

Trustworthy and Confidential SBOM Exchange

Eman Abu Ishgair[†] Chinenye Okafor[†] Marcela S. Melara[‡] Santiago Torres-Arias[†]

[†]Purdue University, {eabuishg, okafor1, santiagotorres}@purdue.edu

[‡]Intel Corporation, marcela.melara@intel.com

Abstract

Software Bills of Materials (SBOMs) have become a regulatory requirement for improving software supply chain security and trust by means of transparency regarding components that make up software artifacts. However, enterprise and regulated software vendors commonly wish to restrict who can view confidential software metadata recorded in their SBOMs due to intellectual property or security vulnerability information. To address this tension between transparency and confidentiality, we propose Petra, an SBOM exchange system that empowers software vendors to interoperably compose and distribute redacted SBOM data using selective encryption. Petra enables software consumers to search redacted SBOMs for answers to specific security questions without revealing information they are not authorized to access. Petra leverages a format-agnostic, tamper-evident SBOM representation to generate efficient and confidentiality-preserving integrity proofs, allowing interested parties to cryptographically audit and establish trust in redacted SBOMs. Exchanging redacted SBOMs in our Petra prototype requires less than 1 extra KB per SBOM, and SBOM decryption accounts for at most 1% of the performance overhead during an SBOM query.

1 Introduction

Millions of software vendors today build products that depend on open and closed source software components distributed through a myriad of channels; each component is, in turn, composed of various software dependencies. While open source software is typically available for inspection, accurately identifying the direct and transitive dependencies that are redistributed in a software product or a deployed system remains a difficult challenge [35, 83].

Motivated by large-scale software supply chain incidents such as the Log4Shell vulnerability [23] and the SolarWinds hack [78], interest in Software Bill of Materials (SBOM) has significantly increased for their ability to provide detailed documentation about the components in a software artifact.

Recent regulations [39, 84] cemented the need for generating SBOMs to enhance the transparency of the software supply chain, prompting software vendors to collect and disseminate SBOMs for their software products.

Yet, software vendors face three key challenges that hamper the adoption of SBOMs in their software development life-cycles. **1) Data confidentiality:** Enterprise and government practitioners have highlighted significant concerns about the disclosure of intellectual property or private data when sharing SBOMs with software consumers [28, 68, 69, 83, 88, 95]. **2) Trustworthy SBOM data:** Another common problem is the lack of standard mechanisms to verify the integrity of SBOM data and establish its trustworthiness [35, 83, 88]. **3) Interoperable exchange:** Additionally, current tools are largely unable to generate SBOMs that can be *composed and (re)distributed* in a standardized and efficient manner that enables cross-vendor interoperability [50, 88, 95].

Current approaches for secure and trustworthy SBOM exchange rely on contractual/legal processes like licensing and non-disclosure agreements [69], or disparate ad-hoc systems [35, 46, 88] (e.g., secure repositories with restricted access [41] or dedicated publish-subscribe systems [70]). While these approaches facilitate the sharing of a vendor’s SBOM, they offer limited scalability and interoperability for secure SBOM exchange because they focus narrowly on distributing SBOMs to *direct* software consumers. Thus, as SBOM data completeness is varied [83, 95], consumers further downstream the supply chain still face challenges discovering and trusting SBOMs for transitive dependencies.

To address these challenges, we propose Petra, an interoperable system for confidential SBOM exchange and trust. Petra is the first-of-its-kind SBOM exchange system to support interoperable composition and redistribution of SBOMs while enforcing confidentiality and enabling cryptographic auditing of redactions. Unlike existing SBOM exchange mechanisms, Petra preserves integrity and access control after SBOMs are redistributed across organizational boundaries.

SBOM exchange architecture for interoperability. At the foundation of Petra’s design is an abstract SBOM exchange

architecture that centers the ubiquitous practice of software composition and redistribution, i.e., the propagation of software components incorporated into other software products. Our model extends this paradigm to SBOMs, enabling the generation and exchange of *composable* SBOMs in a manner agnostic to SBOM data format or software ecosystem.

Petra achieves this goal at two levels. At the system level, Petra defines a set of abstract SBOM exchange roles and protocols that can be instantiated in existing software supply chains (§3). At the data level, Petra breaks down SBOM information into an *SBOM tree*, a tree-based data structure that mirrors the documented relationships between software components and their metadata (§4.1).

For instance, a software vendor *Acme* builds a software package `foo` that depends on package `bar`. With Petra, *Acme* will generate an SBOM tree for package `foo` that at a high level will have `foo` at its root node, and a child node for `bar`.

Petra SBOM trees enable efficient searching and *merging* of SBOM information while ensuring that software consumers can trace a particular dependency back to its origin. This allows Petra to support SBOMs that reflect the highly compositional nature of today’s software and retain transitive dependency information about an artifact within a single SBOM, obviating the need to distribute per-dependency SBOMs.

Confidentiality with selective redaction. SBOM trees allow Petra to balance the tension between transparency and confidentiality in a customizable way. We retain the basic model of SBOM generators creating SBOMs for software products, but enable vendors to control access to their own SBOM information through selective redaction, even if this information is contained within another vendor’s SBOM.

Specifically, leveraging attribute-based encryption, Petra is capable of redacting information at multiple levels of the SBOM tree (§4.2). Petra then captures organization-level requirements via namespace attributes (e.g., company domains).

Cryptographically establishing trust in SBOM data. While placing an entire SBOM within a single Merkle tree node (analogous to binary transparency [2], for example) would enable detection of unexpectedly issued SBOMs, such a structure would still need ad-hoc mechanisms for data lookups within SBOMs, SBOM composition, and auditable SBOM re-distribution. Thus, to detect modifications [28, 88] of the contents of an SBOM, even when redacted, Petra computes an *authenticated Merkle tree* on top of an SBOM tree (§4.3).

c While verifying the integrity of a Merkle SBOM tree requires reconstructing the entire structure to validate its root, users may also delegate this operation to trusted verifiers, not unlike monitors in Certificate Transparency [52]. On the other hand, the Merkle SBOM tree also allows Petra to efficiently prove *sameness* between redacted and plaintext SBOM data while maintaining the confidentiality of the remainder of the tree’s contents.

Checking stronger *completeness* and *accuracy* properties of SBOM information is considerably harder to achieve in

an interoperable, data-agnostic manner as this requires cross-referencing the contents of an SBOM with various application-aware data sources [28]. Thus, Petra establishes cryptographic integrity as the foundation for trustworthy SBOMs.

Petra implementation. We instantiate Petra in a prototype system, which includes an ecosystem-agnostic redacted SBOM generator, as well as a client-side application for SBOM consumers and verifiers. The generator implements Petra’s SBOM tree, selective redaction based on CP-ABE [27] and sameness proof generation, while the client-side tool supports queries on redacted SBOMs and proof verification. A separate verifier component showcases Petra Merkle SBOM tree integrity checking. A Fulcio-backed [65, 80] key management server handles signing and redaction keys for generators and consumers. We study the storage overhead of our evaluation and find that it adds acceptable storage overhead — 13% compared to their non-encrypted counterparts. Similarly, encryption, decryption, and membership routines add less than one second overhead for average SBOM sizes, and two seconds at a maximum.

2 Related Work

2.1 Software Supply Chain Security & SBOMs

The last half-decade has evidenced a widespread exploration of supply chain security techniques from industry, academia and government. Various efforts aim to provide security enhancements through transparency, integrity, and separation [72]. Frameworks such as in-toto [86] provide end-to-end integrity auditing across phases of the software development lifecycle. Other frameworks like The Update Framework (TUF) [51], CHAINIAC [66], and Sigstore [65] focus on ensuring the integrity of distributed software.

Binary transparency [24] provides cryptographic commitments to software binaries on a public, append-only log, making it easier to detect misbehavior. In contrast, Petra aims to balance transparency and confidentiality of software through selective redaction and cryptographic integrity.

Software Bills of Materials (SBOMs) have been recognized as a critical aspect of software supply chain security as they provide visibility into the components of a software product. SBOMs are meant to enable informed decision-making on security risks, license compliance, and overall software quality. Government organizations have developed standards for SBOMs, outlining minimum elements for dependency tracking [63] and defining various SBOM types [34] to address varying needs at different stages of the software lifecycle.

Although SBOMs are now recognized as crucial for software supply chain security, their wide-scale adoption in practice remains limited. A 2022 study by the Linux Foundation revealed that only 18% of organizations use SBOMs extensively across their operations [46]. A separate study focusing on open source projects on GitHub, where only 186 projects

utilizing SBOM generation tools were found [67].

This limited adoption has been attributed to social factors like unclear benefits and use cases [31,59] and a lack of awareness about SBOMs [67], concerns about sensitive information disclosure [35,88,95] and the lack of effective SBOM analysis tooling [83,88]. Our approach in Petra specifically seeks to address unwanted information disclosure.

Interoperable SBOM exchange. Compatibility issues between SBOM generators and consumers arise because tooling has inconsistent support for the two common SBOM data exchange formats SPDX [18] and CycloneDX [7]. To address these issues, open source initiatives such as the Open Source Security Foundation’s ProtoBOM [14] and bomctl [3] projects seek to provide mappings between fields on either standard to allow for translation.

Although limited options for standardized SBOM exchange platforms hinder interoperability across tools and vendors [28,88], efficient SBOM and supply chain metadata exchange remains a relatively understudied area. Frameworks like the CycloneDX SBOM Exchange API [73], Federated PURL (package URL) [76], DBOM [8] and GUAC [10] exist to facilitate metadata exchange, but there is a lack of uniformity and established sharing mechanisms [35,89] and academic literature exploring this problem remains limited. Approaches such as Manufacturer Usage Descriptions (MUD) for IoT devices [53] exist to share information with consumers, but have limited broad deployments.

Petra builds upon existing efforts on interoperable tools that handle SBOM data exchange [50,88,95]. However, existing SBOM exchange frameworks fundamentally differ from Petra in the security guarantees they provide. Systems such as the CycloneDX SBOM Exchange API, GUAC, DBOM, and related metadata exchange platforms primarily address interoperability and metadata exchange, but offer little support for confidentiality, auditable redaction, or composition of SBOMs originating from different vendors. Relying on repository access controls or API authentication, these alternatives to Petra cannot guarantee integrity once SBOMs are redistributed. Petra addresses this gap by binding selective disclosure and integrity guarantees directly to the SBOM structure itself, enabling confidentiality- and integrity-preserving SBOM composition and redistribution without relying on trusted repositories or centralized enforcement.

Much like ProtoBOM, Petra is format-agnostic and uses a graph-based SBOM representation and transformations, supporting both SPDX and CycloneDX formats. The OWASP Transparency Exchange API (TEA) [73] provides SBOM data "on the fly," effectively acting as a generator for a particular stream. In practice, Petra complements the current tools, and we have engaged with the ProtoBOM developers to integrate its redaction features in bomctl [4].

SBOM quality & accuracy. Studies have raised concerns about the quality and accuracy of SBOMs due to incomplete data [50], inconsistent results from different tools [75],

and a lack of validation mechanisms [50]. These inconsistencies were attributed to reliance on heuristic methods, incomplete package manager metadata, and difficulties in handling statically linked binaries during SBOM generation [95] compounded with unclear definitions of SBOM quality [87]. While Petra does not aim to address these issues, we focus on integrity and confidentiality as foundational preconditions for quality, because SBOM quality and accuracy cannot be meaningfully established without trust in the authenticity and protection of SBOMs.

2.2 Privacy-preserving Data Exchange

Privacy-enhancing techniques have been a central topic in cryptography for many decades. Various approaches to secretly share private information have been widely explored in fields such as Private Information Retrieval (PIR) [92], as well as Homomorphic Encryption (HE) [32] and Secure Multiparty Computation (SMC) [26]. There are applications of these in the area of supply chain security [30,93]. Similarly, other supply chain security systems that further privacy are Speranza [61] and Gringotts [91], which protect software distribution security and privacy and version control system security and privacy respectively.

As for non-cryptographic approaches, a promising alternative is Trusted Execution Environments (TEE [22,49]), which can allow to encrypt SBOMs wholesale and provide access to users using a messaging interface. Perhaps the most important distinction is that attribute-based encryption (ABE) allows Petra to not require a live service to re-distribute SBOM information, and instead leans on key management to issue keys to parties in the ecosystem.

Applications of CP-ABE Several works have applied Ciphertext-Policy Attribute-Based Encryption (CP-ABE) in other domains like cloud data sharing [48,54,55], electronic health records [38], and IoT/edge computing [45,90]. These systems highlight the applicability of CP-ABE in practical settings in real world environments. In the SSC domain, a concurrent work [89] proposes a blockchain-based SBOM exchange architecture that uses verifiable credentials to enable verification of SBOM generation, and ABE for selective SBOM data disclosure. More specifically, [89] introduces an oversight authority to determine generators’ adherence to SBOM regulations, and rely on various on- and off-chain components to coordinate SBOM generation and distribution.

While Petra shares the goal of selectively disclosing SBOM data, our Merkle SBOM tree-based design is meant to integrate and interoperate with existing tools and software supply chain components, rather than requiring dedicated and costly blockchain-based infrastructure to provide integrity checking. As such, Petra can offer software producers additional flexibility in defining SBOM access control policies.

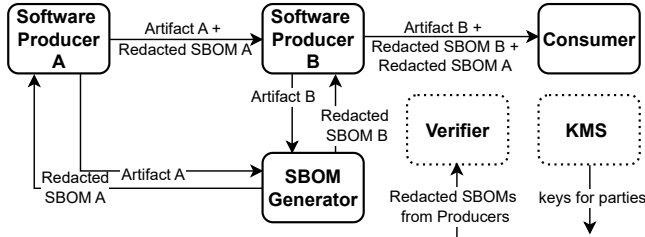


Figure 1: Abstract SBOM exchange system. A generator creates redacted SBOMs according to producers policies’, which are then redistributed along artifacts to other producers or consumers through the redistribution chain. We omitted distributors for simplicity.

3 System and Threat Model

Petra aims to provide a SBOM exchange system that facilitates trustworthy and interoperable SBOM generation and distribution that meets the diverse confidentiality needs of today’s software vendors.

SBOM Definition. Per the U.S. NTIA [62], we define a Software Bill of Materials (SBOM) as a record that unambiguously describes a software artifact’s components, their relationships, and relevant metadata, such as vulnerability information, each dependency’s license, and version. Thus, the baseline data fields used to describe a software artifact in an SBOM include supplier name, component name, component version, unique identifiers, dependency relationships and SBOM author [62]. In Petra, we assume that SBOMs follow a given schema which specifies the *fields* that contain the different elements of an SBOM. The common schemas used to generate and consume SBOMs are SPDX [18], CycloneDX [7], and SWID tags [17]. We include a snippet of an SPDX SBOM in Appendix A.

3.1 SBOM Exchange Roles and Processes

As depicted in Fig. 1, we propose an abstract SBOM exchange system involving the following roles and processes:

Software Producer. An important prerequisite for SBOM exchange is the creation of software artifacts by software producers. In practice, a software producer may be an individual engineer on an open source project, or a larger organization such as an enterprise or government entity.

Generally, the software producer aims to release the software artifact publicly, to a select set of recipients, or internally within an organization. Thus, depending on the publication scenario, the software producer may seek to restrict access to the software artifact as a whole, or to (portions of) the software’s metadata, including its SBOM. In confidential SBOM exchange, the producer is responsible for authoring the redaction policy that governs which parts of the SBOM are visible to whom. This policy encodes confidentiality requirements

based on vendor preferences or regulatory obligations.

Generator. Following the creation of a software artifact, the SBOM generator creates an SBOM that lists the components and related metadata of a given software artifact on behalf of its software producer. In a confidential SBOM exchange system, the generator also applies and preserves the producer’s SBOM data access policy requirements, even if the SBOM is composed with another and/or redistributed at a later time.

The separation of generator and producer is deliberate: Petra aims to accommodate two different SBOM generation use cases. First, there are practitioners for whom a single SBOM generator is not sufficient to produce a comprehensive and composed SBOM for their software, as evidenced by the CISA SBOM types [34], where multiple generators record different elements of an SBOM document. Second, and more crucially, is the plethora of third-party generator tools and services such as FOSSA, GitHub, and others (e.g., [25, 29, 42, 82, 85]). Thus, the explicit generator role of Petra offers a flexible architecture, but this design choice does not preclude merging of the two roles in specific implementations.

Distributor. Once an SBOM is generated, the software producer or generator submits the SBOM to a distributor making it available for discovery. To this end, the distributor provides a system for SBOM storage and management, which commonly allows for retrieving an SBOM for a target artifact, e.g., via a global identifier such as a pURL [76], a standardized identifier format used across major SBOM standards (e.g., `pkg:npm/express@4.17.1`). Such SBOM databases or hubs (e.g., [36, 37, 56]) may also support analysis queries on stored SBOM information without delivering a full SBOM document.

Alternatively, SBOM distribution occurs when an SBOM is delivered with the software artifact (i.e., in-band distribution). For instance, the distributor is hosted on the same platform that delivers software artifacts, such as social coding platforms [5] and package repositories [12]. Similarly, if a software producer distributes an SBOM along with their software artifact, they act as a distributor.

Consumers. Individual software developers and organizations consume software artifacts and their associated SBOMs, either for direct deployment or to integrate an artifact as a dependency and redistribute it in their own software product. In both cases, consumers leverage SBOMs to gain security insights about consumed software by interacting with distributors to obtain the needed SBOM documents.

When acting as a software producer, the consumer creates a software artifact that is composed of various dependencies, resulting in a *composed SBOM* from a generator. That is, say the consumer’s software package `foo` includes dependency package `bar`, the SBOM S_{foo} should contain SBOM S_{bar} .

Verifier. To ensure that SBOM information is trustworthy, verifiers obtain SBOMs from distributors for the purpose of validating their format and contents. Specifically, verifiers check that a distributed SBOM was not tampered with and

matches the output of a generator, as well as checking whether an SBOM reflects a software artifact’s composition.

While other roles may act as verifiers, we envision trusted third-party verifiers will commonly perform this task on their behalf. For example, a government agency or dedicated auditor organization may act as a verifier checking individual SBOMs or portions of the software supply chain.

Key Management Service (KMS). To reduce the burden of key management for other participants, the KMS acts as a trusted central authority for key generation, user onboarding, key distribution, revocation, and expiration.

3.2 Threat Model

Confidential SBOM exchange is concerned with the following threats, labeled as (T1–T7) for reference in § 7.

Software producers and generators are untrusted as they may be driven by business incentives to circumvent security or compliance audits for targeted consumers. That is, software producers and generators may **tamper with SBOM data (T1)** [74], either by omitting certain fields such as vulnerabilities, or altering fields such as dependency versions. Similarly, distributors are untrusted in confidential SBOM exchange as they might tamper with stored SBOM data to undermine trust in the software supply chain of a major software vendor, or to give false sense of security regarding a software artifact.

Distributors and generators may also seek to mount a **split view attack (T2)** [58] in an attempt to fool targeted consumers into using a compromised artifact. Specifically, generators may create inconsistent copies of redacted SBOMs for the same software product, while distributors may provide modified or outdated copies of an SBOM to different consumers. In the case of SBOM redistribution, malicious distributors or consumers may cause a denial-of-service (DoS) by **corrupting SBOMs redistributed to downstream consumers (T3)** [66].

Consumers have also been shown to leverage SBOM information, such as package versions and reported vulnerabilities, for malicious purposes [35, 50, 88, 95]. Hence, software producers are concerned with **unauthorized SBOM data leaks (T4)** by untrusted consumers.

Consumers may further seek to circumvent any access restrictions or redaction on SBOM information put in place by generators. That is, consumers with partial access to SBOMs may attempt a **dictionary attack (T5)** [77] on encrypted fields to infer their values. Additionally, consumers with different access permissions to an SBOM may **collude to decrypt inaccessible information (T6)** [27] about a competitor’s products in an SBOM, or collude to share SBOM data they can each access to **infer inaccessible SBOM data (T7)** [64].

Trusted roles and components. Verifiers are trusted to faithfully audit redacted SBOMs. Similarly, we assume Key Management Services (KMS) faithfully authenticate principals, bind attributes, issue keys, and enforce revocation/ex-

piry. KMSes can leverage existing key transparency mechanisms [57, 58, 65] for additional cryptographic integrity and compromise detection. Further, we assume all parties employ cryptographic schemes and compute systems (hardware and OS) that are free of known vulnerabilities.

Out of Scope Threats. Detecting maliciously crafted SBOMs that conceal malicious components is out of scope, as is detecting malicious components. These threats are the subject of extensive prior work [33, 40, 71, 79, 96]. Similarly, detecting flaws with generator tools and processes (e.g., those who may introduce mistakes) is out of scope [94]. Confidential SBOM exchange also does not attempt to evaluate whether SBOM access policies are excessively restrictive. Detecting or mitigating overly strict is a dedicated research area [89] and requires separate governance or policy-auditing mechanisms.

3.3 Design Requirements

Derived from our architecture and threat model, a confidential exchange system should meet the following functional, confidentiality, and integrity requirements.

Functional Requirements.

(F1) Relational SBOM Representation: SBOMs must be represented using a data structure that preserves the inherent relationships between all software artifacts and associated metadata recorded in an SBOM, regardless of the software ecosystem or data exchange format. This supports composition and redaction operations on SBOM data without losing the relational information needed for meaningful software dependency analysis.

(F2) SBOM Composability: To reflect software composition, SBOMs must be composable. That is, leveraging the relational SBOM representation property, if software artifact B depends on a third-party component A , SBOM S_A can be inserted into the contents of SBOM S_B .

(F3) Indexability: To support interoperable distribution of SBOMs, they must be associated with specific software artifacts via immutable indexes.

Confidentiality Requirements.

(C1) Semantic confidentiality: To address a core barrier to practical SBOM adoption in enterprise and government settings, SBOMs must not leak SBOM contents to parties without explicit permission to access specific pieces of dependency or metadata information. This property must hold even during SBOM composition or redistribution. In other words, software producers must be able to compose and redistribute SBOM contents without requiring access to any information contained therein.

(C2) Collusion Resistance: As consumers may attempt to circumvent semantic confidentiality by colluding to reveal redacted SBOM data, the selective redaction mechanism must be capable of resisting collusion between parties that

have access to only a subset of the needed permissions.

Integrity Requirements.

(I1) Sameness: Consumers of a software artifact must be able to verify that redacted SBOM information faithfully represents and does not omit information from the *plaintext SBOM*, even for redacted data that is only partially accessible.

Specifically, the original relational SBOM structure must be maintained while ensuring that only the information accessible to a verifier is revealed. In practice, while access restrictions would allow most consumers to only partially verify sameness, we expect that software producers and select verifiers are permitted to access the full plaintext SBOM and validate full-SBOM sameness.

(I2) Redistribution Congruence: To support composability and interoperable distribution of SBOMs, a redistributed SBOM must retain its sameness, and preserve semantic confidentiality as well as collusion resistance. That is, an SBOM S_A distributed for software artifact A must remain identical to the SBOM S_A redistributed in SBOM S_B for a larger software artifact B that depends on A .

(I3) Non-equivocation: Consumers must be able to verify they received the same SBOM as others for a given artifact.

4 Foundational Data Structures

We instantiate our confidential SBOM exchange architecture in Petra. At its core, Petra uses a series of data structures to achieve our design requirements §3.3. Through a three-stage SBOM process, we use these data structures to transform a textual SBOM document into an auditable, selectively redacted SBOM representation.

4.1 SBOM Tree

The first stage in Petra translates an SBOM document (e.g., in SPDX [18] or CycloneDX [7] format) into a tree data structure that is suitable for efficient redaction and integrity checking. We draw upon the insight that an SBOM is a collection of claims about a software artifact. Namely, each SBOM describes an aspect of a software product, such as qualitative (e.g., vulnerabilities), legal and compliance (e.g., licenses) and functional (e.g., configurations).

Using this notion, we first *decompose* an SBOM into a series of nodes that describe these claims, and construct an *SBOM tree* that represents the relationships between the elements in the SBOM. We distinguish two types of tree nodes: *Complex nodes* hold *composite* claims, e.g., the collection of metadata about a dependency recorded in an SBOM. These nodes allow Petra to decompose claims further into a set of *field nodes*, and provide nesting.

Specifically, a field node represents exactly a single claim in an SBOM, such as a license claim, and is inserted as leaf nodes into an SBOM tree with the corresponding complex

node as its parent in the tree. For example, a dependency within an SBOM can be decomposed on a series of field nodes, one for each type of metadata about a dependency such as its name and version number.

At the root of the SBOM tree, Petra constructs an *SBOM node* containing a unique indexable identifier for the SBOM to aid in its discovery. The leftmost tree in Figure 2 (a) represents an SBOM tree containing another SBOM node.

This construction allows us to meet property (F1), since the structural properties of the tree describe the inherent semantics of the SBOM. Further, through this complex node decomposition, Petra can provide property (F2) since a consumer can insert an SBOM within another SBOM by adding it as a complex node within their own SBOM tree.

4.2 Attribute-Based Redaction

To provide selective redaction of SBOM information, Petra utilizes a variant of Attribute-Based Encryption (ABE) [44] called Ciphertext-Policy Attribute-Based Encryption (CP-ABE) [27], which enables fine-grained access control via selectively encrypted data. Unlike traditional public-key encryption, which uses a predefined key for decryption, ABE determines decryption keys based on *attributes* associated with principals and the data itself, effectively enforcing access control on encrypted data. These attributes are opaque strings/byte arrays, and encode information about consumers (e.g., role, organization, location).

The CP-ABE scheme used in Petra builds on vanilla ABE schemes by passing the access policy as an encryption parameter. The policy is then embedded within the ciphertext, while consumer attributes are embedded in decryption keys. This gives software producers the flexibility to define who can access the data based on the attributes of different consumers.

Thus, Petra redaction relies on the following operations:

1. $\text{Setup}()$: Generates the public encryption parameter P_k and a master decryption key M_k .
2. $\text{Encrypt}(P_k, M, A)$: Given the public parameter P_k , a message M , and an access policy A , produces a ciphertext.
3. $\text{Key Generation}(M_k, S)$: Derives a secret decryption key based on a set of consumer attributes S from the master key M_k .
4. $\text{Decrypt}(P_k, CT, S_k)$: Given the public parameters P_k , a ciphertext CT , and a consumer's secret key S_k , produces a plaintext message only if the consumer's attributes satisfy the access policy embedded within the ciphertext.

CP-ABE on SBOM trees. To provide semantic confidentiality (C1), the second step in Petra performs a CP-ABE pass to selectively encrypt individual SBOM tree nodes. Specifically, we attach an access policy to each node and perform an SBOM tree walk to encrypt the SBOM data in each node,

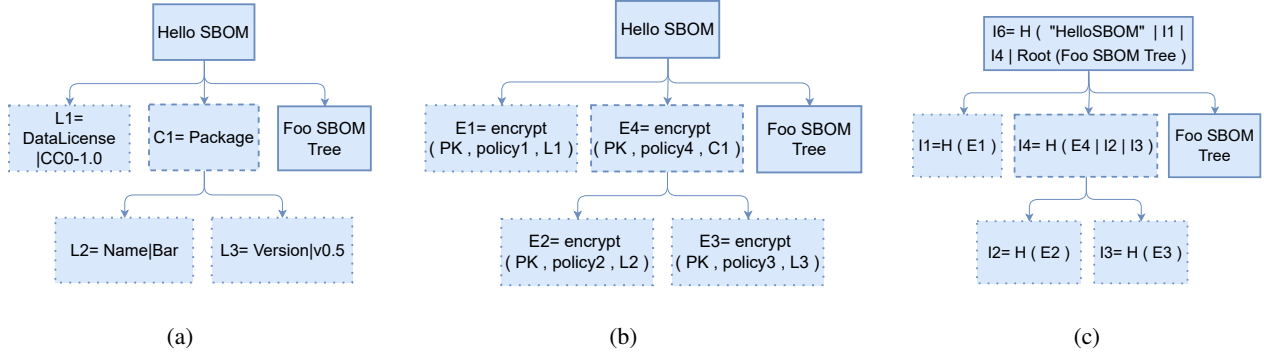


Figure 2: The Petra SBOM tree is an efficient representation of an (a) SBOM that preserves relational information between dependencies and their metadata. This representation also allows Petra to (b) selectively redact and (c) integrity check SBOMs within a single data structure.

namely field names and values, according to the node’s policy; multiple nodes may be redacted under the same policy. We do not encrypt the references to the children of complex nodes, as we still aim to expose the structural relationships of the SBOM. This allows Petra to maintain the functional properties above as well as support SBOM composition.

To reduce the cost of CP-ABE encryption and decryption, we utilize CP-ABE for key encapsulation (CP-ABKEM) [43]. Instead of directly encrypting each node with a CP-ABE secret key, we generate a random AES key for each distinct node access policy, and use the AES key to encrypt all nodes the policy applies to. Each AES key is then encrypted using CP-ABE under the policy. This ensures that only consumers who meet a given policy can decrypt the corresponding AES key and, in turn, decrypt the data. Using CP-ABKEM, we leverage the fine-grained access of CP-ABE, and the efficiency of AES. Figure 2b represents the SBOM tree after encryption.

The choice of CP-ABE is crucial to meet (C2), since the scheme requires a decryption key to hold all the attributes needed in order to decrypt a node. We elaborate further on these security properties in the security analysis (§7).

4.3 Merkle SBOM Trees

The third stage in Petra computes cryptographic integrity checks to detect tampering with pre- and post-redaction SBOM trees. This is done by adapting Merkle tree primitives [60] to our irregularly structured SBOM trees.

A Merkle tree is a tree data structure where each leaf node is referenced by the cryptographic hash of their data block (including children). Each non-leaf node, on the other hand, combines the hashes of its child nodes. The tree root, known as the Merkle root hash, represents the hash of all the data in the tree.

Merkle trees can be used to represent authenticated graphs or trees thanks to the fact that each element’s identifier is recursively comprised of the hashes of content and children nodes.

This property enables Petra to provide indexability (F3) as the Merkle root hash can serve as a content-addressable identifier over an entire (redacted) SBOM [6].

Petra achieves sameness (I1) by applying Merkle tree primitives twice. First, we compute Merkle hashes over cryptographic commitments (in our design, a hash with a random salt) to the SBOM tree prior to redaction. The commitments are used to address the limited keyspace of some SBOM fields and values. In a second Merkle pass over an attribute-based redacted SBOM tree, Petra captures hashes used to detect tampering with the redacted tree.

Membership Proofs. Petra uses Merkle trees to compute an efficient proof that a particular node is contained within an SBOM tree based on its Merkle root hash. A prover computes an *authentication path*, which consists of the hashes of the sibling nodes along the path from the node in question to the root [58]. To verify whether a node is a member of the tree, a verifier sequentially recomputes the hashes combined with the sibling hashes, starting from the node in question up to the root, effectively recreating the Merkle root hash. If the computed root hash matches the expected Merkle root, the node must be in the tree.

Thus, Petra uses membership proofs to validate redistribution congruence (I2) without requiring full decryption or reprocessing. That is, during SBOM composition, Petra includes the Merkle root of the inserted SBOM tree as a complex node in the parent SBOM tree. A membership proof is then generated and embedded into the complex node, linking the inserted redacted SBOM Merkle root to the parent SBOM Merkle structure. This allows downstream consumers to verify that the embedded SBOM remains intact and matches its original redacted version (I1).

4.4 Building Merkle Redacted SBOM Trees

In summary, prior to selective redaction, Petra computes the Merkle hashes over the commitments to the plaintext data

in an SBOM tree using a collision-resistant hash function $H()$. Then, we construct a selectively redacted SBOM tree by encrypting the SBOM tree nodes using a per-node access policy AES key K_A , which is then encapsulated using CP-ABKEM via the CP-ABE $Encrypt(P_k, K_A, A)$ operation; for brevity, we use the shorthand $E()$ to represent AES-based node encryption. In a third pass, Petra computes the Merkle hash over the redacted tree node contents.

The construction is as follows:

Field nodes containing a field *name* and field *value* pair are hashed as:

$$h_{F_{plain}} = Commit(name||value)$$

prior to redaction, and afterwards as:

$$h_F = H(R||A_n||E(salt||name||value)||h_{F_{plain}})$$

where R is a constant value used to indicate a redacted node and A_n represents the node's specific CP-ABE access policy to cryptographically bind a node's redaction policy to its contents. $salt$ is the hashed random number used to compute the cryptographic commitment to the plaintext.

Even though including the access policy A_n results in some information leakage, an observer may at most *infer* the field name from the attributes allowed by the access policy, but cannot learn the value of the field.

Complex nodes represent an artifact-level SBOM element of type t that has one or more field nodes F_0 to F_n as its children. They are computed as:

$$h_{C_{plain}} = H(Commit(t)||h_{F_0_{plain}}||\dots||h_{F_n_{plain}})$$

representing the pre-redaction SBOM element relationships, with post-redaction ones captured as:

$$h_C = H(R||A_n||E(salt||t)||h_{C_{plain}}||h_{F_0}||\dots||h_{F_n})$$

Because Petra does not aim to hide the structure or relationships between elements in an SBOM to support sameness verification, we can reveal the complex node's access structure and child nodes without leaking specific field values.

SBOM nodes identify an SBOM document and may have field, complex or other SBOM nodes as its children C_0 to C_n . Redaction of SBOM node contents depends on whether the software producer wishes to control access to SBOM-level data such as its resolvable index or redaction policy.

Regardless, Petra hashes the contents of SBOM nodes, redacted or not, to allow for integrity checking, authentication, but also indexability of the SBOM. Their hash is:

$$h_{S_{plain}} = H(Commit(r)||h_{C_0_{plain}}||\dots||h_{C_n_{plain}})$$

where r is the plaintext of any redactable SBOM data. After redaction, an SBOM node is structured as:

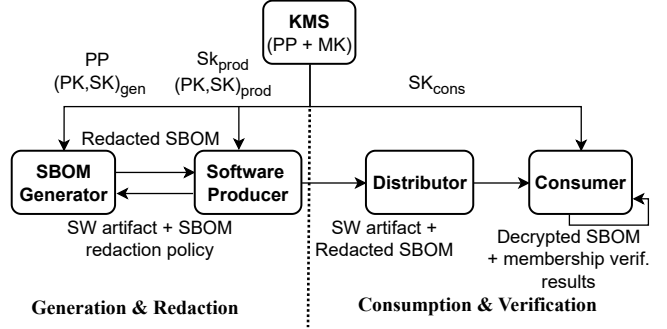


Figure 3: In Petra, a central KMS generates their redaction and/or signing keys. A producer sends the artifact and its redaction policy to a generator, which returns both the pre-redaction plaintext SBOM and the redacted SBOM for cryptographic sameness verification. Consumers fetch the artifact and redacted SBOM from distributors, obtain their decryption key from the KMS, and decrypt only authorized fields.

$$node_S = Sig_{K_s}(H(Encrypt_{all}(P_k, K_{A_n}, A_n)||h_{S_{plain}}||h_{C_0}||\dots||h_{C_n}))$$

where K_s is the generator's SBOM signing key, and $Encrypt_{all}(P_k, K_A, A)$ is the CP-ABE encryption of the entire set of per-policy AES encryption keys K_A .

Unredacted nodes. In practice, we expect SBOM redaction policies that allow for certain SBOM fields to be publicly visible. In this case, Petra replaces the R redaction indicator with a P constant indicating an unredacted node, computing the Merkle hash over the plaintext field or complex node data directly post-redaction.

5 Petra Processes

This section presents the following core Petra protocols represented in Figure 3.

5.1 Redaction Setup

Before redaction, the KMS generates a series of cryptographic secrets needed for selective encryption/decryption of individual SBOM fields; these are directly derived from the CP-ABE construction in §4.2.

Generators Secrets. At startup, the KMS runs CP-ABE setup to generate the public parameters PP used in CP-ABE encryption and decryption, and the master key (MK_{kms}) used to derive decryption keys. The KMS also generates the generator signing key pair (SK_{gen}, PK_{gen}) used to authenticate generated SBOM trees.

Producers Secrets. The KMS generates the producer signing key pair (SK_{prod}, PK_{prod}) used to countersign generated SBOM information, in practice, this can be the same signing

key used to authenticate/sign SBOM information or attestations in existing ecosystems [1, 16]. The KMS makes the public parameter PP , the generator’s public key PK_{gen} , the producer’s public key PK_{prod} available to all Petra parties.

Consumer Secrets and Onboarding. To access redacted SBOM data, software consumers in Petra must obtain decryption keys associated with their access attributes. To do so, the consumer authenticates herself and her attributes A to the KMS (e.g., using OIDC [81]). The KMS issues a private decryption key SK_A associated with the presented set of attributes A , if the consumer can be authenticated. The KMS performs key update process for all non-revoked users, such that decryption keys are updated after T_{end} .

$$SK_A \leftarrow \text{CPABE.KeyGen}(PP, MK_{kms}, A \cup \{\text{expiry} : T_{end}\})$$

Petra also makes it possible for the KMS to delegate decryption key issuance to consumers with specific attributes D (i.e., a dedicated secret key SK_D) to issue a secret sub-key for a subset of their attributes. This allows consumers to grant access to a particular subset of attributes (or fields) for members of their organization.

5.2 Redacted SBOM Generation

Petra extends non-confidential SBOM generation to support selective redaction and SBOM data integrity checking. That is, Petra relies on legacy SBOM generation tools (e.g., Trivy [21]) that evaluate a particular software supply chain step or artifact to generate an SBOM that follows a common format such as SPDX or CycloneDX. In this case, a generator receives a set of artifacts and dependency information from the software producer to generate a corresponding SBOM.

After generating legacy SBOM data, the Petra generator performs a series of passes over it. As depicted in Algorithm 1, a software producer submits an *Artifact*, a series of SBOMs S for dependencies, and a redaction policy p for the artifact. After computing the redacted SBOM, the generator provides the redacted tree to the producer.

Defining the access policy. Recall that in CP-ABE, permissions are described by an *access tree*, where nodes of the access structure are composed of threshold gates and the leaves describe access attributes where AND gates can be constructed as n-of-n threshold gates and OR gates as 1-of-n threshold gates. Access attributes are a set of descriptive properties or credentials for a principal, which will be associated with the issued secret key, for example, the user’s role in an organization or project, or an SBOM usage application (vulnerability scanner, license scanner). Fig. 4 represents an access tree for the version number field of dependencies, such that version numbers across all dependencies can be accessed by a FEDRAMP [9] approved vulnerability scanner, an auditor or a federal agency.

In Petra, a policy is a combination of SBOM tree paths and corresponding access trees. Specifically, a software producer

Algorithm 1: Redaction and Composition

```

Data: Artifact
         $S \leftarrow S_1, S_2, \dots, S_n$ 
         $p \leftarrow \text{Policy}$ 

Result: PlainSBOM, RedactedSBOM

1 /* Compose SBOM into tree */
    $\text{PlainSBOM} \leftarrow \text{new SbmTree}()$ ;
2 for  $S_i \in S$  do
3    $st \leftarrow \text{new SbmTree}(S_i)$ 
4    $\text{PlainSBOM.appendChild}(st)$ 
5 end
6  $\text{proof} \leftarrow \text{computeSamenessProof}(\text{PlainSBOM})$ 
7 /* Apply Encryption using policy  $p$  */
    $\text{EncSBOM} \leftarrow \text{Encrypt}(\text{PlainSBOM}, p)$ 
8 /* Merkelize Tree & Sign */
    $\text{RedactedSBOM} \leftarrow \text{MHT}(\text{EncSBOM}, \text{proof})$ 
9  $\text{Sign}(\text{PlainSBOM}, SK_{gen})$ 
10  $\text{Sign}(\text{RedactedSBOM}, SK_{gen})$ 

```

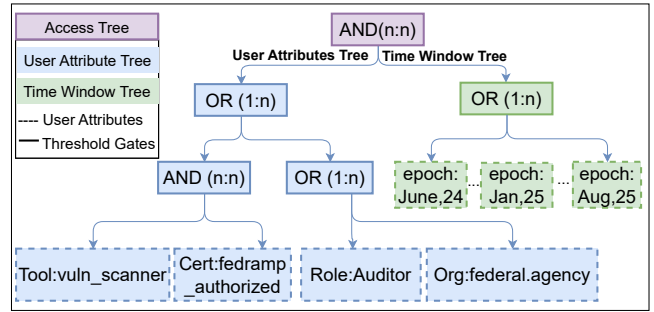


Figure 4: Petra Access tree. Two subtrees are used, one for properties tied to identities and roles, and a second one providing a time-bound vector.

can describe a policy for the tree by describing an access tree policy for a set of fields. For example, a software producer may allow a vulnerability scanner to access package version numbers in a redacted SBOM tree, only if the scanner is FEDRAMP approved. However, a consumer with licensing compliance attributes does not need to know version numbers, but may need to know package names and their corresponding licenses.

Verifying sameness and counter-signing. Once SBOM generation is complete, the software producer ensures that the redacted SBOM adheres to the expected access policy over the plaintext SBOM. This is done by comparing the decrypted SBOM tree to the redacted tree using the embedded sameness proof. If this sameness proof check passes, the producer countersigns the redacted SBOM prior to distribution, allowing consumers to gain confidence that the SBOM is a faithful representation of the expected SBOM information, as generated by the generator, and endorsed by the producer.

5.3 SBOM Distribution

A producer distributes the redacted SBOM by providing the distributor a resolvable index value (e.g., a pURL) to allow consumers to discover the SBOM for the artifacts of interest. In addition to receiving an SBOM, a Petra distributor can also verify the signature and counter-signature from the generator and producer, respectively, to ensure that only authenticated SBOMs are available to consumers.

5.4 SBOM Consumption & Verification

Consuming and verifying an in Petra SBOM requires a series of verification passes as depicted in Alg. 2. First, a consumer must check the authenticity of the *RedactedSBOM* by verifying both the signature and counter-signature using their public keys PK_{gen} and PK_{prod} , respectively (lines 1-6). After, the AES keys for which the consumer has access with her key SK_A are unpacked from the top-level SBOM node (line 7). Then, the consumer performs a full traversal of the redacted SBOM tree, evaluating each node’s ciphertext policy against its attribute key. Each SBOM node embeds a CP-ABE access tree in its policy field, and the check $satisfies(node.policy)$ determines whether the attributes bound to the consumer’s secret key SK_A satisfy this structure. Each node whose policy is satisfied is decrypted using the corresponding AES key, while all others are retained in encrypted form, yielding an SBOM tree containing only the consumer-authorized plaintext (lines 8–16). Lastly, a consumer can verify a membership proof for a specified decrypted SBOM field f to ensure the generator did not omit information.

Verifier pass. In practice, verifiers may represent trusted parties that have permission to access a whole redacted SBOM. By appending an OR Verifier with verifier being an access subtree representing the access attributes of a, a trusted third-party, such as an auditor or government agency, can be allowed to decrypt the SBOM and verify its sameness and redistribution congruence integrity properties.

5.5 Re-distribution & Composition

Lastly, a redacted SBOM can be also included inside of *another* SBOM — a crucial use case for the confidential SBOM exchange problem, as artifacts are composed and distributed through the software supply chain. Perhaps the most notable composition example is that of dependencies being included inside of another artifact.

Diving deeper into the composition step outlined in Algorithm 1, Petra must handle sameness proof generation specially for SBOMs not created by the generator, but rather are provided by the producer wishing to compose multiple SBOMs. Specifically, because Petra computes sameness proofs upon initial SBOM ingestion, to enable sameness proof verification of a composed SBOM tree

Algorithm 2: Consumption & Verification

```
Data: RedactedSBOM, SKA, PKgen, PKprod, f  
Result: DecryptSBOM  
1 if !VerifySignature(RedactedSBOM.countersig, PKprod) then  
2   | return FAIL_UNTRUSTED_SBOM  
3 end  
4 if !VerifySignature(RedactedSBOM.sig, PKgen) then  
5   | return FAIL_UNTRUSTED_SBOM  
6 end  
7 /* Obtain decryption AES Keys */  
   AESSet ← DecryptKeys(SBOMNode, SKA)  
8 /* Traverse tree and decrypt accessible nodes */  
   DecryptSBOM ← new SbomTree()  
9 for node ∈ RedactedSBOM do  
10  | if satisfies(node.Policy) then  
11    | PlainNode ← Decrypt(node, AESSet)  
12  | else  
13    | PlainNode ← node  
14  | end  
15  | DecryptSBOM.insert(PlainNode)  
16 end  
17 /* Compute membership using decrypted tree */  
   proof ← computeMembershipProof(DecryptSBOM, f)  
18 if !verify(proof) then  
19   | return FAIL_GENERATOR_PRODUCER_LIED  
20 end  
21 /* Return decrypted SBOM tree */  
22 return DecryptSBOM
```

CompSBOM without needing the full plaintext tree for an inserted *RedactedSBOM_i*, the generator first recomputes the sameness proof of *CompSBOM* incorporating the value of the sameness proof at the root of *RedactedSBOM_i*.

Then, the updated sameness proof is embedded when computing the Merkle hash over the remainder of the redacted *CompSBOM* tree. This way, the sameness proof can be verified by the producer of the top-level *CompSBOM*, but under the assumption that the inserted *RedactedSBOM_i* is valid.

6 Implementation and Evaluation

6.1 Implementation

We implemented Petra’s SBOM parsing and tree construction in Python, while using Rust for CP-ABE functionality and other cryptographic routines. We used the Maturin framework [11] as well as PyO3 [15] to provide bindings between the two. Further, we use vendor-provided SBOM libraries for SPDX and CycloneDX translation and support. In total, the Petra implementation uses less than 1,600 Python LoC and less than 190 Rust LoC.

For artifact indexing, we use Package URLs (pURLs) [76], which defines a structured and ecosystem-agnostic format for naming software components (e.g., pkg:npm/express@4.17.1). This usage is unrelated to “persistent URLs”; our use of pURLs aligns with conventions in both major SBOM standards [7, 13].

We implemented a KMS for Petra by extending the Sigstore

project [65]. Our modifications allow for sigstore identity proofs (using OIDC) to also pass user attributes A for SK_A generation. The core set of attributes that are provided by the KMS are usernames, and domains (e.g., `foo@bar.com` is decomposed as `user:foo` and `namespace:bar.com`). To further provide specific attributes, the KMS can be configured to match particular claims within an OIDC token (e.g., to show control of a particular workload identity, or account-specific attributes). Lastly, the identity token provided to the KMS is used to derive temporal attributes at a time-window granularity (currently set at one month).

6.2 Performance Evaluation

First, we set off to characterize Petra’s practicality by studying the overhead of running Petra. We posit that the overhead of using Petra falls within three major categories: 1) The timing to convert regular SBOMs into SBOM trees, 2) The storage overhead of a redacted SBOM relative to a regular SBOM, and 3) The timing for encrypting and verifying properties within a Redacted SBOM.

Experimental Setup We carried out the evaluation on a 2.10 GHz Intel Xeon E5620 machine with 256 GB of RAM using a standard Ubuntu 22.04.2 LTS Docker container, a stable build of Rust (rustc 1.75.0), and Python 3.10.

Evaluation Policies In addition, we aim to understand the effect of specific policies as they apply to a redacted SBOM. We expect that more complex policies may result in a larger ciphertext, and thus increase the overhead.

Identifying real-world examples of SBOM access policies was a major challenge. To construct representative policies for our evaluation, we analyzed literature and identified two primary categories of sensitive data in SBOMs: proprietary information and weaknesses information [28, 68, 69, 83, 88, 95]. Then for each type of sensitive data, we map it to claim types in an SBOM (e.g., dependency claims might threaten intellectual property rights, and vulnerability claims might expose security weaknesses). We included security weaknesses, and intellectual property policies in Appendix B.

SBOM Dataset We use SBOM datasets from previous literature [87]. After normalization and de-duplication, these datasets contain 3,380 SBOMs from both generated SBOMs (e.g., by a third-party tool), as well as in-the-wild SBOM collections (e.g., identified by indexing platforms such as sourcegraph [19]). This dataset contains SBOMs that cover between 1 and 5002 packages, with an average of 242 per SBOM, a median of 122, and a standard deviation of 408 packages. Our evaluation against this highly variable range of SBOMs reflects the structural complexity and real-world diversity of software artifacts that Petra supports.

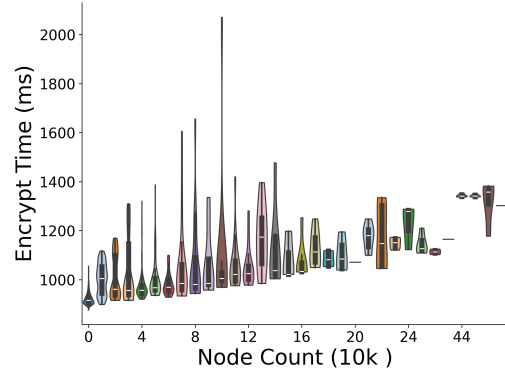


Figure 5: Encryption Overhead for Intellectual Property policy: Decryption overhead matches encryption overhead, with an average deviation of 120.05 ms.

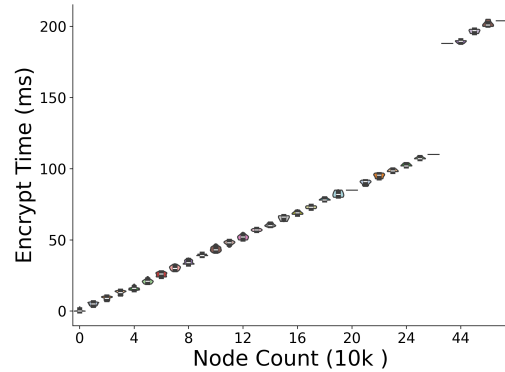


Figure 6: Encryption Overhead for the weaknesses policy: Decryption overhead matches encryption overhead, with an average deviation of 1.16 ms.

6.2.1 Storage Overhead

We measure the resulting size of Petra trees for the SBOMs in our dataset based on each policy, as well as a non-encrypted SBOM size. We find that redacted SBOM trees incur less than 13% increased storage cost on average compared to their unencrypted counterparts. This cost is dominated by the complexity of the policy. In the complicated policy case, the cost of encrypting a medium-sized SBOM (i.e., 12MB) is rather stark with more than 13% increased in the space of the tree, which may make Petra an undesirable system for highly-complicated access scenarios.

However, for our two illustrative policies, a medium-sized SBOM requires only 12% and 51% encryption and decryption storage overhead, respectively. This is acceptable for a system that provides confidential SBOM Exchange according to informal discussion with industry stakeholders.

6.2.2 Performance Overhead

We now study the Performance overhead of Petra. In this case, we measure the cost incurred by each processing pass (tree, encryption, Merkle-hashing) to provide a granular understanding of each foundational technique.

Figs. 5 and 6 illustrate the timing overhead of encrypting an SBOM under the intellectual property and weaknesses policies, respectively. We observe a visible discontinuity in the encryption timing results (Fig. 6) for a small subset of SBOMs. This gap corresponds to a few large SBOMs containing more than 440K nodes. We theorize that these SBOMs exceed CPU-cache capacity, triggering additional memory accesses and a visible discontinuity in the graph. This is an artifact of the "in the wild" dataset we are operating with, where large SBOMs (in terms of tree size) also carry substantial additional data and thus are also larger in storage.

Unsurprisingly, the decryption cost in our experiments closely mirrors the encryption time, with a mean deviation of 120 ms, indicating symmetry in processing times for encryption/decryption operations. Much like the storage overhead costs, the processing cost is dominated by the CP-ABE pass.

6.2.3 Comparison with Protobom

To evaluate Petra against existing tooling, we compare it to Protobom [14], which generates a format-agnostic graph given an input SBOM document. Using the SPDX 2.3 SBOMs in our dataset (due to Protobom's format support), we measured the number of nodes of their internal graph representations. We find that Protobom on average produces 93% fewer nodes than Petra. This reflects the design goals of each system: Protobom prioritizes efficient format translation, while Petra emphasizes preserving compositional relationships. While resulting in a more "verbose" tree, our design also facilitates fine-grained access control.

6.3 Practical Evaluation

We developed Petra through a research collaboration with an enterprise team managing large-scale SBOMs during an internship at Intel Labs, and its design reflects direct operational and practical security feedback. In enterprise settings, SBOMs are frequently redistributed across organizational and trust boundaries.

One of the Petra design features requested by the enterprise team is a KMS and verifier as foundational components for practicality: they provide a user-friendly and centrally managed way to issue decryption keys, enforce organizational authorization policies, which relieve producers and consumers from handling cryptographic keys directly.

However, these services alone are not sufficient. A KMS can issue keys and a verifier can check signatures, but neither enables fine-grained access control, nor do they allow SBOMs

to remain verifiable while being redistributed. While an implementation could, in principle, combine decryption with the KMS if fully trusted, our design keeps decryption and verification distributed so that SBOMs can be safely re-distributed and independently verified by arbitrary parties.

While Petra has not yet been deployed in a production environment, CP-ABE based systems for electronic health records [38] show that such deployments are practical.

7 Security Analysis

Petra's ability to achieve our desired security and confidentiality goals is central. Here, we study how Petra's design goals in § 3.3 address threats (T1–T7) in §3.2.

At a high-level, we consider two areas of attack in the confidential SBOM exchange threat model. First, generators or distributors may seek to compromise the integrity of an SBOM to mislead consumers (§7.1). Second, consumers may attempt to violate confidentiality by gleaning information from an SBOM without having explicit access (§7.2).

7.1 SBOM Integrity Attacks

Tampering with SBOM data (T1): A generator, or a producer that acts as or colludes with the generator, may misbehave by adversarially tampering with SBOM data. For instance, the generator might make critical elements inaccessible to consumers by removing a node from the SBOM tree or altering node data to bypass security or compliance audits. Through this, consumers may make a wrong determination about the security posture of a software artifact.

Petra's sameness property (I1) addresses this threat by making structural changes between plaintext and redacted SBOMs detectable by consumers using Merkle trees and membership proofs. That said, consumers cannot detect tampering if a leaf (and any sibling) nodes are removed for SBOM fields that a consumer cannot decrypt, because this scenario looks indistinguishable from an SBOM with a stricter access policy.

Split view attacks (T2): A distributor may collude with a generator to show different consumers two different SBOMs for the same artifact. In this attack, the attacker tries to fool a victim consumer V into using artifact A by sending her a compromised SBOM S_V (e.g., omitting a vulnerability). At the same time, they send another consumer C the faithful SBOM S_A , allowing her to detect the vulnerability.

Petra provides non-equivocation (I3) by giving V and C a means to observe each other's SBOMs during SBOM redistribution of the faithful S_A via another uncompromised SBOM, revealing this attack. That is, because the redistributed S_A and S_V would have different Merkle root hashes, even though they represent the same artifact A , the victim consumer V would detect the attack. In addition, the signature/counter signature scheme, provides non-repudiation for generators and producers, which allows consumers to verify they are

both looking at an SBOM provided by the same party. Lastly, signed Merkle tree roots can optionally be published to a public transparency log to support global visibility and facilitate split view detection.

Corrupting SBOMs redistributed to downstream consumers (T3): A distributor, or compromised consumer during redistribution, may tamper with an SBOM while in transit. This causes a denial-of-service (DoS), rendering the SBOM unusable to downstream consumers. However, the redistribution congruence property (I2) of SBOM trees addresses this threat because this tampering would be detectable by any consumer with sufficient access to decrypt the SBOM and validate its contents or structure.

7.2 SBOM Confidentiality Attacks

Unauthorized SBOM data leaks (T4): A consumer with access to a plaintext SBOM may misuse or maliciously share sensitive information contained therein to identify points of access to a system or vulnerability information for competitive analysis or more a harmful software producer compromise [35, 50, 88, 95]. Thus, Petra provides semantic confidentiality (C1) by enabling software producers to redact SBOMs using selective encryption, limiting the risk of unauthorized SBOM data leaks in transit and at rest.

Dictionary attacks (T5): An attacker could attempt to carry out a dictionary attack on specific nodes. This can happen by using the sameness proof hashes, and the fact that it is likely some fields have a very limited “key space”. License elements, for example, have a limited number of field values — 677 on the latest SPDX list [20] — which could be enumerated by an attacker.

However, Petra preserves semantic confidentiality (C1) through sameness proofs that use a random salt that is hashed and encrypted along with the node data, requiring an attacker to not only enumerate possible values, but also the nonce key space (256 bits of key space in the current Petra implementation), which makes such attacks computationally infeasible.

Collusion to decrypt inaccessible information (T6): A group of consumers with non-overlapping attributes may try to collude to access more information than what they can individually access. However, Petra’s use of CP-ABE for redaction offers collusion resistance (C2), which addresses this threat because the scheme uses single decryption keys for data accessible to holders of specific $N : N$ combinations of attributes (i.e., permissions) to resist collusion between parties that have access to only a subset of the needed permissions.

Inferring inaccessible SBOM data (T7): A consumer may attempt to infer redacted SBOM data using partial access to the same SBOM or through other related SBOMs. Petra addresses two classes of information inference attacks.

Intra-SBOM Inference Attacks: An attacker with access to a portion of an SBOM using their attributes tries to infer other aspects of the redacted SBOM.

Of these, we identify two variants:

- **Structural Inference:** An attacker seeks to infer the *field name* in encrypted nodes.
- **Semantic Inference:** An attacker seeks to infer the *field name and value* in encrypted nodes.

An attacker may, for example, study the padding properties of field names and values. This risk can be reduced in practice by using a reasonable padding value at a slight cost on storage.

Further, an attacker can attempt to infer information by looking at the structure of the redacted SBOM tree. Certain aspects of an SBOM will be biased to some tree structures; for example, a dependency list is often represented by a complex node with a large number of children. As such, it is possible an attacker learns certain parts of the tree are likely particular complex fields.

Petra preserves semantic confidentiality (C1) in this attack scenario because SBOM tree leaf nodes, while carrying the most crucial information, are unlikely to leak any information from their structure because individual field nodes are encrypted in a uniform way using a non-deterministic encryption scheme.

Inter-SBOM Inference Attacks: An attacker with access to an SBOM may use it to try to infer information about another SBOM they do not have access to. In this scenario, the attacker has the key to decrypt certain SBOM fields in $SBOM_A$ and tries to use this information to infer the value of those fields in an $SBOM_B$. However, Petra preserves semantic confidentiality (C1) by reducing the probability of an attacker inferring a field *value*, even if they were equal in the two SBOMs. That is, by employing encryption schemes that are non-deterministic and dependent on the input access policy, an attacker is highly unlikely to succeed in correlating nodes between two SBOMs.

8 Discussion

The Petra construction outlined above helps further both SBOM confidentiality and transparency. We discuss the following limitations and future extensions.

SBOM quality & accuracy. Petra’s cryptographic guarantees and access control do not depend on the quality and completeness of underlying SBOM data. Nevertheless, improvements in SBOM quality and accuracy could enable Petra to enforce richer SBOM redaction policies.

Expanding to other types of metadata. Currently, Petra focuses mainly on SBOMs, and a promising future direction is adding support for other supply chain metadata formats such as VEX, OpenVEX, or SLSA. By expanding Petra to handle these other formats, the system could offer a more comprehensive solution to secure and audit software supply chains. However, integrating these formats would introduce challenges related to the diversity in their structure, semantics,

and encryption needs. Future work could explore efficient ways to harmonize these differences, ensuring Petra remains both flexible, semantically consistent, and scalable.

Software update history. A desirable property of SBOM exchange is capturing the over-time evolution of software in SBOMs [88]. Though at first sight the composition system within Petra may allow to *append* new information in redacted SBOMs (e.g., new vulnerability information), this is not sufficient. Instead, a mechanism that allows for appending new information, but also removing outdated information, is necessary. For this, other backing data structures such as a *Merkle*² [47] tree may allow for both additions, erasures and replacements on an SBOM tree.

Policy updates post-publication. Petra’s design enables partial updates and avoids full SBOM regeneration in response to policy update. Since each node in the SBOM tree is individually encrypted and tied to its specific access policy, only nodes whose access policies have changed need to be re-encrypted. Corresponding Merkle proofs are updated for those nodes and any intermediate nodes affected by the change. Importantly, redistribution does not need to be repeated if policy changes are made to individual fields, updating a policy and re-encrypting a single field does not invalidate redistributed copies. This design allows producers to modify or refine access control over time (e.g., granting broader visibility to a previously restricted field) without invalidating already-redistributed SBOMs. However, redistribution becomes necessary when structural changes occur.

Semantic-aware policies. Petra currently supports SBOM redaction policies at the granularity of fields and complex SBOM elements, and treats the SBOM as a non-semantic tree (i.e., a collection of cryptographically protected claims rather than a semantically interpreted document). As a result, a redactor with a weak policy could encrypt data in a way that undermines SBOM utility without violating any cryptographic guarantees. For example, if a redactor encrypts all related fields with a different policy, no consumer can make any meaningful verification.

This limitation reflects a deliberate separation between cryptographic verification and semantic interpretation. Petra supports this separation by allowing producers and designated verifiers to access full-SBOM plain-text to validate complete sameness, while consumers receive structurally verifiable views limited to the information they are authorized to access. A malicious producer cannot undetectably alter or remove information within the verifiable portion of an SBOM, as any modification would invalidate the corresponding Merkle proofs or expose a redaction marker, assuming Petra uses a collision-resistant hash function. However, determining whether redaction conceals semantically relevant information is inherently a policy-level concern. Thus, a semantic-aware SBOM translation algorithm could further limit the possibility of undermining SBOM utility and should be trivial to implement on top of the Petra SBOM code.

Consumer policies. A potential future extension of the current construction is allowing consumers to define security policies for consumed SBOMs. A separate role such as a policy issuer could obtain policy requests from consumers and enforce their usage at redaction. For example, by using Homomorphic-Policy Attribute-Based Encapsulation [32], a consumer may further append attributes to the SBOM tree node policy before redistribution to limit access to sub-elements on SBOMs under their discretion. However, this feature may further introduce complexity and risk practical deployment.

Further confidentiality properties. There are a number of additional confidentiality properties that may be of interest in future extensions of Petra, such as node compression or encrypting node references to hide the structure of an SBOM. However, we argue that the current design allows for easier membership proof, and a simpler composition flow.

9 Conclusion

In this paper, we presented the first formally-defined abstract SBOM exchange model. From it, we extended this model to consider confidentiality expectations within the involved parties. Ultimately, these two notions present a way to develop new systems and protocols to address confidentiality and integrity expectations of software users.

Beyond the modelling effort, we presented Petra, the first system to provide confidential SBOM Exchange capabilities, while maintaining tool compatibility expectations. While Petra is not the only design that can achieve these goals, it satisfies the majority of design expectations for a system that provides SBOM Exchange. Due to its design, Petra can be integrated with other SBOM distribution efforts within industry and open source. We are currently engaging with open source projects to integrate Petra structures and processes to achieve stronger confidentiality properties.

Acknowledgment

We thank anonymous reviewers for their constructive feedback. We also thank Intel Corporation for supporting this research through a single-PI grant.

A Ethical Considerations

Petra is primarily concerned with confidentiality in the exchange of SBOM data. We do not expect any ethical harms from providing software producers with better control over their data. Instead, Petra supports a goal that aligns with widely accepted values of security and transparency, namely enabling organizations to protect confidential details (e.g., intellectual property and vulnerability information) while still providing accountability through auditable SBOM exchange.

To help ensure that our design meets real-world needs and expectations, we engaged with both open-source and enterprise stakeholders during system development. Their feedback helped shape the design of our redaction and verification mechanisms, and practical adoption concerns. We believe Petra provides a balanced solution that protects confidentiality, enables trustworthiness, and avoids ethical pitfalls that could arise from restricting information access arbitrarily.

Affected stakeholders and implications: Petra impacts several stakeholder groups.

Software producers (open source and enterprise vendors) benefit from the ability to share SBOMs without disclosing sensitive information, addressing a core barrier to practical SBOM adoption. Petra’s cryptographic sameness proofs limit a producer’s ability to misrepresent SBOM contents without detection by authorized verifiers. However, Petra does not prevent producers from choosing overly restrictive redaction policies; determining whether a policy is appropriate remains a governance and compliance responsibility.

Software consumers including those that subsequently act as producers by redistributing composed artifacts, gain stronger assurances that redacted SBOMs faithfully correspond to the plaintext SBOM, improving confidence in analyses like SBOM-based risk assessments. At the same time, consumers may receive only partial visibility depending on access policies, which may constrain certain analyses. Petra makes these limitations explicit and verifiable, rather than implicit.

Verifiers (auditors, regulators, and compliance authorities) are privileged stakeholders that Petra enables to verify SBOM integrity without requiring universal disclosure. This supports regulatory oversight while minimizing unnecessary exposure of sensitive data. Petra does not replace regulatory judgment or audit policy; it provides cryptographic mechanisms that they may choose to rely on.

Other SBOM exchange roles (generators, distributors, KMS) are affected differently by Petra’s design. Distributors benefit from reduced exposure to confidential SBOM data, as they are expected to store only redacted SBOMs and don’t get decryption keys for redacted fields. Generators, by contrast, operate on software artifacts and plaintext SBOMs in order to perform SBOM generation, redaction, and adding integrity proofs, and therefore remain trusted with access to sensitive information. Key management services (KMS) issue and manage cryptographic keys and enforce access control via attributes, but do not require access to SBOM contents.

Petra reduces the need for distributing plaintext SBOMs by restricting plain-text SBOM access to authorized roles, but it assumes correct cryptographic implementation and key management practices; operational failures in these components remain outside the scope of Petra’s ethical guarantees.

B Open Science

Petra artifacts are open source including source code and evaluation, under Apache License 2.0 and are available at <https://doi.org/10.5281/zenodo.17906622>. The bom-shelter dataset we used is available at <https://doi.org/10.5281/zenodo.17859760>.

References

- [1] Archivista. <https://github.com/in-toto/archivista>.
- [2] Binary Transparency. <https://binary.transparency.dev/>.
- [3] bomctl/bomctl. <https://github.com/bomctl/bomctl>. original-date: 2024-01-10T18:42:31Z.
- [4] bomctl/ROADMAP.md at main · bomctl/bomctl. <https://github.com/bomctl/bomctl/blob/main/ROADMAP.md>.
- [5] Build software better, together. <https://github.com>.
- [6] CISA issues request for comment on software identification ecosystem analysis white paper | CISA. <https://www.cisa.gov/news-events/news/cisa-issues-request-comment-software-identification-ecosystem-analysis-white-paper>.
- [7] CycloneDX. <https://cyclonedx.org/>.
- [8] DBOM: Standardizing Attestation Sharing. <https://dbom.io/>.
- [9] FedRAMP marketplace. <https://marketplace.fedramp.gov/products>.
- [10] guac. <https://guac.sh/>.
- [11] Maturin. <https://github.com/PyO3/maturin>. GitHub repository.
- [12] npm | home. <https://www.npmjs.com/>.
- [13] OWLDoc. <https://spdx.org/rdf/spdx-terms-v2.1/>.
- [14] protobom/protobom. <https://github.com/protobom/protobom>. original-date: 2023-06-04T20:11:52Z.
- [15] Pyo3. <https://github.com/PyO3/pyo3>. GitHub repository.
- [16] sigstore. <https://www.sigstore.dev>.

- [17] Software Identification (SWID) Tagging. <https://csrc.nist.gov/projects/software-identification-swid>.
- [18] Software Package Data Exchange (SPDX). <https://spdx.dev/>.
- [19] Sourcegraph | code intelligence platform. <https://sourcegraph.com/>.
- [20] SPDX license list | software package data exchange (SPDX). <https://spdx.org/licenses/>.
- [21] Trivy home. <https://trivy.dev/>.
- [22] Widevine. <https://widevine.com/>. Google.
- [23] Mitigating log4shell and other log4j-related vulnerabilities | CISA. <https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-356a>, Dec 2021.
- [24] Mustafa Al-Bassam and Sarah Meiklejohn. Contour: A practical system for binary transparency. In *International Workshop on Data Privacy Management*, pages 94–110. Springer, 2018.
- [25] Amazon Web Services. Amazon Inspector SBOM Generator. <https://docs.aws.amazon.com/inspector/latest/user/sbom-generator.html>.
- [26] Fattaneh Bayatbabolghani and Marina Blanton. Secure multi-party computation. In *Proceedings of the 2018 ACM SIGSAC conference on computer and communications security*, pages 2157–2159, 2018.
- [27] John Bethencourt, Amit Sahai, and Brent Waters. Ciphertext-policy attribute-based encryption. In *2007 IEEE Symposium on Security and Privacy (SP '07)*, pages 321–334, 2007.
- [28] Tingting Bi, Boming Xia, Zhenchang Xing, Qinghua Lu, and Liming Zhu. On the way to sboms: Investigating design issues and solutions in practice. *ACM Transactions on Software Engineering and Methodology*, 33(6):1–25, 2024.
- [29] Blue Goat Cyber. SBOM (Software Bill of Materials) Services. <https://bluegoatcyber.com/services/sbom-software-bill-of-materials-services/>.
- [30] Justin Cappos. Avoiding theoretical optimality to efficiently and privately retrieve security updates. In *Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers 17*, pages 386–394. Springer, 2013.
- [31] Anesu Chaora, Nathan Ensmenger, and L Jean Camp. Discourse, challenges, and prospects around the adoption and dissemination of software bills of materials (sboms). In *2023 IEEE International Symposium on Technology and Society (ISTAS)*, pages 1–4. IEEE, 2023.
- [32] Jérémy Chotard, Duong Hieu Phan, and David Pointcheval. Homomorphic-policy attribute-based key encapsulation mechanisms. In *International Conference on Information Security*, pages 155–172. Springer, 2017.
- [33] Lucian Constantin. Npm attackers sneak a backdoor into node. js deployments through dependencies. <https://thenewstack.io/npm-attackers-sneak-a-backdoor-into-node-js-deployments-through-dependencies/>, 2018.
- [34] Cybersecurity and Infrastructure Security Agency (CISA). Types of software bill of materials (sboms). <https://www.cisa.gov/sites/default/files/2023-04/sbom-types-document-508c.pdf>, April 2023.
- [35] Gregorio Dalia, Corrado Aaron Visaggio, Andrea Di Sorbo, and Gerardo Canfora. Sbom ouverture: What we need and what we have. In *Proceedings of the 19th International Conference on Availability, Reliability and Security, ARES '24, New York, NY, USA, 2024*. Association for Computing Machinery.
- [36] DataTrails. RKVST SBOM Hub. <https://www.datatrails.ai/rkvst-sbom-hub/>.
- [37] Endor Labs. Manage 3rd party risk with Endor SBOM Hub. <https://www.endorlabs.com/product/endor-labs-sbom-hub>.
- [38] ETSI. ETSI TS 103 532 V1.1.1 (2018-03): CYBER; Attribute Based Encryption for Attribute Based Access Control. Technical specification, European Telecommunications Standards Institute, March 2018. DTS/CYBER-0025.
- [39] European Commission. Cyber Resilience Act. <https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act>, Sep 2022.
- [40] Gabriel Ferreira, Limin Jia, Joshua Sunshine, and Christian Kästner. Containing malicious package updates in npm with a lightweight permission system. In *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*, pages 1334–1346. IEEE, 2021.
- [41] Enduring Security Framework. Securing the Software Supply Chain: Recommended Practices for Managing Open-Source Software and Software Bill of Materials. <https://www.cisa.gov/sites/default/files/>

2024-08/ESF_SECUREING_THE_SOFTWARE_SUPPLY_CHAIN%20RECOMMENDED%20PRACTICES%20FOR%20MANAGING%20OPEN%20SOURCE%20SOFTWARE%20AND%20SOFTWARE%20BILL%20OF%20MATERIALS_508.pdf, Dec 2023.

- [42] GitHub. actions/attest-sbom. <https://github.com/actions/attest-sbom>.
- [43] M Choudary Gorantla, Colin Boyd, and Juan Manuel González Nieto. Attribute-based authenticated key exchange. In *Australasian Conference on Information Security and Privacy*, pages 300–317. Springer, 2010.
- [44] Vipul Goyal, Omkant Pandey, Amit Sahai, and Brent Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *Proceedings of the 13th ACM conference on Computer and communications security*, pages 89–98, 2006.
- [45] Qi Han, Yinghui Zhang, and Hui Li. Efficient and robust attribute-based encryption supporting access policy hiding in internet of things. *Future Generation Computer Systems*, 83:269–277, 2018.
- [46] Stephen Hendrick. The State of Software Bill of Materials (SBOM) and Cybersecurity Readiness. <https://www.linuxfoundation.org/research/the-state-of-software-bill-of-materials-sbom-and-cybersecurity-readiness>, Jan 2022.
- [47] Yuncong Hu, Kian Hooshmand, Harika Kalidhindi, Seung Jin Yang, and Raluca Ada Popa. Merkle 2: A low-latency transparency log system. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 285–303. IEEE, 2021.
- [48] Junbeom Hur and Dong Kun Noh. Attribute-based access control with efficient revocation in data outsourcing systems. *IEEE Transactions on Parallel and Distributed Systems*, 22(7):1214–1221, 2011.
- [49] Patrick Jauernig, Ahmad-Reza Sadeghi, and Emmanuel Stempf. Trusted execution environments: properties, applications, and challenges. *IEEE Security & Privacy*, 18(2):56–60, 2020.
- [50] Berend Kloeg, Aaron Yi Ding, Sjoerd Pellegrum, and Yury Zhauniarovich. Charting the path to sbom adoption: A business stakeholder-centric approach. In *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security*, pages 1770–1783, 2024.
- [51] Trishank Karthik Kuppusamy, Santiago Torres-Arias, Vladimir Diaz, and Justin Cappos. Diplomat: Using delegations to protect community repositories. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 567–581, Santa Clara, CA, March 2016. USENIX Association.
- [52] Ben Laurie. Certificate transparency. *Communications of the ACM*, 57(10):40–46, 2014.
- [53] Eliot Lear, Ralph Droms, and Dan Romascanu. Manufacturer usage description specification. <https://datatracker.ietf.org/doc/rfc8520>. Num Pages: 60.
- [54] Ming Li, Shucheng Yu, Yao Zheng, Kui Ren, and Wenjing Lou. Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *IEEE Transactions on Parallel and Distributed Systems*, 24(1):131–143, 2013.
- [55] RUIXUAN LI, Chenglin Shen, Heng He, Xiwu Gu, Zhiyong Xu, and Cheng-Zhong Xu. A lightweight secure data sharing scheme for mobile cloud computing. *IEEE Transactions on Cloud Computing*, 6(2):344–357, 2018.
- [56] Lineaje Inc. Lineaje SBOM360 Hub. <https://www.lineaje.com/sbom360-hub>.
- [57] Harjasleen Malvai, Lefteris Kokoris-Kogias, Alberto Sonnino, Esha Ghosh, Ercan Oztürk, Kevin Lewi, and Sean Lawlor. Parakeet: Practical Key Transparency for End-to-End Encrypted Messaging. In *Proceedings 2023 Network and Distributed System Security Symposium*, San Diego, CA, USA, 2023. Internet Society.
- [58] Marcela S Melara, Aaron Blankstein, Joseph Boneau, Edward W Felten, and Michael J Freedman. {CONIKS}: Bringing key transparency to end users. In *24th USENIX Security Symposium (USENIX Security 15)*, pages 383–398, 2015.
- [59] Marcela S. Melara and Santiago Torres-Arias. A Viewpoint on Software Supply Chain Security: Are We Getting Lost in Translation? . *IEEE Security & Privacy*, 21(06):55–58, November 2023.
- [60] Ralph C Merkle. A digital signature based on a conventional encryption function. In *Conference on the theory and application of cryptographic techniques*, pages 369–378. Springer, 1987.
- [61] Kelsey Merrill, Zachary Newman, Santiago Torres-Arias, and Karen R Sollins. Speranza: Usable, privacy-friendly software signing. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, pages 3388–3402, 2023.
- [62] National Telecommunications and Information Administration. Minimum elements for a software bill of materials (sbom), 2021.
- [63] National Telecommunications and Information Administration (NTIA). The minimum elements

- for a software bill of materials (sbom). https://www.ntia.doc.gov/files/ntia/publications/sbom_minimum_elements_report.pdf, July 2021.
- [64] Muhammad Naveed, Seny Kamara, and Charles V Wright. Inference attacks on property-preserving encrypted databases. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, pages 644–655, 2015.
- [65] Zachary Newman, John Speed Meyers, and Santiago Torres-Arias. Sigstore: Software Signing for Everybody. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, pages 2353–2367, Los Angeles CA USA, November 2022. ACM.
- [66] Kirill Nikitin, Eleftherios Kokoris-Kogias, Philipp Jovanovic, Nicolas Gailly, Linus Gasser, Ismail Khoffi, Justin Cappos, and Bryan Ford. {CHAINIAC}: Proactive {Software-Update} transparency via collectively signed skipchains and verified builds. In *26th USENIX Security Symposium (USENIX Security 17)*, pages 1271–1287, 2017.
- [67] Sabato Nocera, Simone Romano, Massimiliano Di Penta, Rita Francese, and Giuseppe Scanniello. Software bill of materials adoption: A mining study from github. In *2023 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pages 39–49. IEEE, 2023.
- [68] NTIA Formats and Tooling Working Group. Software Consumers Playbook: SBOM Acquisition, Management, and Use. https://www.ntia.gov/files/ntia/publications/software_consumers_sbom_acquisition_management_and_use_-_final.pdf, Nov 2021.
- [69] NTIA Formats and Tooling Working Group. Software Suppliers Playbook: SBOM Production and Provision. https://www.ntia.gov/sites/default/files/publications/software_suppliers_sbom_production_and_provision_-_final_0.pdf, Nov 2021.
- [70] NTIA Multistakeholder Process on Software Component Transparency Framing Working Group. Sharing and Exchanging SBOMs. https://www.ntia.gov/files/ntia/publications/ntia_sbom_sharing_exchanging_sboms-10feb2021.pdf, Feb 2021.
- [71] Marc Ohm, Henrik Plate, Arnold Sykosch, and Michael Meier. Backstabber’s knife collection: A review of open source software supply chain attacks. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 17th International Conference, DIMVA 2020, Lisbon, Portugal, June 24–26, 2020, Proceedings 17*, pages 23–43. Springer, 2020.
- [72] Chinenye Okafor, Taylor R. Schorlemmer, Santiago Torres-Arias, and James C. Davis. SoK: Analysis of Software Supply Chain Security by Establishing Secure Design Properties. In *Proceedings of the 2022 ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses, SCORED’22*, pages 15–24, New York, NY, USA, November 2022. Association for Computing Machinery.
- [73] OWASP CycloneDX. Owasp cyclonedx launches sbom exchange api, standardizing sbom distribution. <https://cyclonedx.org/news/owasp-cyclonedx-launches-sbom-exchange-api--standardizing-sbom-distribution/>, October 2024.
- [74] Can Ozkan, Xinhai Zou, and Dave Singelee. Supply chain insecurity: The lack of integrity protection in sbom solutions, 2025.
- [75] Eric O’Donoghue, Brittany Boles, Clemente Izurieta, and Ann Marie Reinhold. Impacts of software bill of materials (sbom) generation on vulnerability detection. In *2024 ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses*, 2024.
- [76] Package URL. Package url (purl) specification. <https://github.com/package-url/purl-spec>, 2024.
- [77] Benny Pinkas and Tomas Sander. Securing passwords against dictionary attacks. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pages 161–170, 2002.
- [78] Sudhakar Ramakrishna. An investigative update of the cyberattack. <https://orangematter.solarwinds.com/2021/05/07/an-investigative-update-of-the-cyberattack/>, May 2021.
- [79] Adriana Sejfia and Max Schäfer. Practical automated detection of malicious npm packages. In *Proceedings of the 44th International Conference on Software Engineering*, pages 1681–1692, 2022.
- [80] Sigstore Community. Fulcio: Sigstore OIDC-pki for code signing. <https://github.com/sigstore/fulcio>, 2025.
- [81] Prabath Siriwardena. Openid connect (oidc). In *Advanced API Security: OAuth 2.0 and Beyond*, pages 129–155. Springer, 2019.
- [82] SOOS. Research Packages. <https://app.soos.io/research/packages>.
- [83] Trevor Stalaker, Nathan Wintersgill, Oscar Chaparro, Massimiliano Di Penta, Daniel M German, and Denys Poshyvanyk. BOMs away! inside the minds of stakeholders: A comprehensive study of bills of materials for software systems. In *Proc. ICSE ’24*, Feb 2024.

- [84] The White House. Executive Order on Improving the Nation’s Cybersecurity. <https://www.nist.gov/itl/executive-order-14028-improving-nations-cybersecurity>, May 2021.
- [85] Eric Tooley and Courtney Claessens. Introducing self-service SBOMs. <https://github.blog/enterprise-software/governance-and-compliance/introducing-self-service-sboms/>, Mar 2023.
- [86] Santiago Torres-Arias, Hammad Afzali, Trishank Karthik Kuppusamy, Reza Curtmola, and Justin Cappos. in-toto: Providing farm-to-table guarantees for bits and bytes. *Proc. of the 28th USENIX Security Symposium*, August 2019.
- [87] Santiago Torres-Arias, Dan Geer, and John Speed Meyers. A Viewpoint on Knowing Software: Bill of Materials Quality When You See It. *IEEE Security & Privacy*, 21(06):50–54, November 2023.
- [88] Boming Xia, Tingting Bi, Zhenchang Xing, Qinghua Lu, and Liming Zhu. An empirical study on software bill of materials: Where we stand and the road ahead. In *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, pages 2630–2642, 2023.
- [89] Boming Xia, Dawen Zhang, Yue Liu, Qinghua Lu, Zhenchang Xing, and Liming Zhu. Trust in software supply chains: Blockchain-enabled sbom and the aibom future. In *Proceedings of the 2024 ACM/IEEE 4th International Workshop on Engineering and Cybersecurity of Critical Systems (EnCyCris) and 2024 IEEE/ACM Second International Workshop on Software Vulnerability*, pages 12–19, 2024.
- [90] Shengmin Xu, Yingjiu Li, Robert H. Deng, Yinghui Zhang, Xiangyang Luo, and Ximeng Liu. Lightweight and expressive fine-grained access control for health-care internet-of-things. *IEEE Transactions on Cloud Computing*, 10(1):474–490, 2022.
- [91] Wenhan Xu, Hui Ma, Zishuai Song, Jianhao Li, and Rui Zhang. Gringotts: An encrypted version control system with less trust on servers. *IEEE Transactions on Dependable and Secure Computing*, 21(2):668–684, 2023.
- [92] Sergey Yekhanin. Private information retrieval. *Communications of the ACM*, 53(4):68–73, 2010.
- [93] Aditya Sirish A Yelgundhalli, Patrick Zielinski, Reza Curtmola, and Justin Cappos. Rethinking trust in forge-based git security. *Network and Distributed System Security (NDSS) Symposium*, 2025.
- [94] Sheng Yu, Wei Song, Xunchao Hu, and Heng Yin. On the correctness of metadata-based sbom generation: A differential analysis approach. In *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 29–36, 2024.
- [95] Nusrat Zahan, Elizabeth Lin, Mahzabin Tamanna, William Enck, and Laurie Williams. Software bills of materials are required. are we there yet? *IEEE Security & Privacy*, 21(2):82–88, 2023.
- [96] Markus Zimmermann, Cristian-Alexandru Staicu, Cam Tenny, and Michael Pradel. Small world with high risks: A study of security threats in the npm ecosystem. In *28th USENIX Security symposium (USENIX security 19)*, pages 995–1010, 2019.

Appendix

A SBOM Example

Figure 7 represents SBOM in SPDX format for the hello package that contains the source file `hello.c`, a `Makefile` for building it, and the compiled binary `hello`. All files are licensed under `GPL-3.0-or-later`. The SPDX document provides checksums for verification and describes relationships, showing that the binary is generated from the source code using the `Makefile`.

B SBOM policies

Figure 8 represents Petra policy used to control access to information that can reveal weaknesses, and Figure 9 represents Petra policy used to control access to proprietary information of the software artifact that an SBOM describes .

```

SPDXVersion: SPDX-2.2
DataLicense: CC0-1.0
SPDXID: SPDXRef-DOCUMENT
DocumentName: hello
DocumentNamespace: https://swinslow.net/spdx-examples/example1/hello-v3
Creator: Person: Steve Winslow (steve@swinslow.net)
Creator: Tool: github.com/spdx/tools-golang/builder
Creator: Tool: github.com/spdx/tools-golang/idsearcher
Created: 2021-08-26T01:46:00Z

```

```
##### Package: hello
```

```

PackageName: hello
SPDXID: SPDXRef-Package-hello
PackageDownloadLocation: git+https://github.com/swinslow/spdx-examples.git#example1/content
FilesAnalyzed: true
PackageVerificationCode: 9d20237bb72087e87069f96afb41c6ca2fa2a342
PackageLicenseConcluded: GPL-3.0-or-later
PackageLicenseInfoFromFiles: GPL-3.0-or-later
PackageLicenseDeclared: GPL-3.0-or-later
PackageCopyrightText: NOASSERTION

```

```
Relationship: SPDXRef-DOCUMENT DESCRIBES SPDXRef-Package-hello
```

Figure 7: An example SPDX document describing the package `hello`. It includes metadata such as licensing information, package verification codes, file checksums, and relationships between files.

[package]

```

version = (org:government_agency and cert:confidential_clearance)
|         or (tool:vuln_scanner and cert:licensed and cert:fedramp_authorized)
|         or (role:auditor and cert:audit_authorization=approved)
|         or (role:security_analyst and cert:confidential_clearance)

```

[vulnerability]

```

* = (org:government_agency and cert:confidential_clearance)
|   or (tool:vuln_scanner and cert:licensed and cert:fedramp_authorized)
|   or (role:auditor and cert:audit_authorization=approved)
|   or (role:security_analyst and cert:confidential_clearance)

```

Figure 8: Policy to restrict the access to information that might reveal weaknesses of the software artifact that the SBOM describes.

[Package]

```
name = "role:compliance_auditor"
```

```
"*" = "owner:ip"
```

```

|         or (role:security_auditor and cert:audit_authorization=approved)
|         or (role:security_analyst and org:federal_agency)

```

Figure 9: Policy to restrict the access to proprietary information of the software artifact that the SBOM describes.