · live site https://engage.mitre.org
7. IQ4hire Research. (2026). The IQ4hire 4D Cybersecurity Framework. Framework paper, version 1.0.

Notices

IQ4hire is not affiliated with, endorsed by or sponsored by NIST, the Center for Internet Security, MITRE, Cisco, OpenAI, Microsoft, Verizon, GreyNoise, CERT-In or any other organization named in this paper. Product, framework and company names belong to their owners and are used only to identify the works discussed. Quotations are taken from the published documents, with sources cited. The scoring in this paper is IQ4hire's own method and opinion, and it is open to anyone to repeat or dispute.