# Interoperable Traceability in Supply Chains: A Use Case in Agritech

Giovanni Farina[8], Gianluca Brunori[1], Stefano Chessa[1],
Alexander Kocian[1](✉), Maria Bonaria Lai[1], Daniele Nardi[2],
Claudio Schifanella[3], Susanna Bonura[4], Nicola Masi[4],
Fiorenzo Ambrosino[5], Angelo Mariano[5], Lucio Colizzi[7],
Giovanna Maria Dimitri[6], Marco Gori[6], Franco Scarselli[6],
and Silvia Bonomi[2]

[1] University of Pisa, Pisa, Italy
`{gianluca.brunori,stefano.chessa,alexander.kocian}@unipi.it`,
`mariabonaria.lai@agr.unipi.it`
[2] Sapienza University of Rome, Rome, Italy
`{nardi,bonomi}@diag.uniroma1.it`
[3] Computer Science Department, University of Turin, Turin, Italy
`claudio.schifanella@unito.it`
[4] Engineering Ingegneria Informatica S.p.A, Rome, Italy
`{susanna.bonura,nicola.masi}@eng.it`
[5] ENEA Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Rome, Italy
`{fiorenzo.ambrosino,angelo.mariano}@enea.it`
[6] DIISM, University of Siena, Siena, Italy
`{giovanna.dimitri,marco.gori,franco.scarselli}@unisi.it`
[7] Dipartimento di Informatica, Università degli Studi di Bari, Bari, Italy
`lucio.colizzi@uniba.it`
[8] Department of Engineering, Niccolò Cusano University, Rome, Italy
`giovanni.farina@unicusano.it`

**Abstract.** In the agri-food sector, traceability is essential to ensure the quality, safety, and transparency of supply chains, where transportation companies are key stakeholders in the overall movement of goods. The multitude of actors involved in supply chains makes it challenging to achieve the above mentioned objectives: each company usually uses its own information system, which is mainly aimed at tracking all relevant information of the single company and is rarely able to interact with other information systems. In addition, when multiple information systems need to exchange data, it is essential to have full control over what data is exchanged and who has access to it.

---

This study was carried out in part within the Agritech National Research Center and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) - MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 - D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

A. Kocian et al. (Eds.): INTSYS 2024, LNICST 608, pp. 29–45, 2025.
https://doi.org/10.1007/978-3-031-86370-7_3

In this work, we introduce the concept of *interoperable traceability*. We have developed an innovative model that enables the seamless integration of data from IoT devices, data management software, and distributed ledgers into a newly designed "data space". We provide an implementation of this concept that maps to a practical use case and provides a demonstrator that facilitates the secure transfer of traceability data between existing systems. This gives stakeholders a whole new way to manage and review data with greater transparency and efficiency.

**Keywords:** International Data Space · Blockchain · Internet of Things · Machine Learning · Intelligent Transport Systems · Agrifood

## 1 Introduction

Agri-food supply chains are logistic networks with particular constraints, such as the perishability and seasonality of goods, that require highly reliable information, communication, and transportation systems. Along with goods, a variety of information is stored and exchanged, and traceability in this sector is essential to reduce some of the associated potential health risks. *Traceability* is defined in [17] as the ability to access any or all information on a product considered throughout its life cycle by recording identification. Traceability, which promotes the culture of sustainability through the responsibility of all actors, becomes a key open-handed transparency of the Agri-Food System (AFS). Transparency provides a common understanding among stakeholders who can access product information without loss, noise, or distortion [20]. AFS benefits from a transparent supply chain in several ways, including improving market efficiency, enhancing information sharing among all supply chain stakeholders, promoting food quality communication, supporting product differentiation, and improving optimization of logistics activities and business processes [20]. Today, the traceability of agri-food products has become an increasingly interesting topic within the technological field. Digital technologies can play an important role in improving traceability and transparency by ensuring the collection of comprehensive, consistent and reliable data along the food supply chain, together with real-time tracking, easy aggregation, integration, analysis and sharing of data [7]. To establish a traceability system along the food supply chain, it is necessary to be aware of a broader ecosystem of actors and infrastructures and to establish a set of agreements with the actors involved [4]. The concept of "Food traceability 4.0" was recently proposed by Hassoun [10] and refers to the implementation of a smart traceability system from farm to fork using Industry 4.0 technologies, specifically blockchain, Internet of Things (IoT), Artificial Intelligence (AI), and big data analytics [11]. The same technologies are beneficial and should be integrated in intelligent transportation systems, which are fundamental actors in supply chains: there is often a co-presence of a multitude of companies, each usually with its own information system to support the business, and there is a need to make this data available by having the ability to control what data is made available and who can access it [16].

This paper presents a unified federated data lake that integrates data from IoT devices, data management software, and distributed ledgers. Product information is shared in an *interoperable fashion* across different systems and platforms, ensuring consistent and transparent data exchange throughout the supply chain. The proposed model enhances the seamless exchange and verification of traceability data, improving transparency and efficiency across the supply chain. The implementation of this concept is demonstrated through the METRIQA platform, which supports the digitization of agri-food processes by providing a secure and interoperable environment for data exchange. The platform's architecture includes components for data storage, decision support, and user interfaces, leveraging technologies such as blockchain, AI, and big data analytics. Three use cases illustrate the practical applications of interoperable traceability in wine production, automated data sensing, and cattle tracking, highlighting the benefits of improved data integration and fraud prevention. Although the use cases described focus primarily on goods producers and transformers, the solution presented can be extended to any actor in the supply chain, including transport companies.

## 2 Background and State-of-the Art

Supply chain traceability has become an important area of research due to increasing global complexity, regulatory requirements, and consumer expectations for transparency. It involves tracking the flow of goods, materials, and information across multiple stages of production, transportation, and distribution. With the advent of Industry 4.0, many disruptive technologies have been successfully deployed in traceability systems across different sectors. Many of these approaches have been proposed in the agri-food sector, given the high demand for transparency and safety not only from customers, but also from companies in the supply chain who want to increase the perceived value of their products.

*Internet of Things (IoT)* plays a key role in capturing real-time data across the supply chain, including environmental conditions, location tracking, and product handling information. By integrating IoT with other technologies, supply chains can achieve end-to-end traceability [15,18]. Blockchain and distributed ledger technology (DLT) enable transparency, security, and certification of supply chain information: they allow all stakeholders in a supply chain to verify and trust data without the need for a central certification authority. The effectiveness of the role of DLT in traceability systems has been demonstrated in several works [2,19]. Artificial intelligence, and more generally machine learning, can detect patterns and inefficiencies, predict problems, and suggest corrective actions [1]. Data analytics enables the analysis of big data generated by modern traceability systems, enabling data-driven decision making.

*Blockchain* technology has emerged as a revolutionary tool in various sectors, including agri-food chains, by offering unparalleled traceability and transparency. Today, blockchain applications are focused on improving interoperability across different platforms to ensure seamless data exchange and integration. Moreover blockchain immutability ensures that data cannot be altered or deleted, preventing fraudulent activities such as mislabeling, counterfeiting, or product substitution. Its decentralized nature and use of cryptography provide enhanced security against various cyber-attacks. This security protects the interests of all stakeholders, from farmers to retailers. Furthermore, blockchain technology enables the sharing of accurate real-time tracking information throughout the supply chain. Recent advances have seen the development of decentralized platforms that not only facilitate secure data sharing, but also enable the integration of Internet of Things (IoT) devices and smart contracts, further enhancing the traceability and efficiency of supply chains. In addition, efforts are being made to standardize blockchain protocols to enable smoother interoperability between different systems, addressing one of the key challenges to widespread adoption in the agrifood sector.

*Data spaces* are a rapidly evolving data sharing concept introduced in 2005 by [8]. The concept of data spaces introduces a drastic change to the systematic and centralized perspective of databases, relying instead on decentralized collections of heterogeneous data. According to the recent European Strategy for Data[1], the creation of a single market for data in Europe will ensure Europe's global competitiveness and data sovereignty, and a key point is the creation of Common European Data Spaces. We refer the interested reader to [6] for a comprehensive presentation of data spaces. More specifically, we focus on an emerging specification for data spaces, namely the *International Data Space (IDS)* [13]. This specification responds to requirements of data sovereignty, trustworthiness, and monetization, and it supports the implementation of different business models centered on data exchange, by enabling data sharing through data spaces characterized by uniform rules, certified data providers and recipients and trust among partners guaranteed by certified components and secure interactions. To define a minimum viable data space, the IDS specification defines a limited set of components: two or more IDS connectors with data source and data provider functionality for data exchange; a Certificate Authority (CA) that provides X.509 certificates to ensure the integrity and authenticity of data exchange; and a Dynamic Attributes Provisioning Service (DAPS) for access control and data security. Details on the implementation of IDS can be found here [13].

IoT, sensors, and computer information systems strongly support the traceability processes of individual supply chain actors, i.e. keeping track of all information relevant to their business. In most cases, these traceability systems are not designed to interoperate with external systems; their primary goal is to store and make available the information to a single stakeholder. In particular, proprietary systems and ad hoc data formats are often used. In other words, the main barriers to full supply chain traceability are the variety of ways data

---

[1] https://digital-strategy.ec.europa.eu/en/policies/strategy-data.

is encoded or transmitted, motivating the need for integration technologies to support interoperability, such as distributed ledger technologies and data spaces.

## 3 METRIQA Digital Information Platform

The METRIQA (MEasurements, TRaceabilIty, and Quality in Agri-food Chains) platform has been developed within the Italian National Research Center for Agricultural Technologies (AGRITECH). The latter is a national implementation of the European Green Deal, to achieve climate neutrality by 2050.

METRIQA supports the digitization of the agrifood sector in Italy in two ways. On the one hand, the platform offers services to research in agri-food by supporting the entire chain of data and information production, storage, and analysis typical of an intelligent environment. Experimental data sets are generated by IoT sensors as well as by researchers during on-site experiments. These data are stored in a distributed fashion and processed by AI tools, to create innovative decision support systems. On the other hand, METRIQA provides a seem-less web-like access to data and services. Additionally, METRIQA offers an open infrastructure for collaboration and innovation in the agrifood sector.

The METRIQA reference architecture, depicted in Fig. 1, is composed of three layers: Storage, Components, and Applications [3,5]. Starting with the Storage layer, METRIQA implements a message broker and a data lake for the (research) data generated by the project partners. At the Components layer, METRIQA introduces a decision support mechanism that facilitates web-based data retrieval services, including a natural language question-answering system. It also incorporates decision support modeling notebooks that utilize scripting languages like Python for data exploration. Additional features include metadata enhancement, Software as a Service (SaaS), Platform as a Service (PaaS), and containerized support for decision support systems crafted by Spoke researchers. At application layer, there are all the User Interfaces (UIs) allowing the end-user to access and interact with the platform. The Application Programming Interfaces (APIs) streamline the data exchange between providers and consumers.

The Blockchain, Security and Data Space components are cross-layer components. Specifically, the blockchain is used in combination with traceability services, and the data space component provides all the elements that support the functioning of the data space according to the International Data Spaces (IDS) specifications [13].

## 4 Interoperable Traceability and Use Cases

The key objectives - and related performance indicators - of traceability processes are related to the fluidity and accuracy of information flows between the actors that populate the supply chain. The main barriers to these objectives are related to the diversity of the ways in which data are codified or transmitted, as actors follow their own codifying rules and use their technologies. This problem is addressed through the concept of *interoperability*, which can be defined as

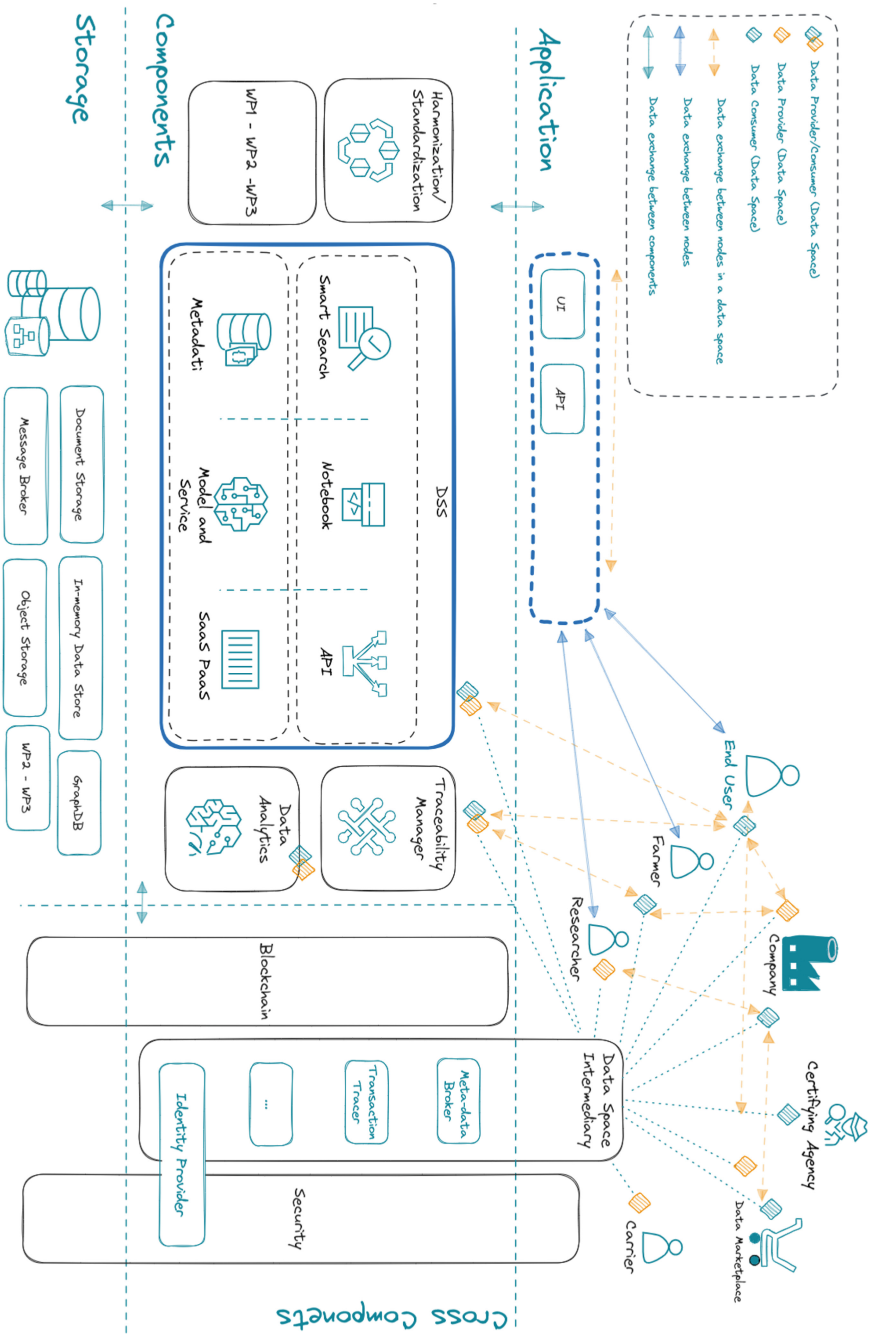

**Fig. 1.** The reference architecture of METRIQA

the ability of systems to exchange and use information, that encompasses multiple dimensions, including legal organizational, semantic, and technical aspects [12]. Although internally for a company the process to achieve interoperability can be based on the power of the hierarchy, when traceability involves independent actors, interoperability requires the development of governance systems. Strongly coordinated supply chains are more likely to have interoperable traceability systems than loosely coordinated ones, as the chain leader can impose its traceability systems and technologies. For smaller firms, the adoption of the chain leader's traceability system could imply a loss of autonomy, as the shift to another customer could imply a switch to different traceability systems.

The concept of Interoperable Traceability aims at addressing this problem. We aim at building a bundle of technologies, namely data spaces, blockchain, IoT, and AI to respond to the needs of interoperability. In this regard, we propose a model that enables the integration of data from IoT devices, data management software, and distributed ledgers into a data space—specifically, an instance of the International Data Space [13] suitable for the agrifood sector—that allows anyone with the appropriate access rights to use, integrate, enhance, and transfer traceability data into their own traceability systems.

To illustrate how the development process aligns with user needs, improves communication, guides design and testing, and manages project scope, consider three use cases that cover different aspects of interoperable traceability.

**Use case 1:** ***Enabling Interoperability in Wine Supply Chain Traceability Systems*** - A wine producer intends to sell wine with data quality to any wine center.

Suppose a wine producer that records various aspects of the wine production process in its information system to ensure quality, authenticity, and traceability. This includes data related to the vineyard, harvest, winery level, packaging, bottling, etc. The sales of its products to different wine distributors, each with its own supply chain management (SCM) system, implies the need to connect to each of these SCM to exchange any relevant information related to traceability along all these supply chains. The lack of interoperability between these systems may become cumbersome for the wine producer, as the technical solutions to implement such a connection imply the development of ad hoc solutions, such as the use of ad hoc middleware for data transformation and APIs for real-time data exchange and/or adaptation to different data formats. The concept of interoperable traceability would instead provide a seamless way to connect these systems, resulting in increased efficiency and reduced costs.

**Use case 2:** ***Automatically sensed data to support supply chain traceability*** -

Automated data sensing increases the reliability and visibility of supply chain operations, making it easier to track and verify the quality, origin, and handling of products, increasing consumer confidence and operational efficiency. Artificial intelligence transforms this data into product quality information for the consumer.

Consider a field equipped with a number of IoT nodes to accurately monitor critical soil and crop substrate parameters in real time to optimize water and fertilizer use. The quality of the data collected by these sensors, which measure key variables such as soil moisture, temperature, nitrogen, phosphorus, potassium, and pH, is critical. These real-time measurements are transmitted to a back-end system, where they are stored in a database and certified through a blockchain system for greater transparency and data integrity. The accuracy and security of this data is critical, as decisions throughout the supply chain depend on it. Consider for example the case of fertigation: when its cycles are managed using high-quality real-time soil and crop data, water and nutrients are applied in optimal amounts at the right time, avoiding inefficiencies and potential over fertigation. The latter, especially with nitrogen-based fertilizers, can lead to the accumulation of nitrates in the soil. These nitrates are absorbed by the plants and can eventually be transferred to the final product, compromising its quality. High nitrate levels in crops are undesirable as they pose a risk to both product safety and consumer health. Therefore, accurate data ensures that irrigation and fertigation cycles are not only efficient, but also safe, keeping nutrient levels balanced and within acceptable limits. As a result, manufacturers who use high-quality data can deliver healthier products with less environmental impact, adding value to their products and building trust with consumers and retailers. Machine learning transforms raw sensor data into traceability information by analyzing and interpreting patterns within the data to gain key insights such as the product's origin, handling conditions, and journey. This traceability information is then presented in a consumer-friendly format, enabling transparency and trust in the quality and authenticity of the product.

**Use case 3:** ***Interoperable cattle traceability for parts of the food supply chain with fraud prevention*** - From Use Case 1 and Use Case 2, the question arises of how to properly combine interoperable traceability with fraud prevention mechanisms.

Consider, as an example, a typical cow grazing in an alpine pasture, equipped with a smart collar that collects position data stored in the farmer's information system, with the purpose of providing traceability data about the animal's pasture location once it enters a given supply chain. Note that in this process, the GPS data can be exchanged in an interoperable manner by connecting to the appropriate data space of the specific supply chain, and that the GPS data can go through a data analysis process to produce information in a consumer-friendly format, as already explained in the previous use cases. Finally, the user accesses information about the origin of the product (say, a steak) through a QR code on the consumer product. However, in this context, the livestock farmer may fraudulently claim ownership of other people's cattle and sell them as his own; or falsely claim compliance with feeding standards (e.g., hay instead of high alpine grass). Thus, in addition to the issues already discussed in use cases 1 and 2, there is also the issue of fraud prevention. To deal with this case, we need to ensure that the GPS data generated by the collar and stored in the farmer's information system is not altered, and this can be achieved by combining the use of a tamper-proof collar with an integrity code of the GPS data that the collar should be able to store in the blockchain through a channel outside the farmer's control. This latter channel can be achieved, for example, when the cow passes through the hands of the next actor in the supply chain, who can read the integrity code of the collar and store it in an appropriate smart contract in the blockchain, along with the certification of this exchange of the cow.

## 5 Design and Development of a Live Demo for Interoperable Traceability

### 5.1 Overview

In this section, we present the demonstrator based on Use Case 3 of Sect. 4 that we are currently developing to illustrate the concept of interoperable traceability. We resemble parts of the food supply chain. Specifically, we focus on the transition from production to processing (and slaughter) of a typical cow in Italy. The functional blocks of the demonstrator are outlined in Fig. 2 and detailed in the following.

We start with the data producer, the cattle farmer. The key components are a *IoT collar* equipped with a GPS sensor connected to an IoT client. Raw location data and its hashes, as well as the ID of the owner and the particular cow, are stored in a local database, the *Farm System*. Suppose that the cattle farmer intends to sell one of his cows to slaughter in the (interoperable) food supply chain. The slaughter opens a dedicated application on the cellular phone and reads out the QR code on the collar. Three types of data flow will occur: *i)* the time series of (hashed) GPS data is sent to the federated dataspace component (described in Subsect. 5.4); *ii)* the daily hashed GPS data is sent to the *Blockchain* smart contract (described in Subsect. 5.3); *iii)* In Italy, the change of ownership of cattle must be reported to the national database (Banca Dati Nazionale - BDN), but this transaction is outside the scope of this demo.

The federated data space component of the METRIQA platform enables secure data exchange between connectors within the AGRITECH Data Space. The process involves deploying core components, ensuring communication between participants, and validating data transfer scenarios. The connector to the data space facilitates data exchange while ensuring data sovereignty.

The data consumer, that is, the *Big Data Analytics* (BDA) component cleans, organizes, and explores historical data before applying machine learning algorithms. The *Machine learning* block tracks the movements of the cow and sets a flag in case the cow was not grazing where it was supposed to be. Similarly, the *Consistency check* performs blockchain-based continuous data integrity checking and sets a flag in case of error. Finally the *Traceability system* block collects all relevant information from the *Machine learning* block as well as the *Consistency check* block via *API Gateway*, in order to be capable of replying to the consumer upon product request via another QR code.

All components of the system operate in real time, except for the cow's movements in space and time. To facilitate reproducibility, the latter has been realized as a theoretical mathematical model.

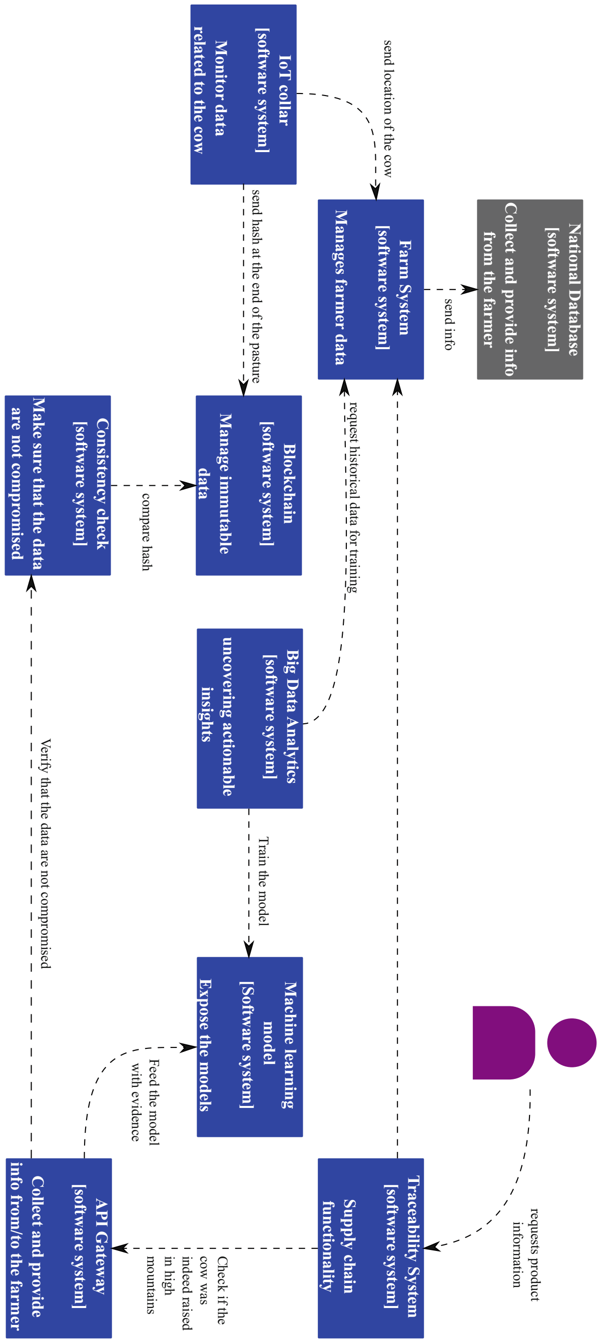

**Fig. 2.** System architecture of our live demonstrator, comprising IoT client, private Ethereum Virtual Machine, a connector to the International Data Space and tools for data analytics and Machine learning

### 5.2 The METRIQA IoT Support for the Live Demo

To get accurate location data, we could simply equip the cow with a smart collar with GPS connectivity. The GPS data is collected periodically, say every 15 min, and recursively hashed (SHA-256). The result is sent to the dataspace API (see Sect. 5.4). At the end of the day, the GPS history of the cow, along with the identifying information of the cow as well as the farmer, i.e. the production data, is embedded in another hash (SHA-256), the transaction is signed with the private key and recursive length prefix (RLP) encoded in binary format. The result is sent to the cloud node where the smart contract resides (see Sect. 5.3).

To make the sensed data reproducible, we decided to emulate the IoT client, i.e., we rely on Markov chain models to simulate the evolution of GPS positions (such as shelter, exploration, foraging) according to [9,14].

### 5.3 The METRQIA Blockchain for the Live Demo

The METRIQA platform integrates a cross-layer blockchain, which provides a programmable decentralized ledger, and a data space component, which regulates the exchange of information between data producers and consumers. Given the nature of the two METRIQA components, by default, the blockchain component should only store data that can be permanently available to all blockchain participants, while the Data Space component should regulate the exchange of all other data.

The blockchain implementation considered for the use case example is *permissioned*, meaning that its access is limited to authorized participants (i.e., accounts) with defined roles, each of which is associated with an identified *entity* (a user, a device, etc.). This is necessary to ensure the authenticity and non-repudiation of the stored data. Each entity is provided with a credential that allows only its owner to perform operations on behalf of the entity. The accounts allow for *pseudo-anonymity* of the entities that interact with the blockchain. This means that the entity associated with each account is not stored in the blockchain by default, but all its interactions with the DLT are auditable. The association between accounts and entities can be made available or not, but it is required for the considered use case. The DLT implementation consists of 3 Hyperledger Besu nodes; this blockchain is compliant with the Ethereum Virtual Machine and uses the Clique consensus mechanism, that is based on proof of authority.

The objectives of the considered use case are pursued through the definition of a *smart contract*, namely a software defined on top of the blockchain network that regulates the ledger updates. More specifically, the smart contract stores data, the *state*, and it exposes a set of interfaces that allow its update, the *functions*. In addition, a smart contract makes it possible to define which accounts can interact with it and how. For the use case under consideration, the smart contract, that is referred to as *Demo-Contract*, keeps track of *1)* a hash (fixed-size string of bytes that is unique to the specific input data) associated with the historical data of each cow's pasture, and *2)* the changes of ownership of a cow. More specifically, the smart contract stores only the hash associated with the cow's pasture data for several reasons: it is not necessary to guarantee the availability of this data, and the amount of data to be stored on a distributed ledger should be minimized; the owner of the cow's pasture data has control over which entity is allowed to access this data. On the other hand, ownership changes are data that must be available to data consumers, motivating its storage in the distributed ledger. Note again that the distributed ledger provides pseudo-anonymity by default; it follows that the association between account and entity (owner) must be made available in order to properly reconstruct all ownership transfers.

The DEMO-contract assigns a *role* to each account that interacts with it. Four roles are defined in order to assign different privileges to different users: *NoPriviledge*, *Administrator*, *Actor*, and *Device*. NoPriviledge is the default role assigned to each account interacting with the DEMO-contract. The Administrator role is assigned to the Contract Deployer (i.e., the account that distributes the smart contract to the DL) and can be extended to other accounts by the initial Administrator. The Administrators in this example have the ability to assign roles to all accounts. The Actor role is the one associated with any account that is controlled by a human. In the DEMO contract, these accounts have the ability to control the devices associated with them and to transfer ownership of the cow they own. The Device role is associated with all IoT devices that interact with the smart contract and are the only ones allowed to register data about the cow's pasture. The DEMO contract state keeps track of 1) the role associated with each account; 2) the ownerships associated with each registered cow; 3) the association between devices, cows, and current owner. Each function can be associated with preconditions, i.e. requirements that must be met in order to execute the functionality. This makes it possible to define access control, i.e. which accounts are allowed to perform certain operations and which state condition must be met in order to update the ledger. More specifically, to name a few examples, the role assignment functionality can only be performed by the administrator account; cow registration and device association can only be executed by the actor account; cow pasture data registration can only be operated by the associated device.

The transfer of ownership functionality requires both the sender and the recipient to acknowledge the transfer, i.e., the transfer is not complete until the recipient confirms the transfer of ownership. Cow pasture integrity data are not stored directly in the states of the contract, but are tracked as events. An *event*, in the context of smart contracts, is a mechanism that allows the contract to communicate that something specific has happened on the blockchain and its occurrence is stored in the DL (outside of the contract state). As a result, all accounts that interact with a smart contract are able to retrieve all events that have been emitted since the contract distribution. This is an alternative mechanism for storing data associated with a Smart Contract and allowing historical data to be offloaded from the contract itself. The DEMO-contract is only responsible for regulating which accounts are allowed to store the cow's pasture-related data and for properly storing a hash associated with a timestamp and a cow.

The software component interacting with the smart contract is responsible for properly generating the data to be stored in the DL.

All requests to update the ledger are made through *transactions*, whose acceptance and execution are governed in a distributed manner by a consensus algorithm. In blockchain terms, a transaction is a signed packet of data that changes the state of the blockchain. It is the primary way to initiate actions on the DL network, especially to interact with a smart contract. To simplify the interaction with the DL and the DEMO-contract, a *Node.js* software component has been developed, which is referred to as *Interaction-Component*. The Interaction-Component provides three main services: it allows to trigger all the functions available on the DEMO-Contract; it simulates the functionality of a smart collar that tracks a cow's pasture data and ensures its integrity (Sect. 5.2); it exposes a remote interface to verify the integrity of the cow's pasture data. The Interaction-Component allows the integrity of the cow's pasture data associated with a specific cow to be remotely verified through a REST API.

### 5.4 The METRIQA Data Space for the Live Demo

The implementation of data space in METRIQA aligns with the IDS specification, and incorporates all its core components. These data space components have been deployed and tested to validate the communication between participants, implementing the scenario of a farmer as a Data Provider that shares GPS data collected by a IoT device (collar in a cow) and a Data Analytics system (as a Data Consumer) to collect data and train an ML model (as in Use Case 3 presented in Sect. 4).

The components are deployed within a Kubernetes[2] cluster, to achieve application scalability and resilience by automating container orchestration, and by ensuring the optimal utilization of resources and high availability of the METRIQA IDS. The adoption of Kubernetes is also motivated in our case to the need of simplifying our deployment and management processes, and of facilitating the continuous integration and delivery and accelerating development cycles. Within the cluster, Minio[3] is deployed as an object storage solution, enabling connectors to access stored data. In our specific demonstrator, we have deployed two connectors: a provider and a consumer. The provider accesses data from Minio once the Data Owner registers it as an Offered Resource on the connector.

The connector is implemented by the TRUE (TRUsted Engineering) Connector[4], which is an open-source connector developed by ENGINEERING and is also part of the FIWARE catalog. Furthermore, this connector adheres to the latest IDS specifications (e.g., IDS Info Model 4) and can be easily customized to fit a wide range of scenarios due to the internal separation between the Execution Core Container and Data App. It integrates with many existing IDS services

---

[2] https://kubernetes.io/.

[3] https://min.io/.

[4] https://github.com/Engineering-Research-and-Development/true-connector.

(Identity Provider, Clearing House, Metadata Broker, etc.) and is fully configurable in terms of data formats and protocols (HTTP, HTTPS, Web Socket over HTTPS, IDSCPv2). It includes a Data Usage Control App to enforce usage policies and a basic Data App that can be customized to meet specific processing and integration needs.

### 5.5 Data Analytics of METRIQA in the Live Demo

The Big Data Analytics Platform[5] is a Data Science (DS) and Machine Learning (ML) solution built on advanced frameworks and open-source technologies for designing, deploying, running, and monitoring Big Data Analytics (BDA) workflows in both streaming and batch modes. Its cloud-native architecture allows for scalable computing and storage by leveraging Kubernetes' capabilities for resource management. The platform offers an extensible catalog of BDA services that covers every phase from data ingestion to data preparation, machine learning analysis, and data publishing. The platform supports data processing from various sources, integrates third-party BDA services, and provides a web-based interface to write BDA applications. Users can design workflows for both batch and streaming data scenarios by combining assets such as datasets and models with BDA services. The Big Data Analytics Platforms supports the assembling, deployment and execution of BDA applications through a graphical user interface. Each BDA application is made of one or multiple BDA services. Each BDA service is a containerized OCI-compliant microservice application. Our proposal, named on the Alida platform, is based upon severl open-source technologies:

- **Kubernetes:** as resource orchestrator;
- **Argo Workflows[6]:** an open source container-native workflow engine for orchestrating parallel jobs on Kubernetes;
- **Apache Kafka[7]:** an open-source distributed event streaming platform;
- **React Flow[8]:** a component on which the graphic designer is based in order to design pipelines of BDA services.

## 6 Discussions and Conclusions

The implementation of interoperable traceability in the agri-food supply chain, as demonstrated by the METRIQA platform, highlights several key advances and challenges in the field.

The integration of IoT, blockchain, and data spaces into a unified system represents a significant technological advancement. This integration ensures real-time data collection, secure data storage, and seamless data exchange, which are

---

[5] https://home.alidalab.it/.
[6] https://argo-workflows.readthedocs.io/en/latest/.
[7] https://kafka.apache.org/.
[8] https://reactflow.dev/.

critical to maintaining the integrity and transparency of the supply chain. In particular, the use of blockchain technology provides an immutable and transparent ledger that increases trust among stakeholders. The ability to track and verify data without a central authority reduces the risk of fraud and increases the reliability of information.

Achieving interoperability between disparate systems and platforms remains a significant challenge. The diversity of data formats, communication protocols, and technology standards requires robust solutions to ensure seamless data exchange. The development of governance systems to manage interoperability is critical. Strongly coordinated supply chains are more likely to achieve interoperability, but smaller companies may face challenges in adopting these systems without losing their autonomy.

The demonstrator effectively showcases the practical applications of interoperable traceability in several scenarios, including wine production, automated data collection, and cattle tracking. These use cases highlight the benefits of improved data integration, such as increased operational efficiency, improved product quality, and increased consumer confidence. Using machine learning to analyze sensor data and provide actionable insights further enhances the value of the traceability system. By transforming raw data into meaningful information, stakeholders can make informed decisions that improve the overall efficiency and sustainability of the supply chain. The combination of tamper-proof IoT devices and blockchain technology provides a robust solution for fraud prevention. By ensuring the integrity of the data collected and stored, the system can effectively prevent fraudulent activities such as mislabeling and counterfeiting. The ability to verify the authenticity of products using QR codes increases consumer confidence and trust in the supply chain.

The study shows that interoperable traceability is a viable and effective solution to improve transparency, efficiency, and trust in agrifood supply chains. The METRIQA platform, with its integration of IoT, blockchain, and data spaces, provides a comprehensive framework for achieving these goals. Key conclusions of the study include:

- **Enhanced Transparency and Efficiency:** Seamless integration of data from multiple sources ensures that all stakeholders have access to accurate and reliable information, improving the overall efficiency of the supply chain;
- **Improved Consumer Trust:** Providing detailed product information through QR codes and ensuring data integrity through blockchain technology improves consumer trust in the quality and authenticity of products;
- **Scalability and Flexibility:** The use of open source technologies and scalable architectures ensures that the system can be adapted to different scenarios and expanded as needed.

Further research is needed to address interoperability challenges and to develop more robust governance systems. In addition, exploring the integration of other emerging technologies, such as artificial intelligence and big data analytics, could further enhance the capabilities of the traceability system. The model, platform, and solutions extend to all transportation systems, fundamental stakeholders in all supply chains.

## References

1. Artificial intelligence in supply chain management: a systematic literature review. J. Bus. Res. **122**, 502–517 (2021). https://doi.org/10.1016/j.jbusres.2020.09.009
2. Agarwal, U., et al.: Blockchain technology for secure supply chain management: a comprehensive review. IEEE Access **10**, 85493–85517 (2022). https://doi.org/10.1109/ACCESS.2022.3194319
3. Bacco, M., et al.: Agricultural data space: the METRIQA platform and a case study in the CODECS project. In: Federated Conference on Computer Science and Information Systems (FedCSIS) (2024)
4. Brunori, G.: Agriculture and rural areas facing the "twin transition": principles for a sustainable rural digitalisation. Ital. Rev. Agric. Econ. **77**(3), 3–14 (2022). https://doi.org/10.36253/rea-13983
5. Chessa, S., Dimitri, G.M., Gori, M., Kocian, A.: WoA: an infrastructural, web-based approach to digital agriculture. In: Novais, P., et al. (eds.) ISAmI 2023. LNNS, vol. 770, pp. 113–122. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43461-7_12
6. Curry, E., Scerri, S., Tuikka, T.: Data Spaces: Design, Deployment and Future Directions (2022). https://doi.org/10.1007/9783030986360
7. FAO: Feasibility study for application of digital technologies for improved traceability and transparency along the agrifood value chains- Case studies in the Near East and North Africa Region. FAO (2023)
8. Franklin, M., Halevy, A., Maier, D.: From databases to dataspaces: a new abstraction for information management. ACM SIGMOD Rec. **34**(4), 27–33 (2005). https://doi.org/10.1145/1107499.1107502
9. Getz, W.M., Luisa Vissat, L., Salter, R.: Simulation and analysis of animal movement paths using numerus model builder (2020). https://doi.org/10.22360/SpringSim.2020.TMS.001
10. Hassoun, A., et al.: Food traceability 4.0 as part of the fourth industrial revolution: key enabling technologies. Crit. Rev. Food Sci. Nutr. **64**(3), 873–889 (2024). https://doi.org/10.1080/10408398.2022.2110033
11. Hassoun, A., et al.: Implementation of relevant fourth industrial revolution innovations across the supply chain of fruits and vegetables: a short update on traceability 4.0. Food Chem. **409**, 135303 (2023). https://doi.org/10.1016/j.foodchem.2022.135303
12. Hodapp, D., Hanelt, A.: Interoperability in the era of digital innovation: an information systems research agenda. J. Inf. Technol. **37**(4), 407–427 (2022). https://doi.org/10.1177/02683962211064304
13. International Data Space Association: Reference architecture model v.3.0 (2019). https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf. Accessed 11 Sept 2024
14. Marshall, B.M., Duthie, A.B.: abmAnimalMovement: an R package for simulating animal movement using an agent-based model. F1000Research (2022). https://doi.org/10.12688/f1000research.124810.1

15. Misra, N.N., Dixit, Y., Al-Mallahi, A., Bhullar, M.S., Upadhyay, R., Martynenko, A.: IoT, big data, and artificial intelligence in agriculture and food industry. IEEE Internet Things J. **9**(9), 6305–6324 (2022). https://doi.org/10.1109/JIOT.2020.2998584
16. Oladimeji, D., Gupta, K., Kose, N.A., Gundogan, K., Ge, L., Liang, F.: Smart transportation: an overview of technologies and applications. Sensors **23**(8), 3880 (2023). https://doi.org/10.3390/S23083880
17. Olsen, P., Borit, M.: How to define traceability. Trends Food Sci. Technol. **29**(2), 142–150 (2013). https://doi.org/10.1016/j.tifs.2012.10.003
18. Shoomal, A., Jahanbakht, M., Componation, P.J., Ozay, D.: Enhancing supply chain resilience and efficiency through internet of things integration: challenges and opportunities. Internet Things **27**, 101324 (2024). https://doi.org/10.1016/j.iot.2024.101324
19. Singh, A., Gutub, A., Nayyar, A., Khan, M.K.: Redefining food safety traceability system through blockchain: findings, challenges and open issues. Multimed. Tools Appl. **82**(14), 21243–21277 (2023). https://doi.org/10.1007/s11042-022-14006-4
20. Trienekens, J., Wognum, P., Beulens, A., Van Der Vorst, J.: Transparency in complex dynamic food supply chains. Adv. Eng. Inform. **26**(1), 55–65 (2012). https://doi.org/10.1016/j.aei.2011.07.007