PNNL Position:

(5) Guidelines for software integrity chains and provenance. See EO Sections 4(e)(ii, vi, and viii).

Documenting software supply chains is a challenging effort, as inputs to source code may come from a standard library, may be created fresh and new by the developer, or may be pulled from a public code share. However, documenting this information is critical to ensuring secure software both during creation and operation; sharing the information is also critical to enable lifecycle-long security for the end users. Before this information can be shared, however, there are two critical prerequisites:

1. **Format.** What data regarding the provenance of the software supply chain exists is currently in completely unique formats, with each vendor or manufacturer documenting and saving different things in different formats. Current efforts are underway to create a standard Software Bill of Materials or SBOM format for critical infrastructure industries, along with sister efforts to standardize Hardware Bill of Material and security testing data reporting templates. These efforts will require broad acceptance across multiple industries.
Data sharing platform. If software has been proven to be of high integrity, but the information which asserts this and the methods of proving such an allegation are not provided, the end user is forced to treat the software as if has low or no integrity. Integrity chains and provenance attestations must be shared.

A comprehensive and centralized data repository containing this information is crucial in both ensuring access for end users and enabling national security focused industry wide data analyses. In operational technology software, ensuring that all operators have the data access required may prevent a significant event such as large-scale power outages or nuclear plan shutdowns. All data should be consistent in a standard software bill of materials (SBOM) format and allow for interaction of analyses with relevant hardware in HBOM format; while a software-only analyses is valuable, it becomes much more actionable when paired with operational hardware.

A distributed data storage method such as blockchain has been proposed as an underlying structure for the sharing of this information between various actors. However, this would not be as effective or accessible as it increases challenges for end users when attempting ensure security across multiple vendors. A distributed data storage method also reduces the ability to streamline and standardize data input. However, a distributed ledger may be a viable alternative to a centralized data repository for ensuring historical accountability and enabling data sharing.

Finally, a trusted third party should be used to hold this data. An external party holding the information reduces bias in any vulnerabilities found and can also help in collating data in a standard format. For critical infrastructure software, the organization should also be responsible for performing larger-scale data analyses across the entire industry, to identify nationwide weaknesses.

Please direct any questions to jess.smith@pnnl.gov.

If requested, we would be happy to present this position and its supplementary information. The coordinating author will be:
Jess Smith, PhD
Jess.Smith@PNNL.gov
509 372 4213