NIST Advanced Manufacturing Series 300-6

Securing the Digital Threat for Smart Manufacturing: A Reference Model for Blockchain-Based Product Data Traceability

Sylvere Krima
Thomas Hedberg
Allison Barnard Feeney

This publication is available free of charge from:
https://doi.org/10.6028/NIST.AMS.300-6

National Institute of Standards and Technology
U.S. Department of Commerce

NIST Advanced Manufacturing Series 300-6

Securing the Digital Threat for Smart Manufacturing: A Reference Model for Blockchain-Based Product Data Traceability

Sylvere Krima
Engisis, LLC
Bethesda, MD

Thomas Hedberg
Allison Barnard Feeney
Systems Integration Division
Engineering Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.AMS.300-6

February 2019

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology

Abstract

Manufacturing is moving to a digital paradigm where data is produced and consumed faster than ever. Managing this quantity of data is critical to daily operations and requires access to trustworthy data. Data management is an rapidly-evolving field that is adapting to support new requirements and the exponentially increasing quantity of data. The technologies and strategies to manage data are diverse – from a centralized database and single source of truth, to distributed systems and cloud storage/computing – and each with their own strengths and flaws. While a single-source-of-truth can easily be corrupted or tampered with, distributed systems often face synchronization issues. Organizations that deal with large amounts of data must identify these strengths and flaws accurately to find the best solutions. In this paper, we discuss blockchain, the foundation of the bitcoin electronic currency, as a data management technology. Blockchain is a distributed storage framework that is virtually tamper resistant, has a native synchronization-discrepancy-resistance mechanism and is already highly praised in the financial world. We explore opportunities in the manufacturing area where digital product data is becoming a critical asset and present our initial assessment and early recommendations on how to record product data transactions on a blockchain.

Key words

data traceability; blockchain; information model.

Table of Contents

List of Tables

No table of figures entries found.

List of Figures

1. Introduction

Manufacturing is moving to a digital paradigm where data is produced and consumed faster than ever. Managing this quantity of data is critical to daily operations and requires access to trustworthy data. Data management is an ever-evolving field that is adapting rapidly to support new requirements and the rapidly increasing quantity of data. The technologies and strategies to manage data are diverse – from a centralized database and single source of truth, to distributed systems and cloud storage/computing – and each with their own strengths and flaws. While a single-source-of-truth can easily be corrupted or tampered with, distributed systems often face synchronization issues. Organizations that deal with large amounts of data must identify these strengths and flaws accurately to find the best solutions. In this paper, we discuss blockchain, the foundation of the bitcoin electronic currency [1], as a data management technology. Blockchain is a distributed storage framework that is tamper resistant and has a native synchronization-discrepancy-resistance mechanism. We explore opportunities in the manufacturing area where digital product data is becoming a critical asset [2] and present our initial assessment and early recommendations on how to record product data transactions on a blockchain.

1.1. Smart manufacturing

Smart Manufacturing (SM) is integration of operating technologies (OT) and information technologies (IT) working together in a real-time. SM requires digital product data be shared and exchanged among numerous engineering applications and information systems [3]. Through its entire lifecycle, a product generates an enormous amount of data in response to different processes (e.g., design, manufacturing, distribution) and needs (e.g., technical, commercial, regulatory). This data is often critical to any organization that plays a role in the product lifecycle. This is where the organizational contribution and value reside. Corrupt or tampered with data can have catastrophic consequences on product development and impact an organization's growth. There is a need to protect the product data and its owner(s) by providing authorization, authentication, and traceability of trustworthy product data through the product lifecycle [4]

1.2. Digital threats to Additive Manufacturing

Due to the abundance of and reliance on digital data, smart manufacturing is subject to a digital threat. Additive manufacturing (AM) is a process that benefits from the smart manufacturing paradigm, in which a physical object is built by growing material, layer-by-layer, to the desired geometry. Unlike conventional manufacturing methods (such as subtractive processes), AM only requires the design of the physical object and a 3D printer, making manufacturing easier and cheaper. These characteristics and benefits also make AM a very appealing target for hackers. Many cyber-threats have been identified[4]. One such threat is digital product data theft. A stolen design and a low-cost printer are enough to produce counterfeit parts potentially incurring loss of revenues and putting customers at risk. Another cyber-threat is digital product data tampering. Due to the nature of the AM process, a product physical structure can be altered to introduce failure points (i.e., internal voids) without any external modification, making a faulty part almost undetectable. A similar approach can be used to corrupt the manufacturing parameters (e.g., change of material or modified toolpath instructions). This paper focuses on digital product data tampering, at the design and manufacturing levels.

2. Using Blockchain to secure the digital threat

2.1. Introduction to the Blockchain technology

The blockchain[5] is a distributed database that links blocks of data and is operated by a network of anonymous peers. These blocks are timestamped and stored in a linear and chronological order, as seen in **Fig. 1**. Each block contains a set of data, a timestamp, and a hash [6] of the previous block.

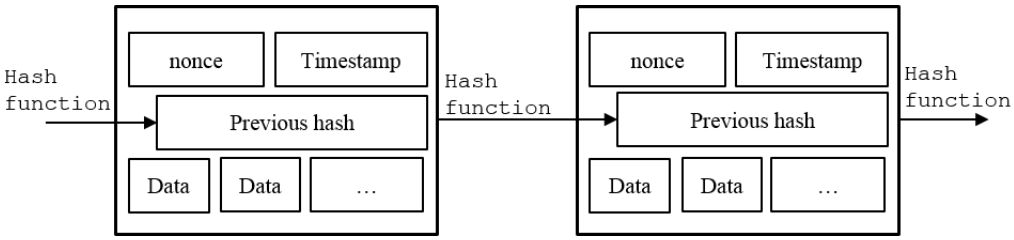

Fig. 1. Illustration of the blockchain principle

Blockchain integrity and trustworthiness are ensured in a two-step validation process. The first step validates the data inside the block against pre-defined domain-specific business rules. For example, bitcoin uses blocks to store financial transactions. A valid block contains transactions that are financially logical – the issuer's credit is still positive after the transaction is processed. The peers must also make sure that the timestamp on the block is within a certain range of the current time.
Because of the distributed nature of the blockchain and latency of the network, peers are often processing different same transactions/data from each other. The second step requires the peers to agree on the (previously) validated data to add in the next block. This agreement is reached through a consensus mechanism[5], preventing malicious peers from adding and/or accepting fraudulent blocks [7].

When two valid blocks (B1, B2) are produced at the same time approximately, the chain needs to eliminate discrepancies or forks. If two peers simultaneously add blocks to the chain, others must use the first block they receive. There are two valid chains (bold chains in left side of **Fig. 2**) at this moment. The next block produced (B3) will only follow one of them, making one chain longer, the official one (bold chain in right side of **Fig. 2**).

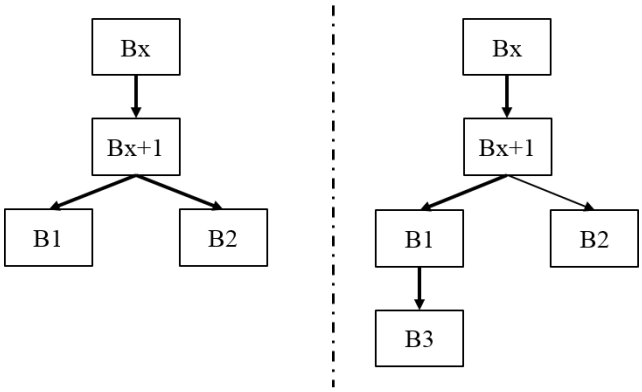

Fig. 2. Blockchain fork resolution

2.2. Reducing the digital threat

Because the chain is tamper resistant and the blocks are timestamped, a blockchain is a robust solution to prove the existence of a specific digital asset at a certain time during the product lifecycle. While the blockchain can act as a distributed-storage mechanism [8], we prefer to avoid this approach for security [9] and latency performance concerns. Any digital asset can generate a unique digital fingerprint using a cryptographic hash function [6]. By storing that fingerprint in the blockchain, one can later prove the existence of the digital asset from when the fingerprint was inserted. Because the asset itself is not revealed, this mechanism can also serve as a data-exchange ledger to record transactions between partners without revealing the content of the transactions. This also helps with proof of integrity: i) a recipient can verify an asset was not altered during the transaction by regenerating the digital fingerprint and comparing it to the one in the blockchain, and, ii) a sender can prove it sent the right asset.

Hedberg, Krima, and Camelio [10] presents a methodology to generate digital signatures of product data using X.509 digital certificates. The generated signature contains a digital fingerprint of the asset that is signed and information about the identity of the signer such as his name and organization. In this paper we present a similar approach in which we secure that fingerprint and associated metadata on the blockchain. Storing the digital fingerprint on the blockchain is a safe way to track both the existence and ownership of a digital asset at a certain time.

Different types of digital assets could be tracked using the blockchain:

- Type A: the blockchain can help to secure proof of existence and ownership of a specific product data that can be critical to solve future engineering and/or legal issues (see **Fig. 3**).
- Type B: by enriching Type A assets with source and destination metadata we can record product data exchanges and demonstrate that a transaction happened.
- Type C: combines Type A and Type B to track physical product ownership and transactions. This could reduce counterfeiting by preventing double-spending [11] of assets.

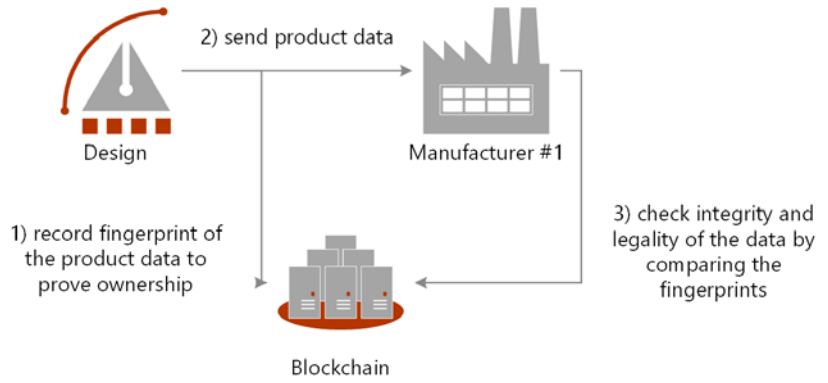

Fig. 3. Using a blockchain ledger to check product data integrity in a simple environment

In a complex and regulated environment, recording Type B assets in a blockchain ledger can be used to identify non-disclosure agreements violations and data breaches through the product lifecycle. When every data transaction is uniquely identified, and recorded in a blockchain ledger, one can easily check the integrity and ownership of the data received, as well as the legality of the data, such as the manufacturer #2 in the last step of **Fig. 4**.

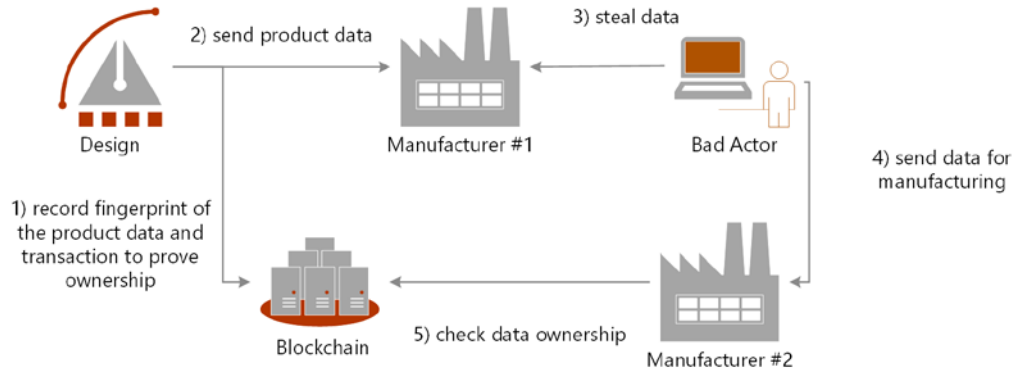

Fig. 4. Using a blockchain ledger to check product data ownership in a complex environment

Our goal is to build a public and trustworthy tamper-resistant repository/registry to record Type B assets. Identifying valid transactions and fingerprints will help identify tampered data before it is used. Such a repository will: i) simplify the traceability of product data transactions due to the immutability of the records, ii) facilitate and automate pre-manufacturing fraud prevention to reduce the complex and expensive post-manufacturing faults detection (using automated and smart contracts), iii) reduce the number of faulty parts distributed, preventing brand reputation damage and loss of revenue from product returns (product not working properly) and liability issues (product endangering its user).

3. Reference information model for a blockchain-based product data traceability

Traceability of product information is a key requirement to secure smart manufacturing supply chain and product lifecycle. As described previously, data tampering can result in damaging consequences to the manufactured/printed part, the system it is installed on, and its owner. One way to mitigate those risks is to record `Data` and `Provenance` information for every single data exchange `Transaction` to: i) ensure the data has not been tampered with, ii) identify if/when the data was tampered with, and iii) track back who tampered with the data.

Our reference information model for blockchain-based solution to support these requirements is divided into the following 3 UML[12] packages:

- `Provenance`: This package contains information about: i) where the data is coming from and going to, ii) who is executing the data exchange, iii) when the data exchange is taking place.
- `Data`: This package contains information about what is exchanged (i.e., type of data)
- `Transaction`: This package contains information about how the data is exchanged (i.e., the data exchange process itself).

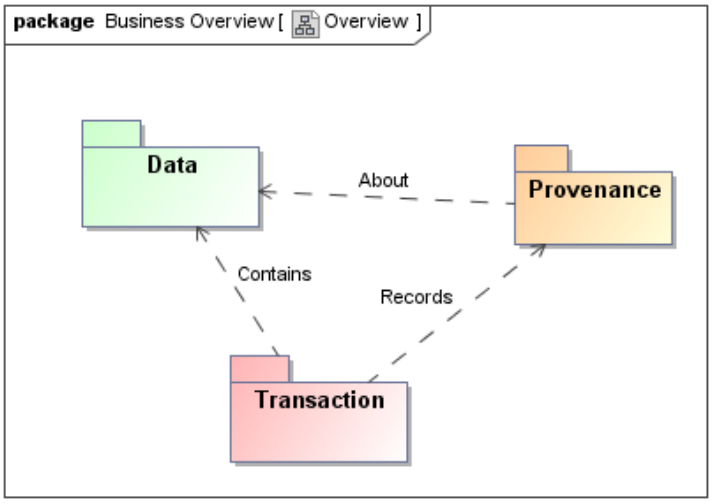

Fig. 5. Reference information model overview

3.1. Data

The `Data` package contains information elements that describe the product data itself:

- The product data content is classified using the **DataType** enumeration. It can describe a 3D model (`GEOMETRY_3D`), a sliced model for additive

manufacturing (`GEOMETRY_SLICED`), a manufacturing system configuration parameters (`CONFIGURATION`) or a set of reference control parameters (`CONTROL_DATA`).

- The product data is a **File** identified by a unique `id`, one or more version identifiers (`versionIds`) and a set of textual `descriptions` related to the content.
- The digital product data (**DigitalAsset**) represent the data being manipulated. It is identified by a unique `fingerprint`. It contains information about the algorithm used to generate that unique fingerprint (`hashAlgo`). A boolean is used to indicate if the content is publicly available (`public`). If it is, the owner can provide a digital object identifier to the content (`doi`), or to more metadata (`doiNoRedirect`).

3.2. Provenance

The `Provenance` package contains information elements that describe the origin and actors involved in the product data transaction:

- A product data transaction is issued by a **Resource,** with a unique identifier (`orgOrPersId`), which can be either an **Organization** or a person in that organization (**PersonInOrganization**).
- A **PersonInOrganization** has an optional first (`firstName`), last name (`lastName`) and `email` but is required to belong to an organization (`belongsTo`).
- An **Organization** has an optional `name` and website `url`, and a mandatory physical `location`.
- A physical location (**Address**) is composed of a `street`, a street number (`streetNumber`), a `town`, a `region`, a postal/zip code (`postalCode`), a postal box number (`postalBox`) and a `country`.

3.3. Transaction

The `Transaction` package record information about the type of data exchange being performed:

- A **Transaction** is identified by a unique id (`txId`), a timestamp (`timeOfTx`) and a payload (`productData`).

Two types of transaction can be recorded:

1. A record of ownership (**RecordOwnership**) is issued by a resource (`issuer`) to claim ownership of a digital asset (`productData`).
2. An exchange of data (**SendProductData**) between a sender (`from`) and a recipient (`to`). A flag (`forward`) can be used to allow the recipient of the data to share the data with others.

3.4. Business rules

The previous sections presented the information artifacts that can be instantiated to represent and describe product data transactions. This section defines a set of business

rules that must be applied to validate instances before they are recorded, in order to maintain a consistent and meaningful repository.

(1) A data exchange can only be initiated by an organization or a person who has previously claimed ownership of the data that is being exchanged. A `SendProductData(SPD1)` transaction is only valid if there is a prior `RecordOwnership(RO1)` transaction such as:

a. `SPD1.from == RO1.issuer AND`
b. `SPD1.productData == RO1.productData AND`
c. `SPD1.timeOfTx > RO1.timeOfTx`

Fig. 7 shows a valid instance of `RecordOwnership(RO1)` that is necessary to create a valid instance of `SendProductData(SPD1)`.

(2) A person or organization can still send data it does not own, if it was explicitly given that right during the initial acquisition of the data. A `SendProductData(SPD1)` transaction can be valid if there is a prior `SendProductData(SPD2)` such as:

a. `SPD2.to == SPD1.from AND`
b. `SPD2.forward == true AND`
c. `SPD2.productData == SPD1.productData AND`
d. `SPD2.timeOfTx < SPD1.timeOfTx`

Fig. 8 shows a valid instance of `SendProductData(SPD2)` that is necessary to create a valid instance of `SendProductData(SPD1)`.

(3) A person or organization can only claim ownership of a data if no other organization or person has previously claimed ownership of the same data. A `RecordOwnership (RO1)` transaction is only valid if there is no prior `RecordOwnership (RO2)` transaction such as:

a. `RO1.productData.fingerprint == RO2.productData.fingerprint AND`
b. `RO2.timeOfTx < RO1.timeOfTx AND`
c. `RO1.issuer == RO2.issuer`

Fig. 9 shows an invalid ownership claim `RecordOwnership(RO1)` because of a prior claim `RecordOwnership(RO2)`.

3.5. Secure and validate product data transactions

The goal of this reference model for product data traceability is to secure the digital thread for smart manufacturing by securing collaboration and underlying data transactions. A secure collaboration requires a consumer to be able to validate data before consuming it (see **Fig. 10**).

For a data consumer to validate a transaction and its content, the content creator must record the transaction on the blockchain before executing it (see **Fig. 11**). The content creator will generate a fingerprint for its data, instantiate the transaction metadata using

the model in **Fig. 6**, record and secure this metadata on the repository and then share the content with the consumer(s).
Prior to consuming data, the consumer must ensure that it has not been tampered and that he/she is the intended recipient of the data (see **Fig. 12**). This validation happens in two steps: 1) one generates a fingerprint from the data received and query the transaction repository to search for the transaction and its metadata. If there is no record returned, one must assume that either the transaction was not properly registered or the data was tampered; 2) If a metadata record is retrieved from the repository, the consumer can look at the provenance, data and transaction information to make an educated decision based on the content expected.

4. Conclusion

The smart manufacturing initiative relies on digitization of the product data to speed up engineering activities. This digitization generates a significant amount of data that is exchanged between the different actors and systems involved in the product lifecycle. Trustworthiness is key and only authentic and valid data should be consumed[10]. We presented a blockchain-based solution to secure and authenticate product data. Due to its tampering resistance, blockchain is an ideal candidate to record and secure data exchanges. We presented a reference data model that represents a set of data exchange metadata necessary to identify invalid transactions and tampered product data that should not be consumed.

References

[1] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System.”

[2] T. Hedberg Jr., J. Lubell, L. Fischer, L. Maggiano, and A. Barnard Feeney, “Testing the Digital Thread in Support of Model-Based Manufacturing and Inspection,” *J. Comput. Inf. Sci. Eng.*, vol. 16, no. 2, p. 21001, Mar. 2016.

[3] B. Kulvatunyou, N. Ivezic, and V. Srinivasan, “On architecting and composing engineering information services to enable smart manufacturing,” *J. Comput. Inf. Sci. Eng.*, Jun. 2016.

[4] L. D. Sturm, C. B. Williams, J. A. Camelio, J. White, and R. Parker, “Cyber-physical vulnerabilities in additive manufacturing systems: A case study attack on the .STL file with human subjects,” *J. Manuf. Syst.*, vol. 44, pp. 154–164, 2017.

[5] D. Yaga, P. Mell, N. Roby, and K. Scarfone, “NISTIR 8202: Blockchain Technology Overview,” 2018.

[6] P. Rogaway and T. Shrimpton, “Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Preimage Resistance, Second-Preimage Resistance, and Collision Resistance,” in *Fast Software Encryption*, 2004, pp. 371–388.

[7] C. Dwork and M. Naor, “Pricing via Processing or Combatting Junk Mail,” in *Proceedings of the 12th Annual International Cryptology Conference on Advances*

in Cryptology, 1993, pp. 139–147.

[8] A. Miller, A. Juels, E. Shi, B. Parno, and J. Katz, “Permacoin: Repurposing Bitcoin Work for Data Preservation,” in *2014 IEEE Symposium on Security and Privacy*, 2014, pp. 475–490.

[9] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, “Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts,” *Cryptology ePrint Archive, Report*, pp. 1–32, 2015.

[10] T. Hedberg Jr., S. Krima, and J. A. Camelio, “Embedding X.509 Digital Certificates in Three-Dimensional Models for Authentication, Authorization, and Traceability of Product Data,” *J. Comput. Inf. Sci. Eng.*, vol. 17, no. 1, pp. 11008–11011, Nov. 2016.

[11] J.-H. Hoepman, “Distributed Double Spending Prevention,” *ArXiv e-prints*, vol. abs/0802.0, 2008.

[12] Object Management Group, “Unified Modeling Language,” 2017.

Appendix A: Information Model UML Class diagram

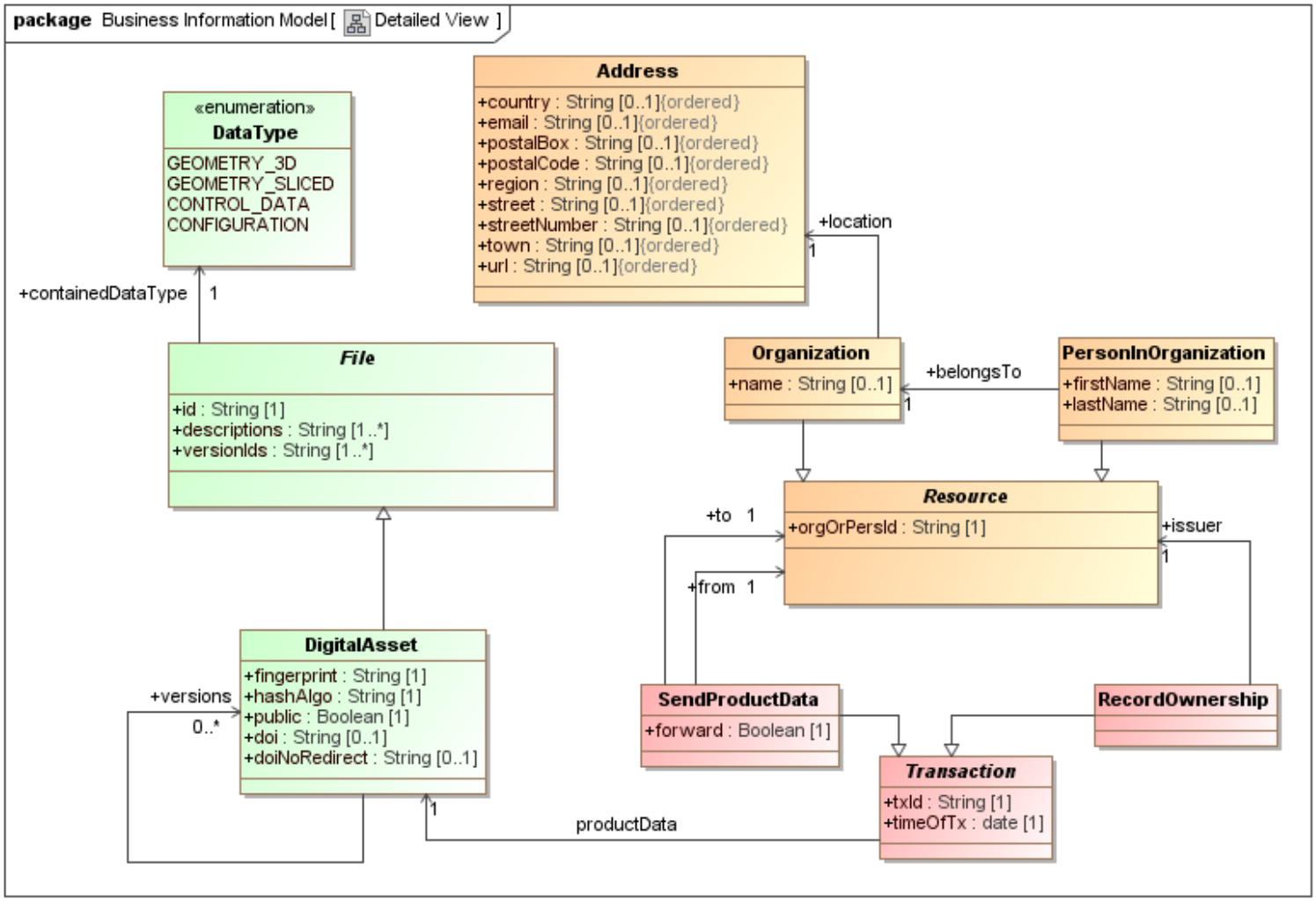

Fig. 6. Reference model for blockchain-based product data traceability

Appendix B: Business rules UML Object diagram instantiations

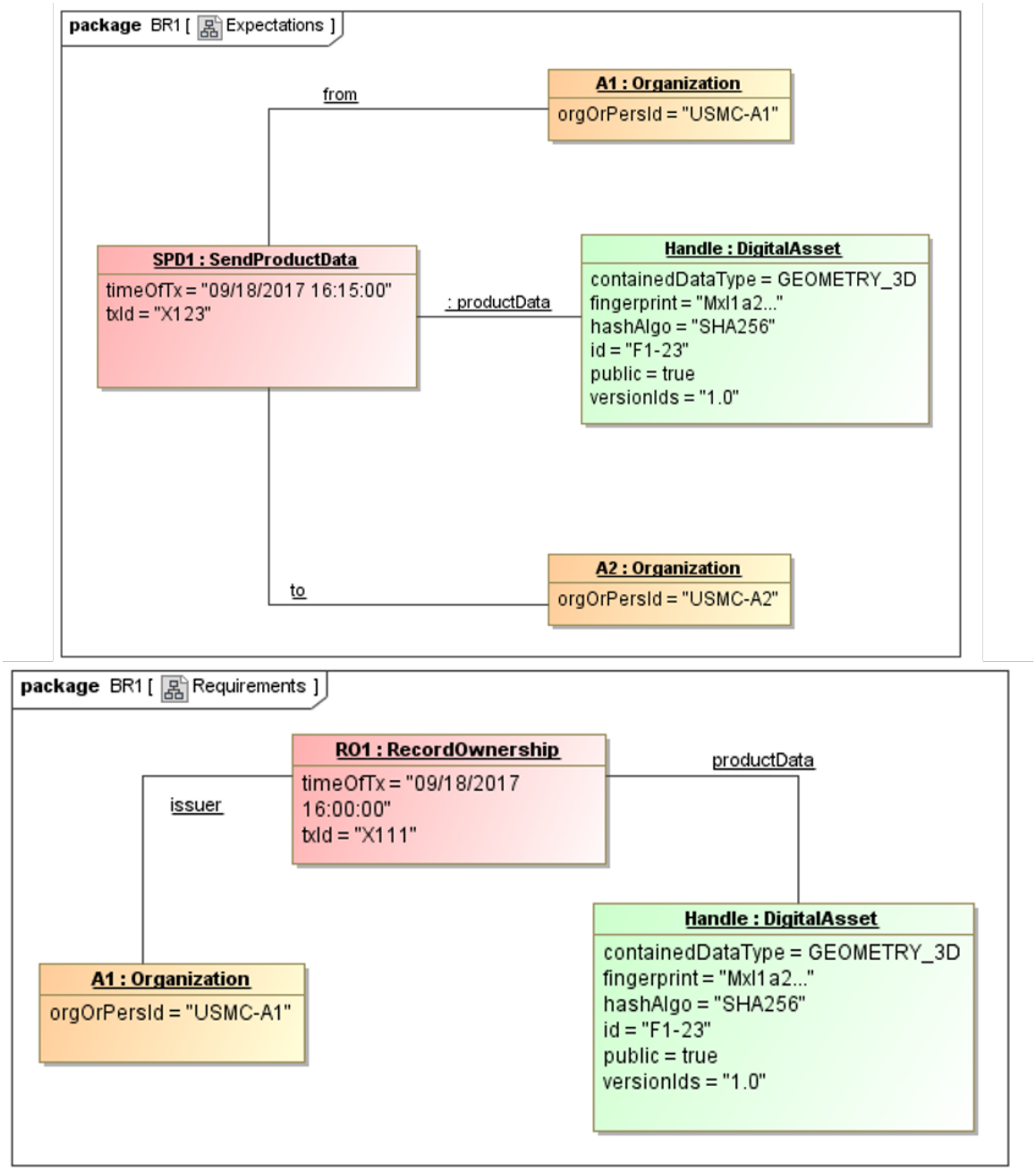

Fig. 7. Business Rule (1) instantiation

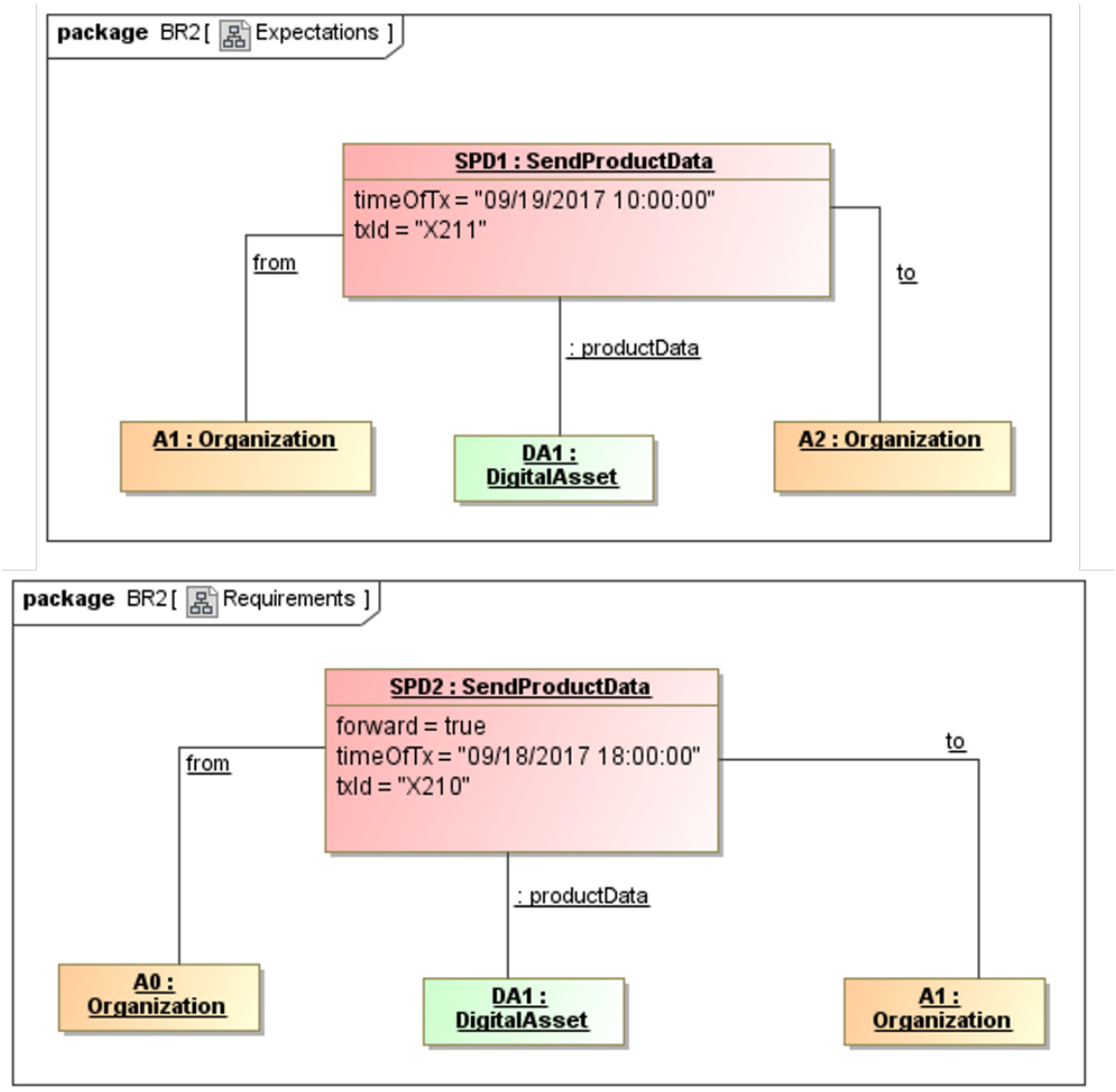

Fig. 8. Business Rule (2) instantiation

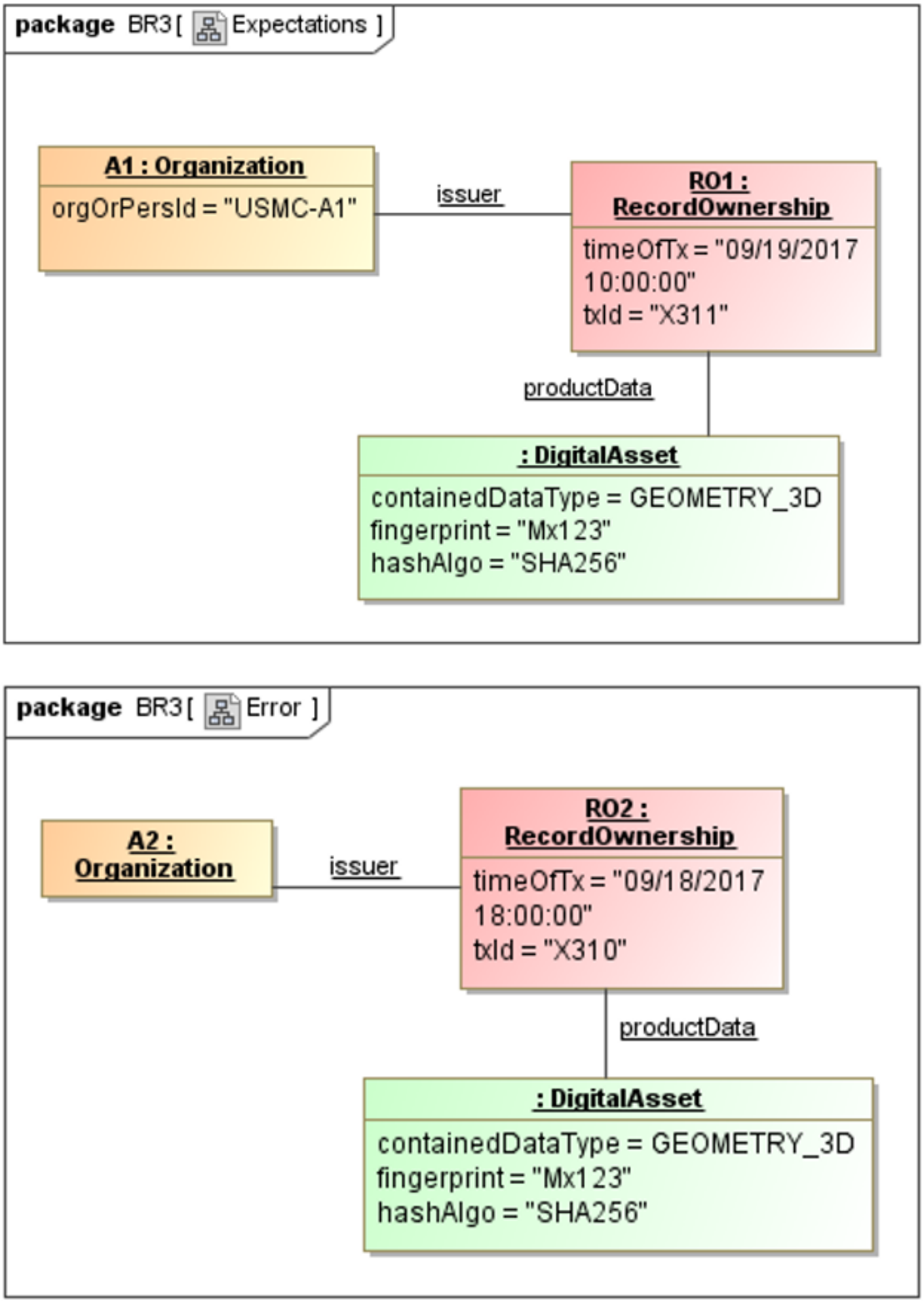

Fig. 9. Business Rule (3) instantiation

Appendix C: BPM Business process

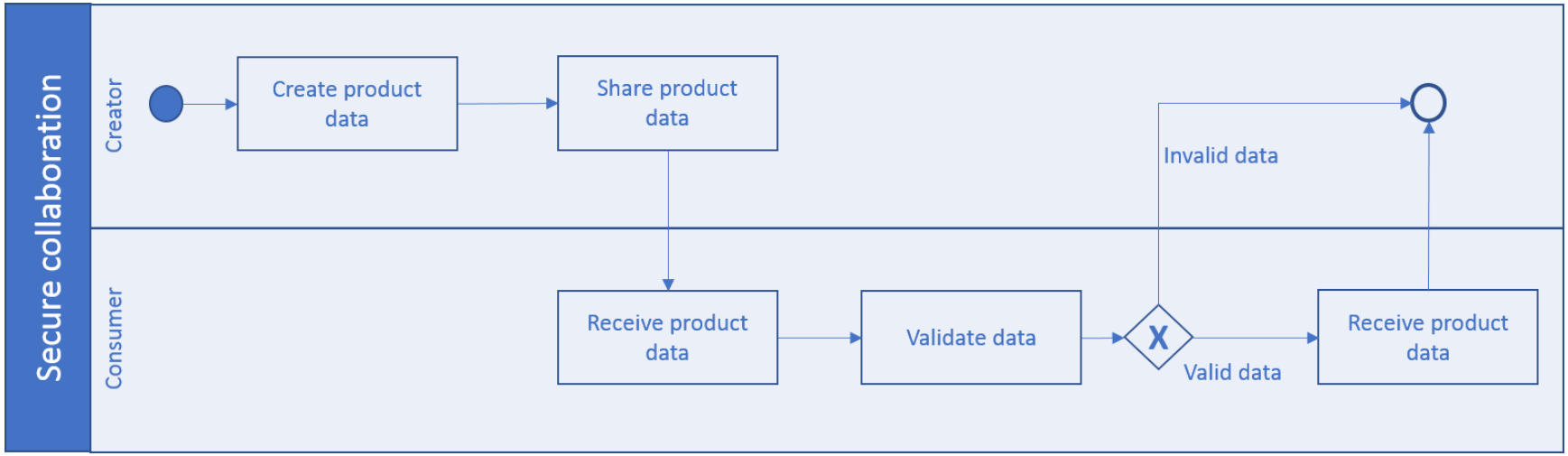

Fig. 10. Secure collaboration and transaction business process

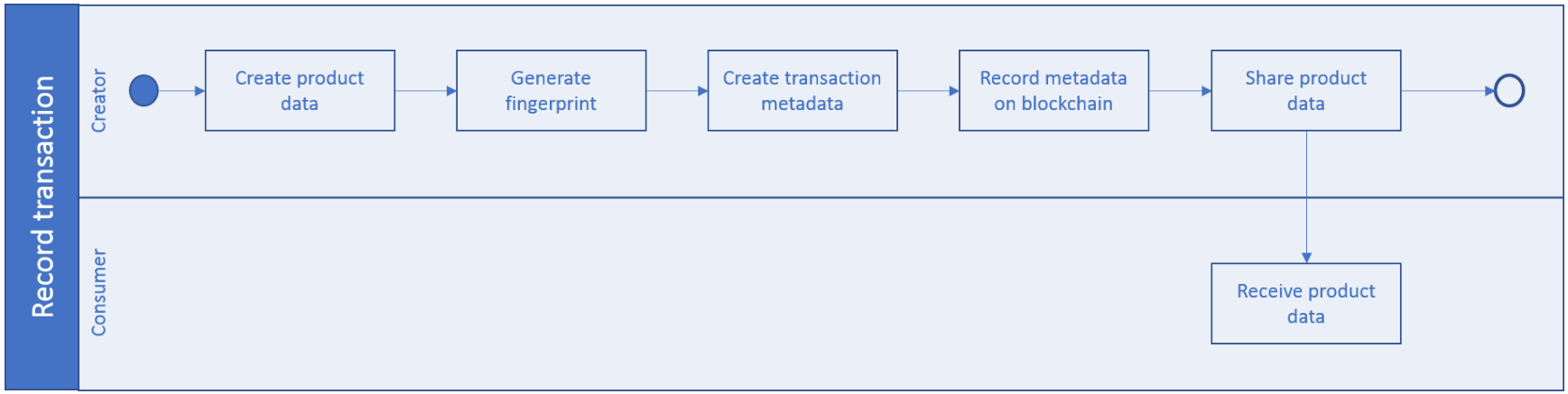

Fig. 11. Recording a transaction on the blockchain

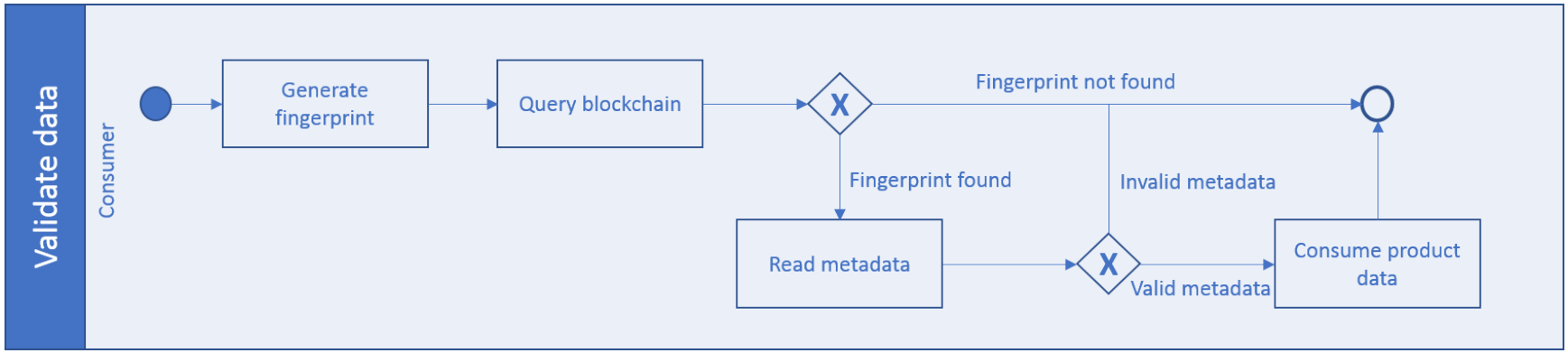

Fig. 12. Validating a transaction using the blockchain