NIST CSF v2.0 Concept Paper Feedback

Jack Jones, Chairman, The FAIR Institute

March 3, 2023

Thank you for the opportunity to review and provide feedback on the proposed changes to NIST CSF. As a community-focused project, it is undoubtedly challenging to reconcile and prioritize the many perspectives and suggestions you receive. With that in mind, I've constrained my feedback to only those considerations that I feel are particularly relevant to the following CSF objectives:

- Flexibility
- Relative simplicity (to use and to explain)
- Enable accurate risk-based reporting
- Enable accurate and reliable benchmarking

Note that "accuracy" and "reliability" are not highlighted as criteria within the CSF, but it seems reasonably safe to assume that the intention of any framework would include these criteria. They have been called out here because there are characteristics of the CSF today, and how it's being used by some organizations, that do not fully support those two criteria.

Subcategory Descriptions

In an effort to avoid overwhelming users with detail, NIST has to-date elected to keep the number of subcategories within CSF as low as possible. This simplicity is undoubtedly welcomed by many users, but it introduces some significant difficulties as well, particularly with regard to accurate scoring.

For example, PR.AC-1 "Identities and credentials are issued, managed, verified, revoked, and audited for authorized devices, users and processes" covers a lot of very different ground. If, for example, an organization could be considered a "2" for the identity and credential issuance component, a "3" for identity and credential auditing, and a "1" for revocation, which score should be used for PR.AC-1? Averaging the scores is a common approach, but many controls (these for example) have Boolean AND relationships with one another, where the low score is the accurate overall characterization of the dependencies within the subcategory.

Clearly, adding greater specificity to the CSF would not be welcomed by some within the community. That said, if the additional granularity was arrived at carefully, and if the reasoning for the choices was made clear, most cybersecurity professionals would likely appreciate the resulting improvements in clarity and measurement quality. Certainly, organizations relying on the CSF would benefit in terms of a significantly more accurate understanding of their cybersecurity profile, and better prioritization.

Tier Scoring

NIST has chosen to not provide specific scoring criteria within the CSF under the belief that doing so would limit organization and industry flexibility in using the framework. Although this does indeed maximize flexibility, it also has some fairly significant downsides:

- Without clearly defined and consistently applied scoring criteria, meaningful and reliable benchmarking simply isn't possible. An exception to this is where a specific service provider has defined its own scoring criteria. In such a case, that service provider would be able to benchmark its

customers consistently. However, outside of that circumstance, an expectation of meaningful and reliable benchmarking is unrealistic.

- Contrary to NIST guidance, many organizations (and even some service and solution providers) try to use the existing CSF Tier scale to rate subcategories. The current CSF Tiers are defined to be used at the program level and are not suitable for accurately rating subcategories. This almost invariably results in unreliable subcategory scoring, which invariably misrepresents an organization's cybersecurity profile.
- Some organizations also try to use Tier scores applied at the subcategory level as inputs for risk measurement and prioritization. Here again, the existing Tier scores do not accurately characterize subcategory risk reduction efficacy, which means these organizations are extremely unlikely to get accurate risk measurements or make well-informed decisions regarding improvement priorities.
- Some organizations instead leverage CMMI-like scoring models for CSF subcategories. Although these scoring models are appropriate for characterizing process maturity, it is a considerable and often flawed leap of faith to believe maturity model scores translate accurately into control efficacy. Yes, higher levels of process maturity surrounding controls will often improve a control's efficacy, but that isn't remotely the same as efficacy.

Despite the angst some members of the community might have regarding a subcategory scoring model published by NIST, many organizations likely would choose to adopt such a model rather than trying to devise their own. At the very least it would provide a solid starting point for organizations that want to develop their own scoring model without starting from scratch.

The following provides an example of a potential subcategory-level scoring model:

- Tier 1 - The outcome is reliably achieved for less than 50% of the organization's environment.
- Tier 2 - The outcome is reliably achieved for between 50% and 75% of the organization's environment.
- Tier 3 - The outcome is reliably achieved for between 75% and 90% of the organization's environment.
- Tier 4 - The outcome is reliably achieved for more than 90% of the organization's environment.

There are a couple of important criteria that need to be defined for this model:

1. "Reliably achieved" means that where the relevant controls are applied (i.e., within the claimed coverage range), the control outcomes are realized more than 95% of the time. In other words, it isn't enough to simply "install" a control in 75% to 90% of an organization's environment, the control also has to be operating as intended at least 95% of the time.
2. In addition, all of the organization's most important assets must fall within the coverage range. For example, if the CSF subcategory outcome "Data at rest is protected" (PR.DS-1) is being realized in 75% to 90% of an organization's environment, but there are sensitive data at rest in the other 10% to 25% of the environment that are not appropriately protected, then that organization would not be able to claim a rating above Tier 1.

It's worth pointing out that these criteria (particularly the second one) are strongly associated with an organization's maturity, as the processes and characteristics generally associated with CMMI-like models ultimately drive control reliability. In other words, an organization operating at higher levels of CMMI maturity will experience higher levels control reliability. The implication here is that if an organization were to score itself at the subcategory level using this model, it would be relatively straightforward to accurately derive the organization's overall maturity based on its subcategory ratings. Killing two measurement birds with one stone, so to speak.

Note, too, that this type of scoring model can be used to evaluate subcategories "subjectively" (e.g., using subject matter expert estimates of the percentages) or more objectively based on empirical data, so it doesn't impose any particular "data availability" requirements.

The ranges also could be adjusted by organizations that feel the need to do so. This retains flexibility, while still providing greater objectivity and clarity, and allowing clear differentiation between organizations using a customized scale versus those using a common baseline.

Note, too, that a model of this nature does not affect the inherent flexibility that comes with CSF's "control outcomes" approach, as it doesn't impose specific controls.

Subcategory Assignments

Another opportunity for improvement would be to reassign some of the CSF Subcategories to different Functions. For example, it would be far more accurate to have PR.IP-4, PR.IP-9, and PR.IP-10 (all of which are recovery-focused) within the Recover function. This would enable more accurate characterization of organization capabilities and maturity, as well as reduce the potential for misusing subcategory scoring in risk measurements.

The second point above regarding risk measurements is particularly important due to the potential for inaccurate reporting and decision-making. Because practitioners commonly equate the Protect function with loss event "prevention", it's common to see organizations include PR scores as factors in estimating loss event probability. However, recovery control outcomes such as PR.IP-4, PR.IP-9, and PR.IP-10, affect loss magnitude (i.e., limit the impact). Consequently, when scores that are mis-assigned to the PR Function are included in this process the result is inaccurate measurement in both event probability and magnitude, which leads to inaccurate risk and profile reporting, and poor prioritization.

The recommendation is to carefully review the CSF subcategories specifically with an eye toward ensuring that their function assignments are accurate from a risk measurement perspective.

Final Note

The FAIR Institute genuinely appreciates having had the opportunity to offer our thoughts and recommendations on this important endeavor. We would be more than happy to answer any questions NIST might have regarding these recommendations, as well as to brainstorm with stakeholders on these topics.