Connecting Commercial Blockchain Platforms and European Customs:

an Interoperable and Self-Sovereign Data Sharing Architecture

Luis Alejandro Cabrera Mosca

TNO TU Delft

DELFT UNIVERSITY OF TECHNOLOGY

MASTER THESIS

Connecting Commercial Blockchain Platforms and European Customs: *an Interoperable and Self-Sovereign Data Sharing Architecture*

Submitted to obtain the degree of
MASTER OF SCIENCE
in Transport, Infrastructure & Logistics

by **L.A. CABRERA MOSCA**
student number 4439295

To be defended in public on September 29th, 2021

Graduation Committee

Chairman:	Prof. Dr. Y. (Yao-Hua) Tan	TU Delft	TPM Faculty	Section ICT
First Supervisor:	Dr. J. (Jolien) Ubacht	TU Delft	TPM Faculty	Section ICT
Second Supervisor:	Dr. B. (Bart) Wiegmans	TU Delft	CEG Faculty	Transport & Planning
External Supervisor:	Dr. B.D. (Boriana) Rukanova	TU Delft	TPM Faculty	Section ICT
External Supervisor:	Dr. Ir. W. (Wout) Hofman	TNO		

An electronic version of this thesis is available at https://repository.tudelft.nl/

TU Delft

Acknowledgements

This thesis represents the culmination of my degree of Master of Science in Transport, Infrastructure and Logistics at the Delft University of Technology. The research has been performed in collaboration with the Dutch Organisation for Applied Scientific Research (TNO) as part of the PROFILE project, an initiative funded by the European Union for the Horizon 2020 research and innovation programme to improve customs risk management.

My participation in this project has been truly an enriching experience. During the last six months, I have had the opportunity to explore the latest developments in information sharing within supply chain and logistics, as well as how European institutions constantly adapt their supervision strategies in view of fast-changing industry practices. From an academic perspective, the personal benefits of this project have exceeded my expectations, reminding me that intellectual reward is always hidden in the most unexpected places. I am in debt to those without whom this milestone would have not been possible.

First, I would like to thank TNO and my supervisors. Thank you Jok Tang for the warm welcome to the TNO family. Thank you Wout Hofman for helping me navigate complex issues and discern the music from the noise. Thank you Yao-Hua Tan, Bart Wiegmans and Boriana Rukanova for being a critical eye behind my progress. I am specially grateful to my main supervisor, Jolien Ubacht, whose patience and empathy have been as important as her guidance throughout the entirety of the project.

Second, I would like to thank those who provided unconditional support. Thank you to my friends for allowing me to gather so much joy and memories during these unforgettable years. Thank you Norma for being an anchor of hope despite the distance.

Lastly, I would like to thank my family for always believing in my endeavours and inspiring me from a young age to seek personal growth through curiosity towards the unknown. Your living examples are the best lessons of kindness and perseverance that I could ever have.

L.A. CABRERA MOSCA
Delft, September 2021

Summary

Increasingly specialised logistic services are triggering the disaggregation of supply chain functions and fostering the generation of information silos. This is perceived by European customs as a threat, since it affects the reliability of the import risk assessments they conduct. The organisation of cargo inspections in European ports is based on the results of these risks assessments. Therefore, their accuracy and reliability are essential to ensure secure borders and legitimate trade in the European Union (*EU*). Additionally, the number of customs agents are limited and excessive inspections can lead to an unsustainable reduction in cargo throughput at strategic transport hubs. Therefore, the number of inspections European customs can conduct is subject to manpower and commercial competitiveness constraints. In this context, European customs is exploring alternatives to collect better declaration data and improve the reliability of risks assessments.

Recently, commercial data sharing platforms based on blockchain technology (*BCT*) are allowing to expedite the verification of trade finance documents. The cross-organisational trust achieved in these platforms has driven the digitisation of the trade documents, from which import declaration data is extracted. European customs see combining data from multiple platforms an opportunity to improve supply chain visibility, turn import declarations more agile and risk assessments more effective. However, it remains unclear how to integrate declaration procedures in these new data ecosystems.

This thesis explores how to overcome two of the major barriers preventing European customs to interact easily with these platforms. Firstly, the lack of interoperability solutions to make platform architectures compatible with each other to share declaration data between peers across multiple platforms across supply chains. Secondly, the need to adapt available identity management solutions to the distributed nature of these platforms to promote trust between declarants and the rest of logistic service providers they interact with, also know as data sovereignty. This research gap involves the study of a ledger-based information sharing architecture that allows the generation of import declarations based on the data stored in multiple blockchain platforms. In order to achieve this, the following main research question is formulated:

What interoperable peer-to-peer data sharing architecture can be used by European customs administrations to gather declaration data from commercial blockchain platforms while preserving the data sovereignty of supply chain actors?

Since the goal is to design the architecture of an information system, design science research is used. It allows to capture the business and institutional needs of a practical problem-solving space, while keeping the design bounded by an academically relevant scope of work. The research is divided into five phases: 1) *Problem Explication*, 2) *Requirement Generation*, 3) *Design & Development*, 4) *Demonstration* and 5) *Evaluation*. First, by means of a literature review, document analysis and the consultation with industry experts, a broader understanding of the researched domain is gained. Second, requirements are obtained to narrow down the potential design directions and describe in detail the architecture functionalities required by the application context. Then, the actual architecture is designed: the requirements and knowledge collected during the two previous phases is converted into concrete technical components. The next phase is the demonstration of the data sharing architecture, which shows that the design can be applied successfully beyond the conceptual plane to solve a practical problem. Lastly, the goal of the design evaluation is to assess whether the proposed architecture complies with the specified requirements and the extent to which it can be considered a feasible solution.

The *Problem Explication* shows how European customs regulations have evolved to accept declaration data in the form of links to private information systems. Despite the evolution of European regulation and the advancement in information technologies, the transport industry has faced challenges for the digitization trade documents, such as the bill of lading (*B/L*). The legal certainty around the digitisation of these documents has been preventing transport processes from reaching its full efficiency potential. However, blockchain technology (*BCT*) has proved its ability to reduce the friction of information sharing by expediting the issuing and processing of *B/L*s while guaranteeing the security of cargo ownership chains. This is the reason commercial blockchain platforms have gained popularity in the shipping industry and show the potential to become the new industry standard for the management of declaration data. The data pipeline concept emerges from this trend, which envisions dynamic and secure inter-organisational information sharing by leveraging the security and verifiability characteristics of these platforms. Therefore, *BCT* offers data sharing incentives and promotes trust in the data provided by other entities. For this reasons, it is considered an enabler of supply chain transparency. The main challenge is equipping commercial blockchain platforms users with the technical capabilities to generate verifiable links to trusted data, which can then be used by customs administrations to perform their institutional duties.

Three design principles are chosen to guide the *Requirement Generation*. The first principle, logistic event visibility, is the essence of the practical value of the design for customs administrations. The two remaining design principles are two of the accompanying implementation challenges addressed in the research: stakeholder data sovereignty and architecture interoperability. Based on the technical characteristics of commercial blockchain platforms and relevant European legislation, a set of functional and non-functional requirements is generated. These requirements are used in the next phase to motivate the selection of architecture components.

The *Design & Development* results in a novel approach to migrate from import declarations based on duplication towards information sharing based on links to original and trusted data stored in blockchain platforms. Three layers are used: *Cross-chain Communication*, *Credential Management* and *Event Visibility*. The *Cross-chain Communication* layer uses an overlay network to enable cross-platform peer-to-peer interactions. A data transfer protocol is used to observe and share ledger states via trusted gateways. Hash time-lock contracts are used to propagate self-sovereign smart contract logic defined by the owners of logistic data. The *Credential Management* layer uses decentralised identifiers (*DID*s) as an alternative to the currently available identity certification solutions, allowing data owners to be in full control of the exposure of their digital identities and business information and avoid challenges related to key rotation and certificate revocation. Lastly, *Event Visibility* proposes a directed acyclic graph (*DAG*) ledger environment where to deploy decentralised applications. These applications are used to combine the ledger states of independent logistic blockchains, model the issue of trade documents throughout the whole cargo custody chain, and eventually detect anomalies in the framework agreements between logistic partners.

The *Demonstration* illustrates how the different architecture layers are used to create links between the *B/L*s issued between exporters, a freight forwarder and a carrier. The latter acts as declarant and creates an import declaration that bundles links to declaration data stored in other blockchain platforms. It is shown how the proposed architecture can be used to reduce the number of intermediaries that process the logistic data eventually included in import declarations, reducing incomplete and contradictory entries in the final import declarations processed by customs.

A naturalistic *ex ante* evaluation is been performed in the *Evaluation* to assess the compliance with the design requirements and the applicability of the data sharing architecture. The link between functional and non-functional requirements and the components is tested. The relevance and suitability of the designed is proved against the industry knowledge of an expert, who validates its implementation potential in a leading commercial blockchain platform. Governance and implementation costs are identified as main design weaknesses, and approaches to tackle both are proposed. First, a framework to link governance and technical requirements is discussed. Then, an overview of expected implementation costs is covered, as well as the identification of potential transaction benefits of a ledger-based design approach for supply chain applications.

Following the completion of the research activities, future research areas are identified. The research has focused on trade document, leaving aside a growing number of data sources, such as sensor-based data collected at transport terminals. Future research should explore how to complement the architecture with these data sources in order to integrate document-based data sharing with logistic data gathered on site. The research has considered the format and content differences between *B/L*s and import declarations negligible to focus on the high-level design of architecture components. This leaves room for further research on the requirements for the cross-reference of documents between platforms to ensure end-to-end semantic compatibility during cross-validations performed by customs. Lastly, additional research should evaluate the long-term performance consequences of the selected components, such as the feasibility of an additional consensus layer an its requirements in terms of speed and scalability. This is important in order to confirm that the transaction throughput required by European customs is in line with the architecture´s ability to maintain consensus on a growing number of cross-platform references.

Contents

List of Figures

List of Tables

List of Abbreviations

aBFT	**a**synchronous **B**yzantine **F**ault-**T**olerant
AEO	**A**uthorised **E**conomic **O**perator
B2B	**B**usiness-to-**B**usiness
B2G	**B**usiness-to-**G**overnment
B/L	**B**ill of **L**ading
BaaS	**B**lockchain-**a**s-**a**-**S**ervice
BTC	**B**loc**h**ain **T**echnology
CCC	**C**ross-**c**hain **C**ommunication
COFE	**C**ustoms **O**ffice of **E**ntry
COU	**C**ustoms **O**ffice of **U**nloading
D/O	**D**elivery **O**rder
DAG	**D**irected **A**cyclic **G**raph
DeFi	**De**centralised **Fi**nance
DDs	**D**esign **D**ecisions**s**
DID	**D**ecentralised **I**dentifier
DLT	**D**istributed **L**edger **T**echnology
DPKI	**D**istributed **P**ublic **K**ey **I**nfrastructure
DPs	**D**esign **P**rinciples**s**
eB/L	**e**lectronic **B**ill of **L**ading
ENS	**En**try **S**ummary Declaration
ETA	**E**stimated **T**ime of **A**rrival
EU	**E**uropean **U**nion
GC	**G**roupage **C**enter
GDP	**G**ross **D**omestic **P**roduct
GDP	**G**eneral **D**ata **P**rotection **R**egulation
HB/L	**H**ouse **B**ill of **L**ading
HTLC	**H**ash **T**ime-**l**ocks **C**ontract
ICT	**I**nformation and **C**ommunication **T**echnologies
IDP	**Id**entity **P**rovider
IoT	**I**nternet of **T**hings
L/O	**L**etter of **C**redit
LOD	**L**inked **O**pen **D**ata
MB/L	**M**aster **B**ill of **L**ading
MMR	**M**erkle **M**ountain **R**ange
NVOCC	**N**on-**V**essel **O**perating **C**ommon **C**arrier
P2P	**P**eer-to-**P**eer
PBFT	**P**ractical **B**yzantine **F**ault-**T**olerant
PKI	**P**ublic **K**ey **I**nfrastructure
PLA	**Pl**ace of **A**cceptance
PLD	**Pl**ace of **D**elivery
PoT	**P**ort **o**f **T**ranshipment
POD	**P**ort **o**f **D**elivery
POL	**P**ort **o**f **L**oading
RA	**R**isk **A**ssessment
RDF	**R**esource **D**escription **F**ramework
RSA	**R**ivest–**S**hamir–**A**dleman

SB/L	**S**traight **B**ill of **L**ading
SC	**S**tripping **C**enter
SSI	**S**elf-**S**overeign **I**dentity
TDAG	**T**ransaction **G**raph
TNO	Nederlandse Organisatie voor **T**oegepast-**N**atuurwetenschappelijk **O**nderzoek
TTL	**T**rusted **T**rade**l**anes
URI	**U**niform **R**esource **I**dentifier
URL	**U**niform **R**esource **L**ocator
URN	**U**niform **R**esource **N**ame
W3C	**W**orld **W**ide **W**eb **C**onsortium

Chapter 1

Motivation & Approach

This chapter is an introduction to the research. Background information about to the research context and the research design can be found in section 1.1. The data management and supply chain visibility challenges faced by European customs administrations that motivate this study are covered in section 1.2. The addressed research gap is covered in section 1.3, and the research objectives in section 1.4. Finally, the selection of the research approach and the detailed research design can be found in section 1.5.

1.1 Introduction

In view of the global interindustrial trend towards digitization and the rapid rise of the smart industry technologies - such as artificial intelligence, blockchain or the internet of things - the logistics sector is expected to undergo an imminent transformation during the upcoming decade [86]. However, despite the benefits this new technological landscape entails, its real implementation potential remains surrounded by uncertainty [122]. In this context, blockchain technology (*BCT*), and distributed ledger technology (*DLT*) in general, are regarded as enablers of more transparent logistic processes with the ability to abruptly redefine the exchange of data across supply chains.

In a nutshell, *BCT* is to transactions what the internet was to information. Being a trustless system based on transparency and visibility, it allows entities to share information without the need to assess their degree of trust towards other participants [107]. It's main advantage is the set of modern cryptographic mechanisms through which transaction records become immutable, meaning that they can not be altered once stored in the blockchain [134]. In the recent past, *BCT* has emerged as a disruptive technology with innumerable applications in the transport and shipping industry. Improved cost effectiveness by simplifying the tracking of items and transactions [65], increasing shipping flexibility [101] and a considerable reduction of supply chain risks [36] have all been shown achievable benefits of integrating *DLT*, *BCT* in particular.

The source of the interest in using *BCT* to tackle the latest challenges in large-scale logistic data dissemination is threefold. First, a dire need to cope with a fast increase in supply chain complexity, both at operational and organizational level [32]. Second, the leading businesses' fear to fail at staying at the vanguard of their industries [169], fueled by the aforementioned trend towards industrial digitization and *BCT*'s potential to reshape the foundations of existing business models [65]. And lastly, the success examples of early adopters. The leverage of *BCT* for non-financial applications is a reality, being widely accepted to have reached sufficient level of maturity to produce tangible results in real-world problems [169]. From supply chain transparency for ethical sourcing in the fashion industry to counterfeit prevention and product authentication in pharmaceutical distribution [36, 72], *BCT* is being slowly adopted in varied logistics applications.

1.2 Challenges faced by European Customs

While the transportation industry adapts to the digital era, *BCT* can play a main role in overcoming some of the latest supply chain management challenges [86]. In order to fully grasp the motivation of customs administrations to tackle these using *DLT*, the remainder of the section introduces current supply chain management weaknesses, including strategic considerations on European trade, the rise of blockchain platforms in the private sector and technical barriers preventing the development of some blockchain applications.

European Union Trade

Traditionally, international trade has had macroeconomic, political and diplomatic connotations until the apparition of the *Digital Economy*. Behind the recently coined term stands the new wave of technical innovations bringing logistic data processing and strategic commerce decision-making closer than ever [113]. New business models have emerged from the need to gather, store and exchange data [166], being now crucial in the development of global value chains [170]. As a result, international commerce has been radically redefined during the last twenty years, converting effective cooperation between institutions and enterprises the cornerstone for the success of many industries. Trade represents 35% of the European Union's *GDP*, who accounts for almost 17% of global trade [76]. Public administrations are thus obliged to seek adequate trade policies to align their interests with new technological paradigms, as well as fostering the implementation of beneficial industry practices.

This growing interaction between European companies and the rest of the world is acknowledged by the *European Commission* as evidence of how products and services are increasingly traded across borders. In this new international commerce scene the exchange of information is as valuable as the exchange of goods, and policies must enable this new reality [170]. Moreover, the particularly complex institutional ecosystem of the *European Union* (*EU*) imposes additional constraints to the translation of current supply chain management practices into executable road-maps for the digitization of activities related to trade.

A good example is the diversity of port governance models used by member states, meaning that European policy-makers must define and promote a common modernization strategy while respecting the members' right to establish their own strategic commerce and transport development agendas [172].

Supply Chain Visibility

The implications of this deep global trade transformation on the design and management of supply chains are vast. Institutions and industry leaders are fully aware that transport systems and infrastructure have become more dependent on innovation and technology than ever [172], but there is still no consensus on the most convenient strategy to navigate this transformation. This lack of long-term vision has not prevented supply chain visibility from becoming a very relevant research area for the re-interpretation of supply chain management in the *Industry 4.0* era [149].

Supply chain visibility can be seen as an enabler of supply chain interaction in terms of cooperation (visibility of essential data and limited alliances), coordination (joint visibility procedures) and collaboration (shared vision and consolidated trust) [153]. However, the realization of these interactions has been limited to a large extent by insufficient infrastructure to effectively access data. As a result, the recent advance of information and communication

technologies (*ICT*) capable of bridging these data sharing voids has lead current research to focus on the opportunities to elevate the standards of supply chain data readiness, usability and shareability [28, 103]. This can be interpreted as the search of novel data dissemination mechanisms that may allow the retrieval and storage of high-quality data throughout more complex supply chains.

Following this perceived urgency to develop better ways to leverage supply chain data, is the strong link between supply chain visibility and the ability to detect and respond to risks [126, 180]. This is applicable to the private sector regarding the uncertainty around operational performance, and ultimately, business objectives. But the public sector, customs administrations in particular, must also prioritize the visibility of logistic processes across their borders to protect their national interests.

As the nature of global trade and commerce shift, so does the structure of the logistic processes driving border risks. For instance, the tendency to disaggregate supply chain functions leads to an increasing number of actors, which in turn produces decentralized knowledge within supply chains [180]. Global cargo mobility currently implies complex cargo custody chains involving actors that provide increasingly specialised services, such as freight forwarding, warehousing, distribution and other commonly outsourced activities.

On the eyes of customs administrations, the result is information fragmented between the providers of these services and potentially hidden by an opaque network of commercial agreements. Therefore, the ability to assess the impact of imports on national interests commences to rely on an unprecedented level of collaboration between customs administrations and a growing number of enterprises. This is perceived by European institutions as a threat, which raises the flag on their obligation to maintain end-to-end visibility on the economic activities carried out across supply chains entering the *EU*.

Commercial Blockchain Platform Ecosystem

Recently, the logistics sector has reduced collaboration friction and enhanced trust with blockchain-based data sharing platforms. These platforms offer logistic service providers with secure information exchange services that help optimising contractual information flows and decreasing risks by improving the reliability of forecasts. An example is *TradeLens*, a joint venture between *Maersk* and *IBM* aimed at improving supply chain visibility, and in doing so, increasing the efficiency in containerised shipping [82].

Since platforms coordinate information flows between all kinds of actors between origin and destination, customs administrations identify the opportunity to improve their supply chain visibility by combining data from multiple platforms for customs declaration purposes. This is possible because, as shown in Figure 1.1, logistic information from each segment of a supply chain is processed by a different platform. Instead of relying on the last cargo custodian to forward the information accumulated downstream a supply chain, customs could benefit from retrieving logistic data earlier at each stage of a supply chain. This abundance of information helps moving from declaration data based on duplication towards information sharing based on links to the original and trusted data stored in these platforms. However, despite the benefits these platforms have brought to the private sector, it remains unclear how customs administrations can integrate current customs declaration procedures with these new information ecosystems. Overall, *BCT* is a very promising approach to achieve global supply chain visibility standards, but it includes its own set of intrinsic barriers that ought to be surpassed.

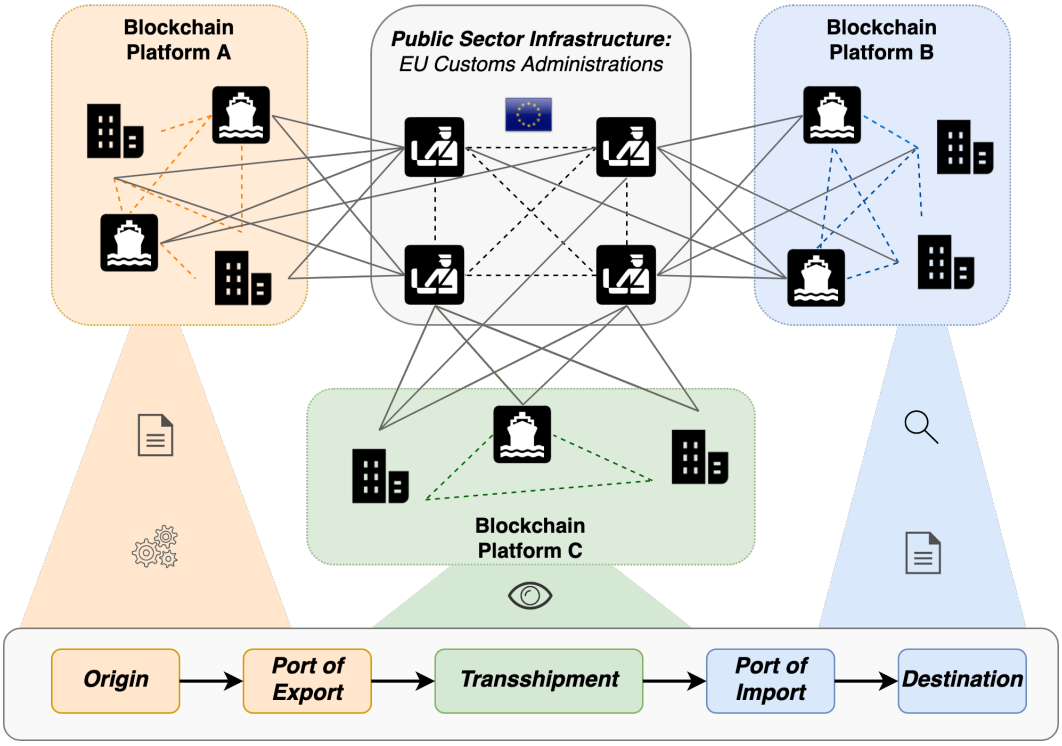


FIGURE 1.1: Blockchain platform ecosystem.

Blockchain Interoperability

In general, interoperability can be defined as the ability of two information systems to communicate, understand and collaborate with each other despite interface and execution platform differences [131]. In the research context, interoperability refers to the communication standards between different blockchain platforms that allow them to agree on the interpretation of information [98]. Interoperability is crucial for long-term industry transformations, because the widespread implementation of *BCT* solutions without sufficient interoperability can prevent blockchain applications from solving multi-actor collaboration issues at a socio-technical scale [164].

Interoperability is thus a main challenge that the mass adoption of *BCT* in supply chain management is facing, specially when new applications are being proposed and tested at a fast pace [48]. Given the fast development of a market for this type of platforms impulsed by private initiatives, law enforcement agencies and customs administrations within the European Union face the challenge of preparing adequate data sharing mechanisms to interact with these providers of data exchange services.

Although blockchain platforms aggregate transactions from multiple information flows, the ability to certify the validity of transactions and the interpretation of data produced in the platform is reserved to its users. A growing number of blockchain platforms would result on siloed nests of supply chain data powered by different blockchain protocols. Therefore, on top of the challenge of adapting current customs supervision procedures to blockchain environments, links between information produced in different platforms should be created.

In case companies and institutions start developing blockchain applications independently, it is impossible to ensure their interoperability following the currently available blockchain frameworks [102], which poses a great barrier for those entities whose activities involve their

participation in multiple blockchain systems. This problem is accentuated for scenarios in which data from multiple blockchain-based sources needs to be combined, such as the risk assessments performed by European customs.

Data Sovereignty

The aforementioned platforms are the result of private initiatives where enterprises have reached consensus on the architecture that best fits their data governance requirements. The data produced in these platforms contains sensitive commercial information, so their security benefits are one of the main motivations for private enterprises to voluntarily join these projects. Each platform guarantees confidentiality standards to achieve trust between participants, which directly hinders the incentives to establish cross-platform data links.

Before European customs can interact with these platforms, preserving the confidentiality of the information processed therein is another challenge to overcome. One of the motivations for enterprises to join these platforms is an increased control over their data. Therefore, even if customs is able to provide the best security measures, supply chain actors would still be reluctant to take the risk of centralising the control of very sensitive commercial information about their partners´ commercial activities. The challenge consists on proposing a data gathering strategy that defines data dissemination rules depending on the commercial relationship between actors [78], while making data access less complex and positively contribute to the institutional duties of European customs.

1.3 Research Gap

The interaction between distributed ledgers is very challenging because virtually all ledger protocols can be considered siloed and research on solutions to increase this interaction is very scare [98]. There is insufficient academic effort on the standardization of communication protocols that allow two different ledgers to share their internal states and coordinate application logic [183]. A notably increasing demand for additional work is addressed by industries and institutions closely linked to commerce, such as healthcare [4], energy [48], agriculture [145] and European customs [109, 128].

This research gap is even more pronounced for the presented *EU* customs challenges, since there is no real evaluation of data sharing architectures to exchange logistic data between multiple private blockchain platforms and European customs. Therefore, research on the components needed to aggregate distributed applications for the generation of *ENS* from verifiable links to the internal ledger state proofs of multiple blockchain platforms is not available. Moreover, customs may require to interact with a specific platform user, who should be able to participate in data exchanges without compromising the confidentiality of the information involving other platform users. This is why the peer-to-peer connection model shown Figure 1.1, or meshed ledger design, requires further research.

Similarly, there is limited research available on the implementation of self-sovereign identities in the supply chain domain: providing enterprises incentives to share data by equipping them with data governance privileges to control the exposure of their information. This is needed because the digitisation of supply chain management implies complementing data exchange architectures with decentralised identity management. The need for interoperable decentralised identity management is also motivated by the trend towards environments that combine public and private data access policies (hybrid blockchains) [48, 107].

Certifying and protecting identities is becoming an integral part of information sharing [43, 59], but architectures able to apply self-sovereign identity management in global supply chains are absent in literature. Methodologies for the translation of logistic processes into data sharing rules within a meshed ledger while accounting for data sovereignty are not available. This approach could however enable customs to gain the trust of enterprises and allow third parties to implement verifiable links towards information stored in blockchain platforms.

1.4 Research Objectives

Based on the challenges faced by European customs and the knowledge gaps presented in the previous section, a main research question is used to build the research activities. Also, a discussion on the expected scientific contribution and the societal relevance of the research is presented in the following sections.

1.4.1 Main Research Question

The main purpose of the research is to tackle a specific problem of logistic data dissemination between the private and public sector, which will be achieved by answering the following main research question:

Main Research Question	*MRQ*
What interoperable peer-to-peer data sharing architecture can be used by European customs administrations to gather declaration data from commercial blockchain platforms while preserving the data sovereignty of supply chain actors?	

In its broadest sense, an information system architecture can be defined as the *"fundamental concepts or properties of an information system in its environment, as embodied in its elements and relationships, and in the principles of its design and evolution"* (pp. 2) [85]. This definition interprets a valid analytical description of an architecture as the aggregated stakeholder perceptions of the physical and logical relationships between the components of an information system that provide value to their needs [159].

However, the divergent stakeholder needs of complex socio-technical systems, such as the described blockchain-platform ecosystem [49], make it difficult to describe every perception in an objective, mutually exclusive, collectively exhaustive way. Instead, the term architecture is used in this research in the context of an architectural model: *"an illustration, created using available standards, in which the primary concern is to represent the architecture of an information system from a specific perspective and for a specific purpose"* (pp. 27) [159].

The main research deliverable is a functional description of a peer-to-peer data sharing architecture from the customs administrations perspective: supply chain actors taking part in commercial blockchain platforms grant customs administrations access to logistic data. The ultimate purpose of the architecture is to allow European customs to combine information from multiple commercial blockchain platforms to be used for customs declaration purposes, while preserving the data sovereignty of the users involved.

1.4.2 Scientific Contribution

This research will contribute positively to the scientific community in different of ways. First, providing an additional use case of *DLT* in which the decentralization trilemma is overcome (pairing decentralisation, scalability and security), by prioritising decentralisation and security requirements in a design applicable to solve a practical problem. Therefore, the design methodology used can be applied in domains outside supply chain management and logistics, due to a better understanding of the interaction between technical ledger specifications and the governing rules of their application environments [68].

The second contribution is linked to the rapid rise of commercial blockchain platforms. A detailed data gathering mechanism that combines data from multiple ledgers under varying data sovereignty constraints will open opportunities for improving interoperability between blockchain protocols. This knowledge addition represents one of the largest obstacles for the widespread use of distributed ledgers [102] and can provide future insights on the adoption of *DLT* in trust-driven and heavily regulated scenarios.

Furthermore, the research sheds light on the synchronisation of information sharing between the public and private sectors. Finding technical solutions to circumnavigate regulation and industry practices in order to align the interests of private enterprises and public institutions is a relevant an challenging researched task [49]. In this context, the research presents an approach to combine *DLT*-based data sharing components to increase the compatibility of information flows between private supply chain actors and the associated interactions with European law enforcement agencies.

1.4.3 Societal Relevance

The research is conducted in collaboration with the Dutch Organisation for Applied Scientific Research (*TNO*) for PROFILE, a project funded by the European Commission as part of the Horizon 2020 program. The goal of this initiative is to help European customs identify high-risk activities earlier and with higher accuracy. This is expected to redefine how customs leverage supply chain data to protect their national interests by upgrading the data gathering capabilities of European agencies.

This does not only enhance the reliability of customs risk assessments, but also the prevention of tax fraud and the maintenance of strategic commerce relations, which affects society as a whole [148, 114]. It also benefits the logistics sector (*e.g.*, exporters, importers or shipping lines) by reducing the bureaucratic friction of customs declarations. and increasing the agility of transport terminals. Moreover, the successful development and implementation of the proposed architecture would showcase the potential of *DLT* to provide public institutions with better tools to engineer and enforce policies that promote public values and societal needs effectively [118].

As covered in section 1.2, ledger interoperability is a key bullet point in the global blockchain research agenda. Achieving compatibility between ledgers will make *DLT* more accessible to the general public and diminish the risk of knowledge monopolies. A reduced number of industry experts controlling the integration of *DLT* in matters of public interest, such as transport [59, 166], energy [48] or agriculture [145], should be avoided. The research represents a small step forward in the accomplishment of this objective by establishing a precedent on the interoperable integration of institutional supervision in large-scale private blockchain environments.

1.5 Research Design

The research methodology should be aligned with the addressed knowledge gap , as well as the scientific and societal contributions expected from the research. This section covers the available research approaches and the selection of the most appropriate methodology. Lastly, the secondary research questions and their link to the research phases are presented.

1.5.1 Research Approach

Since *BCT* is a developing research field and its range of applications wide, there is a lack of research frameworks tailored to the characteristics of all application domains [44]. The research combines the understanding of logistic processes, the data acquisition procedures of European customs, and the implementation of state-of-the-art *ICT* to preserve data privacy requirements across supply chain actors. The heterogeneity of approach angles makes it interesting to compare the alternatives that best fit each field included in the problem.

The first option is to use the logistical background of the problem as reference. There has never been consensus on the preferred research practices in this field, which vary from very detailed modelling techniques for simulation studies aimed at optimizing processes, to surveys and other qualitative methods related to organizational sciences aimed at covering the strong managerial implications of the field [67]. However, although one of the potential long-term benefits indirectly associated with the research includes the improvement of process performance - port terminal efficiency, for instance - the source problem does not reside in modelling in detail the ecosystem of logistic processes in which the solution will be implemented. For this reason, this is thought to be an invalid research approach.

Another option is to consider the problem to belong to the information systems research, which is the study of *ICT*, software and data systems to carry out specific tasks and interact with a number of entities in varied organizational and social contexts [19]. Among the possible subcategories found within information system research, the *work system* view by Alter [7] represents the essence of the problem at hand: *"an information system whose process and activities are devoted to processing information, that is, capturing, transmitting, storing, retrieving, manipulating, and displaying information"* (pp. 451).

This view focuses on how information artefacts, in which a number of actors are engaged, can be supported by innovative *ICT*, which describes the core of the problem treated in this research. Under this view, there are two main relevant information system research approaches worth discussing: action research and design science research. Action research entails focusing on problems with complex and direct implications in socio-technical systems and the human context, rather than the design of information system artefacts and their detailed technical characteristics [84].

On the contrary, design science research is suitable for the assessment of technically driven problems, and can be considered as the dominant approach in the discipline of information systems [40]. It is considered a valuable tool for bridging the gap between new relevant application environments and unexplored areas of the field's knowledge base using rigorous design practices [140]. Furthermore, literature shows that the design science research methodology has been systematically adopted in a considerable number of blockchain studies, having proved to be particularly useful in revealing adoption drivers and hindrances of novel blockchain applications [136]. This arguments are considered sufficient to justify the use of the design science research methodology in the research.

1.5.2 Research Framework

The chosen outcome based research approach can be described from two perspectives: a framework for the research methodology itself and a framework for the scientific and practical nature of the problem being addressed. This section elaborates on these two frameworks and their application in the proposed research topic. The research context framework shown in Figure 1.2 captures the essence of the design science methodology. It represents the interaction between the application environment of the researched artefact and the scientific knowledge base. The design science approach allows to capture the business and institutional needs of a practical problem-solving space, while keeping the design bounded by an academically relevant scope of work.

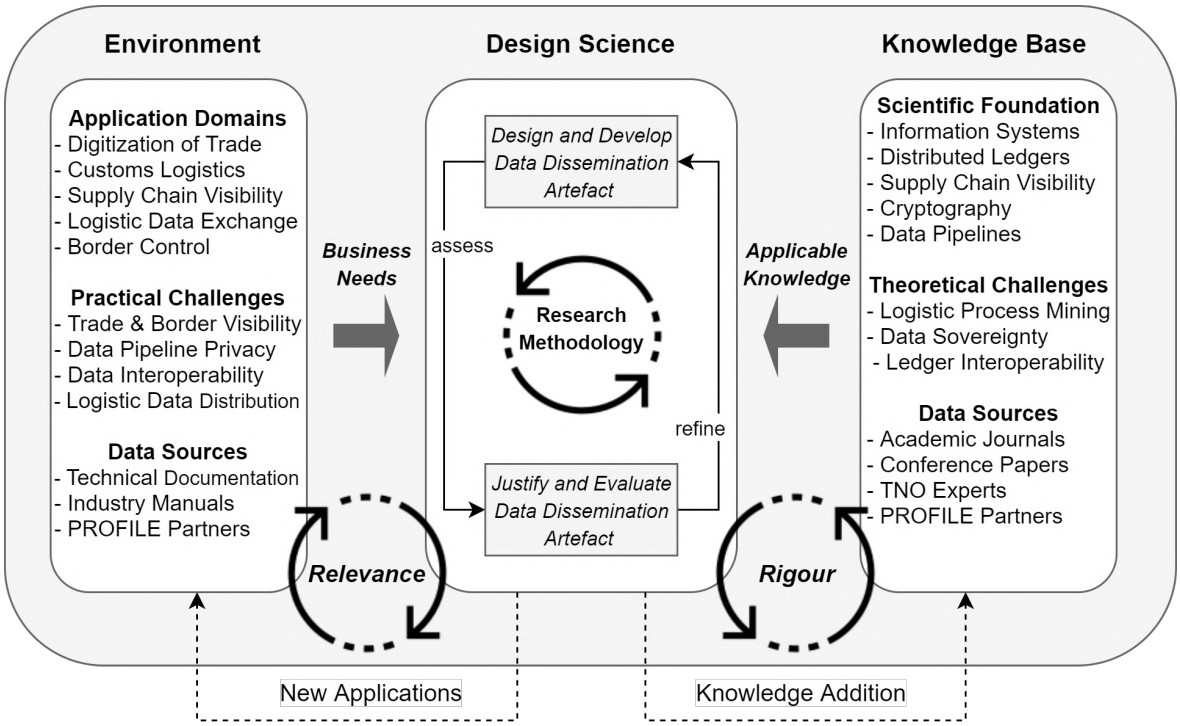


FIGURE 1.2: Research context framework, adapted [173].

The environment contains the application domains of the envisioned solution from different levels of detail, ranging from the digitization of trade to the technical nuances of ledger-based data access and distribution for border control. The functional requirements of the solution originate from this environment layer and is used to monitor the suitability for its assigned application. Similarly, the non-functional requirements of the solution originate from the knowledge base, which includes the scientific foundations of the academic disciplines covered in the research, such as supply chain visibility or data pipeline models using the latest *ICT*. This means that the combination of available scientific knowledge and the experience of TNO, TU Delft industry experts will be used iteratively. The next step is introducing the research methodology shown in Figure 1.3.

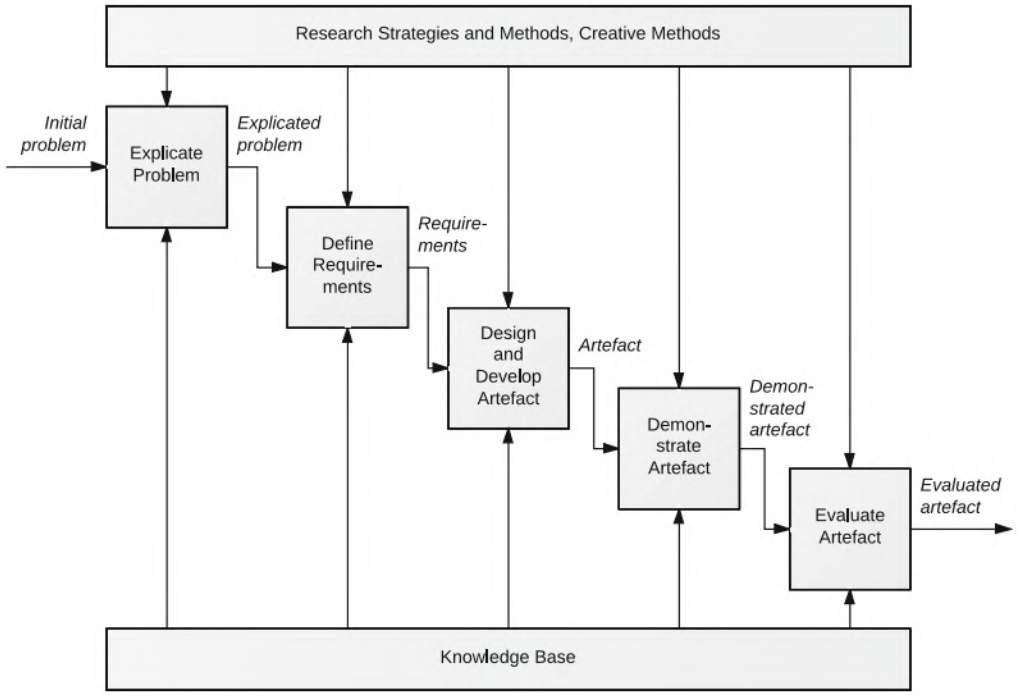


FIGURE 1.3: Design science research methodology framework [90].

1.5.3 Research Phases & Questions

Based on the book *"An Introduction to Design Science"* [90], this subsection introduces the work performed in each phase of the design science research framework and their associated research questions.

Problem Explication

In this phase, the practical problem must be broken down into narrower sub-problems. By means of extensive literature review, document analysis and the consultation with TNO experts, a deeper analysis of the problem at hand will be gained. This will be done by answering the following research question:

RQ1: ***What is the relationship between supply chain visibility, import declarations and the risk assessments performed by customs administrations?***

- *How is data shared between supply chain stakeholders?*
- *How is declaration data shared between supply chain stakeholders and European customs?*
- *What is the role of blockchain technology in supply chain visibility?*

By assessing these sub-questions, a deeper understanding of the current data exchange procedures will be gained. Also, by answering the second sub-question, a clearer definition of the current data confidentiality constraints will be achieved. The last sub-question will show the current data exchange used, as well as examples of interaction between customs and enterprises using *BCT*. During this phase, the advantages and disadvantages of using meshed ledger networks in the research context will be explored from a technical perspective, and the suitability of the available ledger technologies supporting meshed configurations will be further assessed.

Requirement Definition

The design directions are narrowed down in this phase to arrive at a sufficiently constrained design space. The following research questions are aimed to obtain a detailed description of the application context and design functionalities of the data sharing architecture. *RQ2* refers to the technical specification in terms of data sovereignty:

RQ2: ***What are the design requirements to preserve the data sovereignty of supply chain actors when creating links to data stored in multiple ledgers?***

The following research question refers to the interoperability requirements. Given that customs administrations needs to interact with platforms that use multiple blockchain technologies, the must be able to share partial views of the internal state of their platforms:

RQ3: ***What are the design requirements to allow multiple blockchain platforms to share interoperable links to their ledger states?***

Based on these requirements, the suitable architecture components and the logical relationships between them can be further developed in the next research phase.

Design & Development

During this phase, the actual artefact is created. The requirements and knowledge collected during the previous phases will be converted into concrete architecture components. The outcome of this phase is the main deliverable of the research, and is expected to consume the most time. An extensive literature review on distributed information systems and the analysis of the documentation of currently available solution will be used to achieve this, and answer the following research question:

RQ4: ***What architecture components can used by customs administrations to gather declaration data stored in multiple commercial blockchains?***

Demonstration

These next phase is the demonstration of the data sharing architecture. It intends to show that the design can be applied successfully beyond the conceptual plane to solve a practical problem. This is addressed by means of an illustrative real-life case [90] in order to answer the following research question:

RQ5: ***How would the current import declaration procedure be implemented using the peer-to-peer data sharing architecture?***

Evaluation

The evaluation of the design is the last research phase. Its goal is to assess whether the proposed architecture complies with the requirements specified in the *Define Requirements* and the extent to which is a feasible solution in practice. First, the design fit will be evaluated by comparing the design decisions documented in the *Design & Development* phase with the design requirements identified in the *Requirement Generation* phase. Then, an expert will be interviewed to assess the relevance and feasibility of using the data sharing architecture for the proposed application. The outcome of these two activities will be used to answer the following research question:

RQ6: ***Does the data sharing architecture comply with the requirements to a sufficient extent to be considered a feasible solution that contributes to the application domain?***

The research is concluded with a rigorous documentation of the work performed and a reflection on the limitations of the final deliverable. Also, based on the key findings, recommendations for future research opportunities to be carried out in next iterations of the design will be also discussed.

1.5.4 Data Collection

Regarding the initial motivation of the research, PROFILE stakeholders provided a description of the practical needs in terms of the desired solution functionalities. TNO and *TU Delft* experts provided recommendations on the preferred scientific models and guided the technical design phase. It is the researcher's duty to critically assess these recommendations by comparing them to existing literature and prediction on industry practices.

The research activities involves a good understanding of the current European customs procedures, customs regulations and underlying logistic activities performed by private enterprises. Regulation decisions and European strategic reports are used to complete the academic knowledge found in journals. The latter provide insights on the link between design approaches and industry practices. Both will be used in order to answer *RQ1*, *RQ2* and *RQ3*. Another source is the documentation of ledger technologies used by commercial blockchain platforms. Open source documentation and white papers are available to understand the components of existing architectures linked to logistic data management. These will be used mainly to discover design directions and answer *RQ4*. Document analysis will play a fundamental role in bridging the gap between practical relevance and scientific innovation. Therefore, a combination of literature reviews and document analysis is used throughout the entire research. An overview of the research methods and the data collection strategies is presented in Table 1.1.

TABLE 1.1: Overview of research methods and data sources.

Research Activity	Research Method	Data Sources
Explicate Problem (***RQ1***)	Literature Review and Document Analysis	Academic Journals & European Regulation
Requirement Definition (***RQ2 & RQ3***)	Literature Review, Expert Surveys and Document Analysis	Academic Journals, Regulations, Technical Documentation, Expert Network
Design & Development Artefact (***RQ4***)	Literature Review and Document Analysis	Academic Journals, Expert Network and Technical Documentation
Demonstration (***RQ5***)	Use Case	-
Evaluation (***RQ6***)	Requirement Analysis and Expert Interview	Expert Network and European Regulation

1.5.5 Research Flow Diagram

A detailed research diagram with the research activities performed and the outcome of the research questions is presented in Figure 1.4. The diagram also includes the data collection and research method used in each phase.

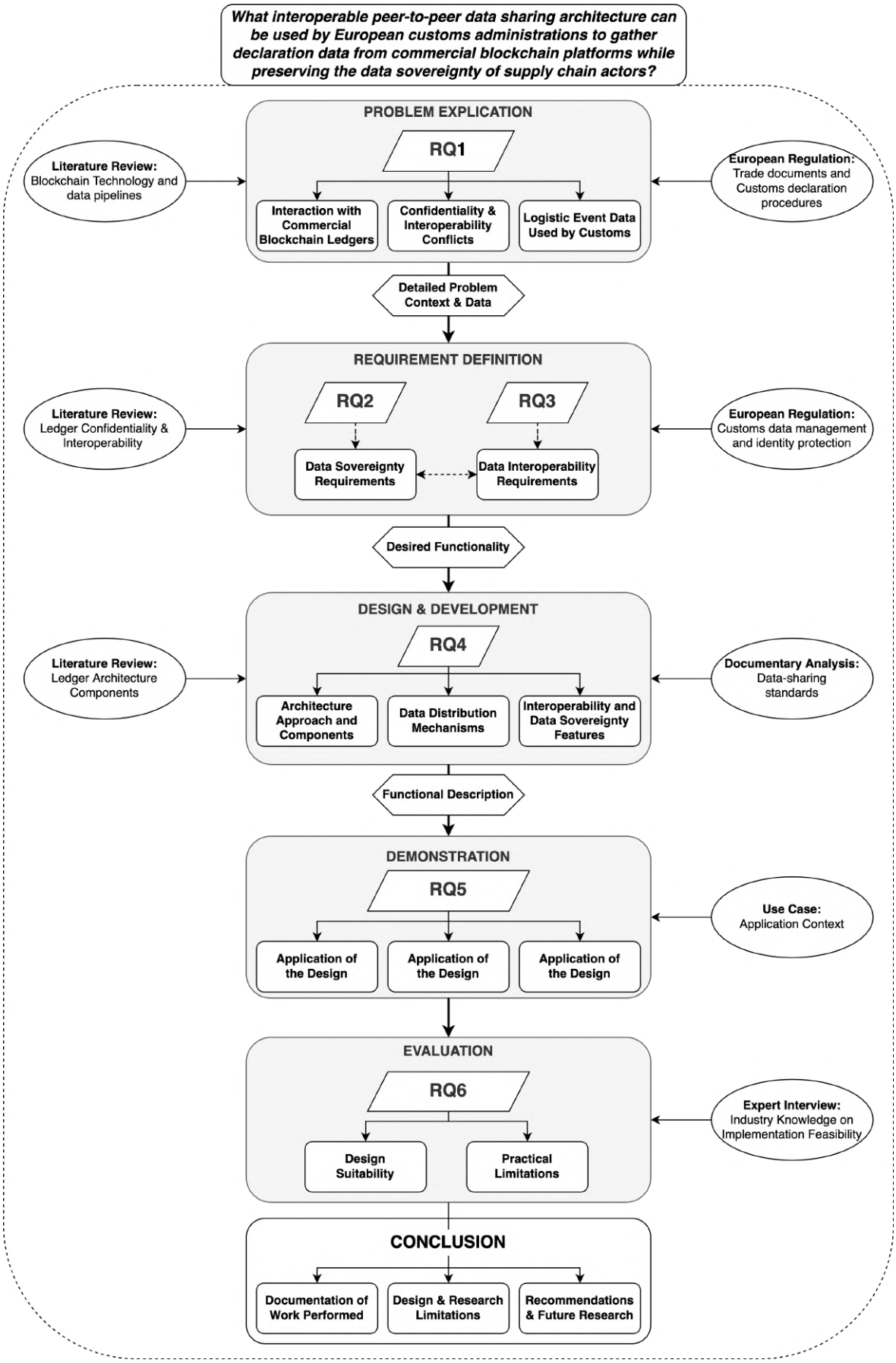


FIGURE 1.4: Research flow diagram.

Chapter 2

Problem Explication

This chapter presents a detailed problem statement to answer *RQ1: What is the relationship between supply chain visibility, import declarations and the risk assessments performed by customs administrations?* It will help understanding the opportunity to adapt declaration procedures to logistic data flows in the new blockchain platform ecosystem. The supply chains entering the *EU* are covered in section 2.1. The commercial relationship between supply chain stakeholders is covered in section 2.2. The flow of contractual information is covered in section 2.3. The customs declaration process and the confidentiality issues that emerge from different data flows are covered in section 2.4. An introduction to the data pipeline concept and its relevance in public-private data sharing is covered in section 2.5. Finally, an introduction to *BCT*, its application in supply chains and its role in the integration of supply chain management and the latest regulatory supervision strategies for trade is covered in section 2.6.

2.1 Structure of Supply Chains Entering the *EU*

The global trade scene through which goods are imported into the *EU* includes innumerable combinations of transport modes, shipping routes and cargo transfer movements around and across European and international customs. The physical flow of maritime cargo can be generalised as a discrete event sequence that all goods must go through before entering European territory. This process occurs in the logistic domain depicted in Figure 2.1.

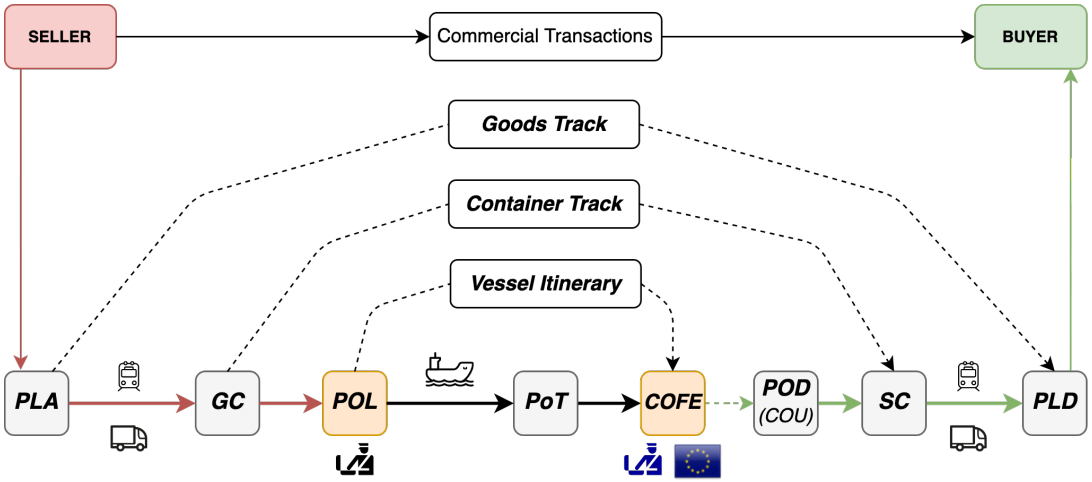


FIGURE 2.1: Logistic domain for *EU* imports, adapted [60, 128].

The seller of goods starts the shipping process from a Place of Acceptance (*PLA*) towards a Place of Delivery (*PLD*), where a buyer expects the transported goods. However, the commercial transactions between these parties are not linked to the production of logistic event data, as other intermediaries are the ones directly involved in the transportation of the goods and the subsequent import, export and carriage responsibilities. This section focuses on the latter, which are presented in more detail in section 2.2.

The *PLA-PLD* segment represents the *goods track*, where individual cargo, such as parcels, can be monitored. The goods are then transported to a grouping center (*GC*), also known as stuffing center, where they are rearranged or repacked and assigned to a container, which will probably carry goods from different sellers. The container is then transported to the Port of Loading (*POL*). This is the first moment in which customs administrations are involved, in this case, the customs office of a non-member state. Once a container has been loaded into a vessel it might be transshipped, meaning it can be shipped to an intermediate destination. This can be the case of a container terminal acting as hub for a shipping line, which is referred to as Port of Transshipment (*PoT*). Throughout this process, a container might visit more than one *PoT* outside the *EU* and be carried by a number of vessels. Therefore, the tracking of the containers is linked to the itinerary of these vessels.

Eventually, the container arrives to a Port of Discharge (*POD*). Here, the appropriate security and safety risks analysis is be performed by an European customs administration (see section 2.4) known as Customs Office of Entry (*COFE*), and possibly transshipped domestically to a European port of unloading. The customs office at the latter is referred to as the Customs Office of Unloading (*COU*). It must be noted that the *COFE* and *COU* can be identical, although a distinction is made for sake of completeness. The possibility of this difference can have a large role in the effectiveness of customs risk assessments, which will be discussed in more detail in section 2.4. Lastly, the container is transported to the stripping center (*SC*), where its cargo is unpacked and prepared for distribution to the *PLD*.

The logistics domain is used as reference to construct logistic events through associations between physical objects in time (a container being unloaded from a vessel) or between physical objects and locations (a container arriving to a terminal). These events might be built from historic data, represent real-time stakeholder interactions or be estimations for the execution of these interactions in the future. Each of these import supply chain *tracks* or *legs* represent different groups of stakeholders and data sharing processes, as well as requirements for their synchronization [128].

2.2 Interaction between Supply Chain Stakeholders

In order to understand the relationship between the cargo flow tracks identified in Figure 2.1 and the exchange of information, the functions and typical commercial relationships between *EU* import stakeholders are covered in this section. For instance, while the goods imported by enterprises are identified individually within the goods track, the pallets transported in a container are associated to this container´s itinerary for further reference beyond the *GC*. This involves complex coordination with a number of intermediaries shown in Figure 2.2, whose duties and responsibilities are important to clarify.

Although the stakeholders covered in this section might act as sellers or buyers, the functions and responsibilities of entities engaged in sales agreements behind the movement of goods are not included in this section. These contracts do not play a leading role in the exchange of physical logistic data, but the legal mechanisms by which sellers, buyers and shippers interact are very relevant in terms of data confidentiality and should be also taken into account, as covered in subsection 2.4.3. The overlap in roles in Figure 2.2 may vary largely between commercial transactions initiated through e-commerce platforms, such as individuals ordering items from an online store, and private international supply chains, where enterprises sell and ship goods to its own business divisions in other countries.

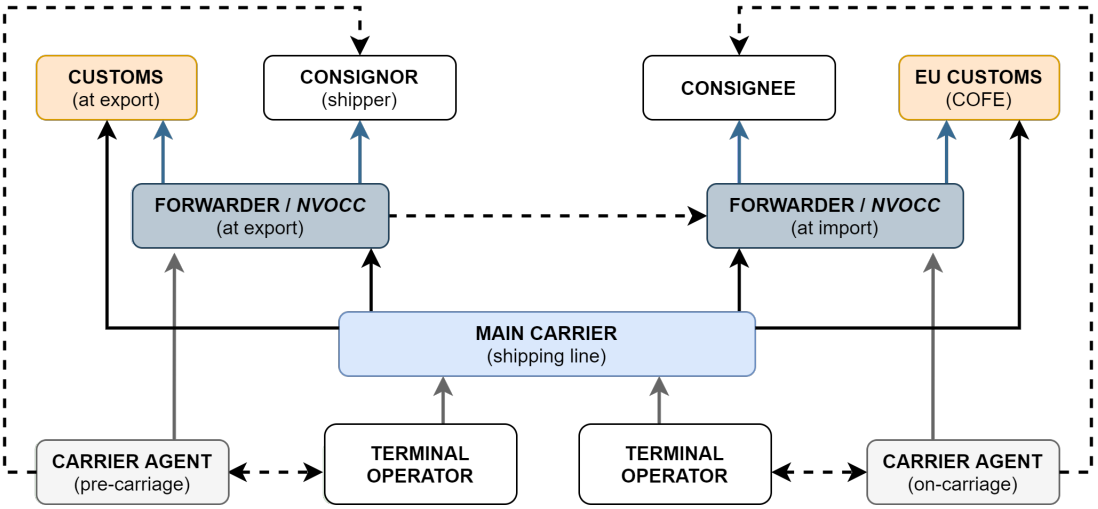


FIGURE 2.2: Stakeholder map for *EU* imports (*solid*: value delivery relation, *dashed*: data flow for process synchronization), adapted [166].

Regardless of the commercial character of these transactions, the physical flow of goods starts with a consignor or shipper, who establishes a contractual agreement known as contract of carriage (or contract of affreightment) with a carrier for the movement of goods towards a consignee or receiver. The consignee is the person entitled to take delivery of the goods, while the receiver is to whom the goods are physically delivered at the *PLD* [30, 176]. Although these two entities might not be the same, the terms are sometimes used interchangeably in maritime transport. section 2.3 covers the relevant aspects of contracts of carriage and their role in the exchange of data between supply chain stakeholders.

International shipping can become an organisational burden exposing enterprises, both large and small, to error and delay risks [161]. Due to the complex customs formalities and procedures to obtain cargo slots with shipping lines, exporters and importers usually make use of freight forwarders to deal with the logistic arrangements behind international cargo transport. Their existing relationships with shipping lines and their experience managing customs procurement in different countries make freight forwarders a key intermediary between the exporters, importers and deep sea shipping companies.

In the context of entry of goods into the *EU* through maritime traffic, a carrier is defined in the *Union Customs Code* as *"the person who brings the goods, or who assumes responsibility for the carriage of the goods, into the customs territory of the Union"* [53]. In other words, the carrier is the actor transporting goods from the *POL* to the *POD*. There are different types of carriers operating under varied contracts and with different responsibilities. A distinction between the *carrier agent* and the *main carrier* will be made for clarification purposes.

The term carrier agent will be used as the party responsible for managing the transport of goods from the *PLA* to the *POL* (pre-carriage) or from the *POD* to the *PLD* (on-carriage), sometimes including the booking of container and cargo slots. The pre-carriage and on-carriage can be executed with a range of modes, including road, barges or rail via intermodal freight terminals. The selection of the pre- and on-carriage modes will depend on the type of cargo transported (containers, bulk materials or dangerous goods) and the type of terminals at the *POL* and *POD*. If these are arranged and performed by the consignor, the consignee or a third party carrier agent, the process is called merchant-haulage. If they are arranged directly by the shipping line, the term used is carrier-haulage.

On the other hand, the main carrier, normally a the shipping line, refers to the operator of the mode of transport by which the goods enter the *EU*. Its functions focus on the organization of vessel operations, such as planning docking schedules with stevedores or synchronising cargo unloading with the port terminal authorities. The rise of vessel-sharing pacts and alliances between shipping lines has made it also necessary to explicitly differentiate between main carrier and shipping line. Two shipping lines operating different routes might decide to share their vessel capacity, leading to vessels carrying containers tied to the carriage contracts of external shipping lines.

When a main carrier does not operate its own fleet of vessels it is known as a non vessel operating common carrier (*NVOCC*), acting as carrier to the shipper and as shipper to the vessel-owning carrier. In terms of cargo responsibilities towards the shippers, these follow the same contracts as shipping lines, and normally lease large amounts of vessel space, but remain as accountable party in a number of billing procedures (see section 2.3 for a detailed analysis). These scenarios entail data-sharing arrangements to comply with the European customs declaration requirements discussed in section 2.4.

NVOCC's are commonly mistaken with freight forwarders. Both may book vessel capacity for reduced shipping rates (slot chartering) or offer warehousing services, either in-house or via third parties. Their main difference is the infrastructure ownership and their responsibilities towards the consignor or consignee. *NVOCC's* operate their own fleet of containers and are obliged to adhere to carriage agreements as main carriers, while freight forwarders limit their services to acting on agency of their clients to increase the bureaucratic and financial efficiency of their shipping activities. The value delivery of freight forwarders resides on their access to a deep network of carrier agents and shipping lines to provide dynamic, secure and cost efficient point to point shipping solutions. Ideally, the only point of contact of a shipper will be its freight forwarder, who should unilaterally facilitate the complex interactions between the stakeholders discussed in the previous paragraphs.

2.3 Flow of Contractual Information in Transport Chains

The stakeholders described in section 2.2 use contracts of carriage. To facilitate their fulfillment, the bill of lading (*B/L*) is a document used to legally bind the activities of supply chain actors [174]. Since their ultimate purpose is to ensure the delivery of goods at a destination to the person entitled to take delivery [150], the rules for issuing a *B/L* are very relevant in the digitization of trade information sharing. Storing and forwarding *B/L's* containing sensitive commercial information, such as goods classification, invoice value, sales contract terms, signatures or vessel voyage [148, 187], has strong data sovereignty implications as discussed later in subsection 2.4.3. Understanding the *B/L* issuing process is thus of special interest for the research, both in terms of data privacy and, as discussed in section 2.4, the European customs declaration process.

The practical function of a *B/L* is to confirm goods are shipped on a specific vessel and port for delivery to a named destination, showing evidence of the contract of carriage by reflecting the terms on which the goods are transported [150]. A *B/L* can be a *negotiable document of title* to the goods, meaning it provides a legal guarantee that the lawful recipient can exercise the right to demand delivery of the goods from the main carrier [38]. A simplified version of the issuing process is shown in Figure 2.3. Once goods have been received by the main carrier, he issues a *B/L* to the consignor via a freight forwarder. This means that the document is filled and signed by the main carrier, although other parties, such as a freight

forwarder or other agents, may fill the document with the carrier's consent [53]. Then, the *B/L* is shared with the consignee via the import freight forwarder, which hands in the *B/L* to the carrier's counter-party at destination in order to unlock and receive the goods. This simplified approach may be valid for scenarios with a reduced number of stakeholders, such as when carrier-haulage is used or when the freight forwarder acts as intermediary for bureaucratic coordination. This can be the case in large private supply chain governance structures, where the different stakeholders represent divisions of the same entity.

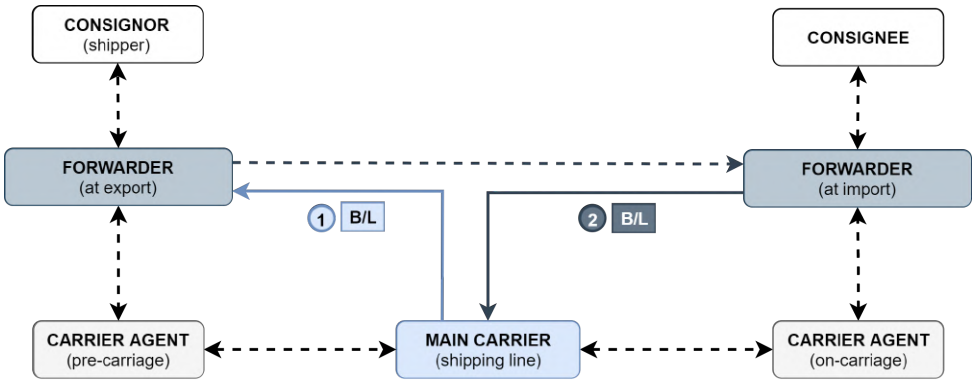


FIGURE 2.3: Simplified bill of lading issuing process, where dashed lines indicate other data flows for process synchronization.

However, a shipper can engage in an initial contract of carriage with a *NVOCC*, who then enters a secondary contract of carriage with a shipping line. In this situation, a *B/L* serves as evidence for agreements other than a contract of carriage between shipper and main carrier. This legal mechanism ensures that the shipper can take advantage of the specialized services of particular companies with the certainty that the right to obtain delivery of the goods will be eventually transferred to the buyer, with whom a commercial duty is held. This is the case when the main carrier is not the vessel operator, under complex multimodal merchant-haulage contracts, or when a vessel is leased under a charter-party agreement [124]. As a result, a sequence of *B/L*'s can be issued by different stakeholders downstream the supply chain based on the commercial nature of the underlying contracts being endorsed. The resulting contractual information flow during the issuing process is depicted in Figure 2.4.

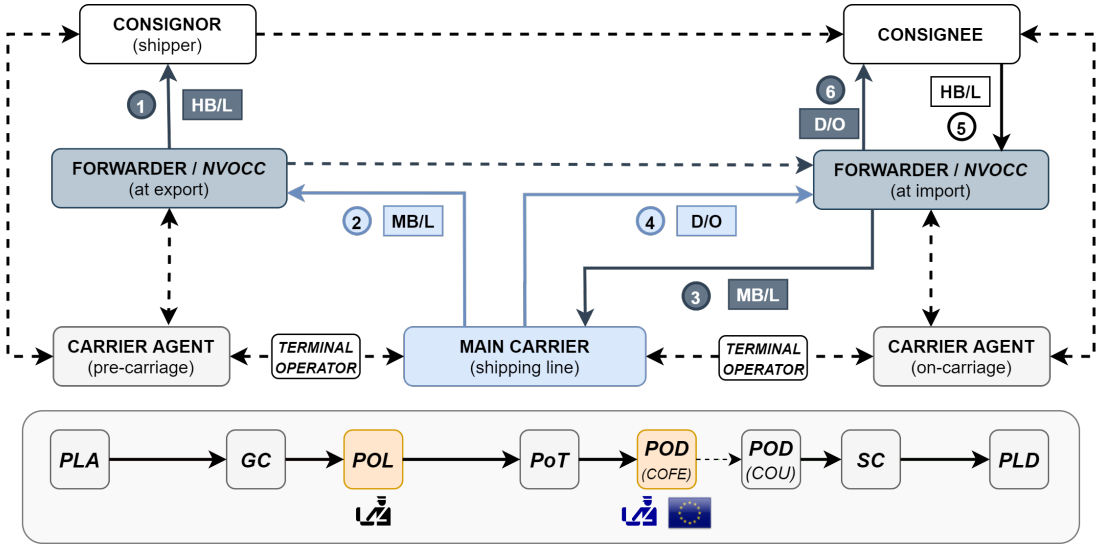


FIGURE 2.4: Bill of lading issuing process, where dashed lines indicate other data flows for process synchronization.

This diagram represents an initial mapping of the physical goods flow occurring in the logistics domain, described in section 2.1, into the private side of the digital data exchange across supply chains entering the *EU*. First, a freight forwarder or *NVOCC* (forward agent) issues a *House Bill of Lading* (*HB/L*) to the consignor once the latter has handed over the goods. In a *HB/L* the official shipper and recipient are normally the consignor and the consignee, and represents the contract of carriage between the consignor and the forward agent. The forward agent will be responsible for ensuring the goods reach the recipient.

After the main carrier has received the goods from the forward agent on behalf of the consignor, a *Master Bill of Lading* (*MB/L*) is issued. In a *MB/L* the forward agent is listed as the shipper, while its counter-party at destination is listed as the recipient. Similarly to the *HB/L*, the main carrier is now responsible for ensuring the goods are handed over to the forward agent's representative in the import side. This shows the aforementioned intermediary role of the forward agent, who acts as shipper to the main carrier and as carrier to the shipper.

The forward agent at import side must then show evidence of its lawful recipient status in order to unlock the goods, which is done by presenting the *MB/L* issued and signed by the main carrier. In exchange, the main carrier issues a *Delivery Order* (*D/O*) to allow the forward agent to access the goods. The original consignee must still show the *HB/L* issued by the forward agent at export side, which in turn produces a second *D/O* with which the consignee can finally take possession of the goods.

As shown in Figure 2.4, the different *B/L*'s need to be shared between stakeholders to facilitate the flow of goods. Certified post has been traditionally used for the exchange of original versions of these documents, but the need to cope with increasing trade by improving the efficiency of supply chains resulted in an industry-wide effort to migrate towards electronic document exchange alternatives during the last two decades. An example is the difference between goods being shipped after payment or goods being paid after delivery. Throughout the payment verification process, *B/L*'s are used, among other documents, to regulate such contract characteristics (see subsection 2.4.3). In this context, it is possible to encounter the scenario in which the goods, already paid by the recipient, arrive earlier than the original *B/L*. This can delay the transit of goods unnecessarily, create additional costs in customs procedures and temporary storage, and decrease the container throughput of terminals.

A solution for this type of issue is the *Straight Bill of Lading* (*SB/L*), which is a non-negotiable document (must not be transferred) and it's normally used in cases where the goods are being shipped between divisions of the same company or when the goods are shipped as a donation. This process might resemble Figure 2.3. However, when using *SB/L*'s, stakeholders are still vulnerable to a number of miss-delivery risks, as well as banking and insurance frauds. Therefore, negotiable or non-negotiable *B/L*'s are used at discretion of the consignor, who must evaluate the risks of the transaction and perform a safety-efficiency trade-off. This is an excellent example to showcase the added value of implementing *DLT*, particularly *BCT*, in the automation of supply chain data exchange.

By eliminating the need for such trade-offs, all parties involved in supply chains entering the *EU*, including European customs, can leverage a number of benefits. Besides the obvious ones, such as faster access to more and better information, these include a global and transparent chain of cargo custody and traceable documentation of compliance with improved data access and authenticity control [32]. This section has summarized part of the essential logistic contract data exchange in the private sector, from which customs administrations can extract very valuable information to perform better risks assessments.

2.4 Customs Data Declaration in the *EU*

Before goods enter European territory, customs administrations must perform a risk assessment. Based on the result, positive or negative loading permission is granted, or physical inspections are planned [53]. By means of these risk assessments, it is decided whether the arrival of the goods entails a risk that surpasses an acceptable threshold.

2.4.1 European Customs Risk Assessments

There are a number of ways incoming goods may suppose a threat to the *EU* and its member states. Different type of customs frauds can have economic, social, health or security consequences [148, 128]. Therefore, the methods used to assess levels of threat and to establish risk thresholds can vary largely per fraud. This is also the case between member states, which despite engaging in common economic and trade policies still exercise their right to use the fraud prevention models they deem more suitable to protect their national interests.

Although it is known that member states do not make use of the same risk assessment methodologies, there is an interest in finding paths towards deeper collaboration between their customs administrations [59]. This is not only due to shared economic interests and trade strategies, but in sufficiently similar identities enabling trust in security matters with profound social and cultural implications. Nonetheless, it is publicly acknowledged that the detection rates of customs procedures are constrained to a great extent by limited customs manpower, as well as the fact that excessive inspections would dangerously decrease the competitiveness of European ports [128]. This creates the need for enhanced passive risks assessments that do not interfere with terminal operations more than necessary. The PROFILE project is an example of such joint initiatives searching for innovative data-sharing technologies to bring European customs and port governance models closer [109, 128].

Under-valuating goods to avoid import tariffs, deliberately miss-classifying cargo or lying about its provenance are common import and export frauds encountered by customs administrations [87]. Interestingly, the detection of customs frauds can be closely linked to the *B/L*'s described in the previous section. Complex fraud schemes can be engineered around weaknesses during the issuing of *B/L*'s, the processing of the information they contain and the authentication of their provenance [69]. As a result, regulations on the customs declaration process are tightly linked to the interaction with *B/L* issuers [55]. This is evidence of the strong link between information sharing with supply chain actors and border protection, upon which the aforementioned interest to connect European customs has built up.

2.4.2 Entry Summary Declaration

The research does not intend to focus on the nuances behind the risk assessments themselves, but on the preceding customs declaration procedure, in which all actors mentioned in section 2.2 take part, either directly or indirectly. To that end, the remainder of the section focuses on the nature and purpose of the *Entry Summary Declaration* (*ENS*).

The *ENS* is formally described in the *Union Customs Code* as *"the act whereby a person informs the customs authorities, in the prescribed form and manner and within a specific time-limit, that goods are to be brought into the customs territory of the Union"* (*Article 5(9)*) [53]. This applies to all goods with some exceptions: electrical energy, goods entering by pipeline, items of correspondence, personal travel baggage, nonhazardous goods entering directly from offshore installations operated by a person established in the *EU*, and other cargo categories

ruled by national security and diplomatic treaties, are exempt from lodging an *ENS* upon arrival [54]. Simply put, an *ENS* is required for those goods being moved under a transport contract, which refers to the contracts of carriage discussed in detail in section 2.3.

Similarly to a *B/L*, there is a person responsible for submitting the *ENS*, but other persons may participate in this process with the consent of the declarant. The term *declarant* refers to the person lodging the *ENS* or the person in whose behalf it is been lodged [53]. In the case of deep sea containerized shipping, the responsible declarant is the person assuming responsibility for the carriage of the goods, i.e., the main carrier issuing a *MB/L* operating the final vessel entering the *EU*. Agents commonly involved in the filling of ENS are freight forwarders and *NVOCC*'s, although the exact distribution of lodging tasks are agreed privately among the relevant stakeholders, and often change with the characteristics of the underlying contracts of carriage for practical reasons.

The goods covered by an *ENS* are those unloaded in European ports and those consigned elsewhere and remaining on-board during European port calls [55]. The *ENS* must be declared to the first customs administration encountered by a vessel on its itinerary, referring to the *COFE* described in the logistics domain of Figure 2.1. The lodging must be completed up to 24 hours before the loading of the goods on the vessel, which does not count foreign transshipment ports [53]. For example, if goods will travel from China to Singapore, be loaded into a second vessel, and then travel to the *COFE*, the declaration of the China-Singapore route is not mandatory. However, if the vessel travelling from China to Singapore and from Singapore to the *COFE* is the same, such itinerary must be included in the *ENS*.

The declaration is rather straightforward, as shown in Figure 2.5. The declarant submits the *ENS* to the *COFE*, who is in charge of performing the risk assessment (*RA*) and sharing it together with the *ENS* with the pertinent *COU*. For cargo staying on-board during all port calls within the *EU* the *ENS* must only be lodged at the *COFE*, who will filter the relevant information needed by the *COU*. Also, in case a *RA* shows evidence that such on-board cargo poses a threat, the scheduled *COU*'s will be informed by the *COFE*.

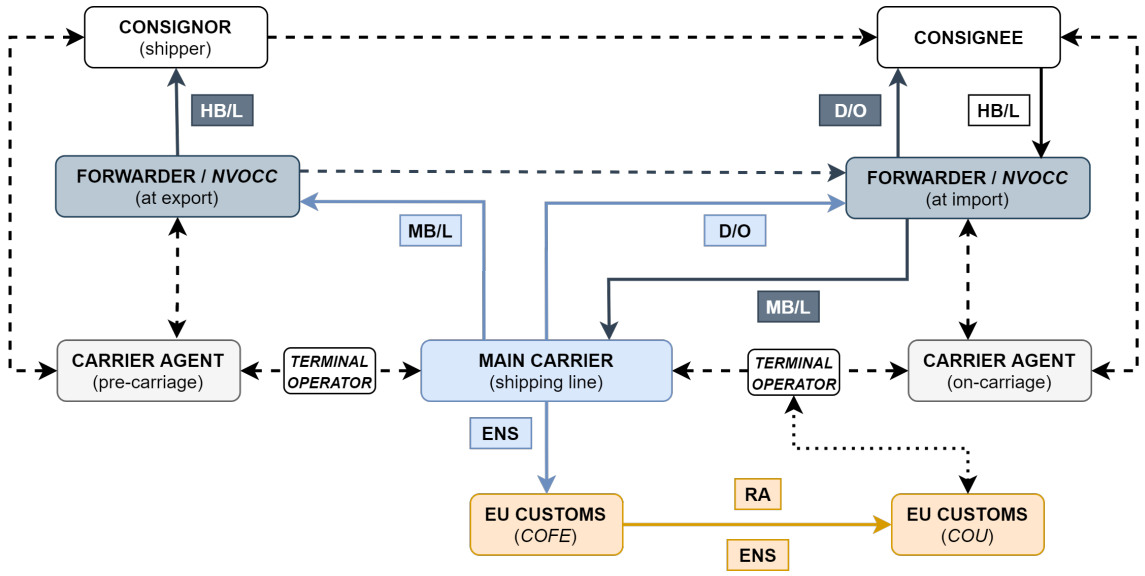


FIGURE 2.5: Combined entry summary declaration and bill of lading sharing, where dashed lines indicate other data flows for process synchronization.

Lodging the *ENS* directly to the *COU* is also a possibility included in the regulation, *"provided that the latter immediately communicates or makes available electronically the necessary particulars to the customs office of first entry"* [53]. Shipping lines are normally reluctant to do so, given that the *COFE* is still responsible for performing the *RA*, processing arrival notifications and emitting "do not load" warnings [54, 55]. Since not all member states support *COU* lodging, carriers tend not to make use of it. This is because they prefer to maintain a unique and standardised process compatible with their clients´ and collaborating partners´ activities regardless of the final cargo destination.

As explained in chapter 1, the private logistics sector is moving towards digital platforms for the exchange of information and documents [168]. The presented *B/L* issuing ecosystem is one of the processes envisioned to be fully digitized in the near future [32], and European customs want to leverage this new trends to incorporate *ENS* lodging and processing features [59]. This can be seen in the evolution of customs legislation. The *Union Customs Code* [53] has been recently updated in order to facilitate the interaction of European customs with commercial data-sharing platforms. This is reflected on the two following articles:

127(7): *"Customs authorities may accept that commercial, port or transport information systems are used for the lodging of an entry summary declaration provided such systems contain the necessary particulars for such declaration [...]."*

127(8): *"Customs authorities may accept, instead of the lodging of the entry summary declaration, the lodging of a notification and access to the particulars of an entry summary declaration in the economic operator's computer system."*

This new declaration approach adopted by the European institutions fosters the exchange of electronic documents through more efficient information systems. Figure 2.5 refers to the process followed for a specific *ENS*. In practice, a shipping line must submit a large number of *ENS*'s corresponding to a large number of *B/L*'s. Unless all cargo slots has been heavily booked by the same *NVOCC* or the vessel is being operated under a charter-party agreement to transport large amounts of goods for the consignor, it is very probable that the main carrier will deal with a document declaration decision-making problem. This scenario is depicted in Figure 2.6.

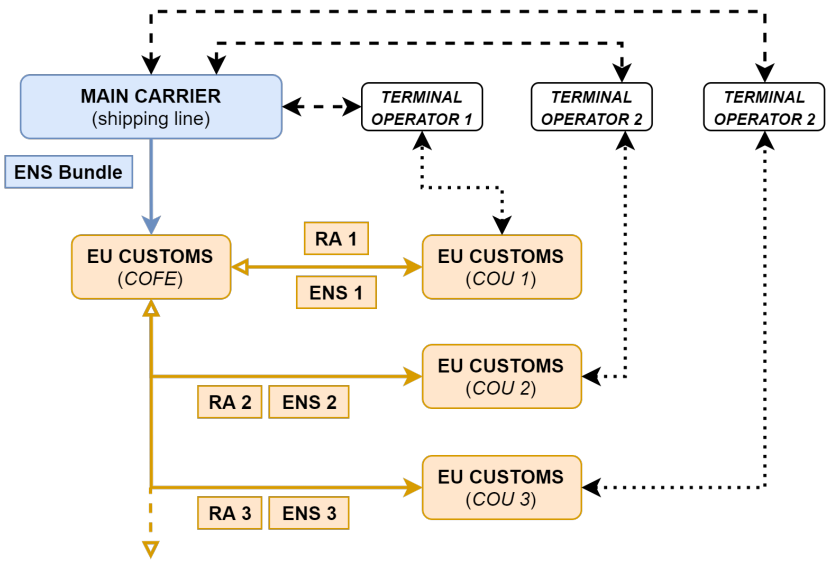


FIGURE 2.6: Entry summary declaration sharing between European customs.

Given the abundance of *B/L*s associated with the arrival of a large vessel, the main carrier decides the level at which *ENS*s are constructed. From one *ENS* for each *B/L* issued to *ENS* covering all *B/L*s issued (or any combination in between) are alternatives the main carrier can choose from. These decisions vary with the total number of *B/L*s issued and the complexity of the itinerary of the vessel, and can greatly affect how the European customs, specially the *COFE*, should disseminate information towards other customs offices (*COU*s).

The necessity of upgrading document exchange procedures between customs administrations and trade actors can also be seen in the case of diversions in vessel itineraries. After an *ENS* has been lodged in a *COFE*, it is possible to amend and/or update the *ENS* information when the scheduled vessel itinerary suffers changes. The main carrier must then inform the originally declared *COFE* of the changes through a diversion notification [55]. The original *COFE* must also forward all the relevant *ENS* information and *RA* results to the newly appointed *COFE*. This does not include goods that may have previously satisfied the declaration requirements for goods calling European ports under a transit procedure or goods being temporarily unloaded in order to accommodate other goods in the vessel [53].

The recursive nature of many *B/L*s and their associated contracts of carriage creates a waterfall of declaration duties. Every time a *B/L* is issued, the responsibility to enable the smooth transfer of goods between actors moves downstream, as well as the responsibility to submit accurate and legitimate customs declarations. However, the obligation to make the necessary information available stays at the bottom of the *B/L* issuing chain. This is reflected on the *Union Customs Code Delegating Act* [54] as follows:

112(1): *"Where, in the case of transport by sea or inland waterways, for the same goods one or more additional transport contracts covered by one or more bills of lading have been concluded by one or more persons other than the carrier, and the person issuing the bill of lading does not make the particulars required for the entry summary declaration available to his contractual partner who issues a bill of lading to him [...], the person who does not make the required particulars available shall provide those particulars to the customs office of first entry [...].*

Where the consignee indicated in the bill of lading that has no underlying bills of lading does not make the particulars required for the entry summary declaration available to the person issuing that bill of lading, he shall provide those particulars to the customs office of first entry."

This system allows exporters and importers to leverage the services of freight forwarders and shipping lines, and outsourcing customs declaration procedures. At the same time, companies avoid being exposed to customs compliance risks in case information is hidden or omitted to them by those to whom they issue a *B/L*. Therefore, a formal notification of the issuing of a *B/L* from the issuer to the recipient is also needed. In fact, the *B/L* issuing chain itself must be declared if the natural information flow downstream the chain does not occur. This is reflected on the *Union Customs Code Implementing Act* [55] as follows:

184(1): *"[...], the carrier and any of the persons issuing a bill of lading shall, in the partial dataset of the entry summary declaration, provide the identity of any person who has concluded a transport contract with them, issued a bill of lading in respect of the same goods and does not make the particulars required for the entry summary declaration available to them.*

Where the consignee indicated in the bill of lading that has no underlying bills of lading does not make the required particulars available to the person issuing the bill of lading, that person shall provide the identity of the consignee."

2.4.3 Role of Trade Confidentiality in Transport Operations

Data provided to customs administrations is protected by professional secrecy [53]. An *ENS* contains different types of information, some of which might be considered more or less sensitive. Regardless of the perceived level of sensitivity, all data provided to European institutions to be used during customs procedures is legally recognized as strictly confidential. Therefore, the information gathering and dissemination systems used in these procedures must adhere to the data protection provisions in force, as specified in the Council Decision of 2009 on the use of information technology for customs purposes [39].

Besides the actors enabling the physical transfer of goods across supply chains, there are entities responsible for underpinning trade transactions from a financial perspective. These play an essential role in the facilitation of trade, and include banks, insurers, credit institutions and other forms of trade financiers [59, 69]. They are responsible for issuing a type of document known as a *Letter of Credit* (*L/C*) [161]. It is considered the most important financial instrument used in international trade, working as a conditional payment guarantee with respect to the commercial contract between the shipper and the receiver of goods [17]. The transaction cycle of a *L/C* is tightly linked to the issuing of a *B/L*, as shown in Figure 2.7.

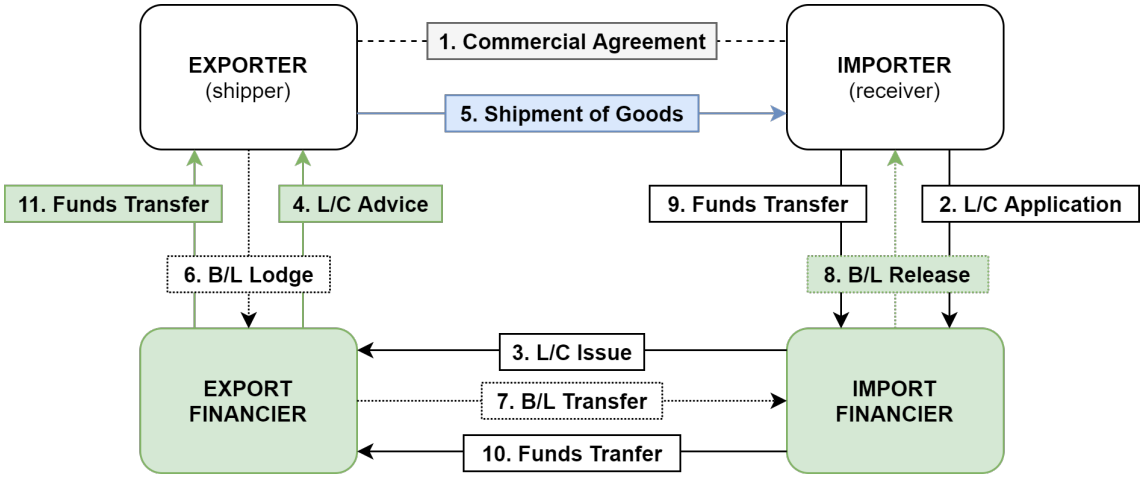


FIGURE 2.7: Letter of credit and bill of lading transaction cycle, adapted [17].

First, a shipper and a receiver enter into a commercial agreement, where the *L/C* is chosen as payment method, which is common practice in international trade as a legal incentive to protect payment obligations between trade actors [18]. Then, the receiver applies for a *L/C* to an import financier who will issue such document reflecting the nature of the contract of sale included in the commercial agreement [17]. After performing the pertinent due diligence, the issuer forwards the *L/C* to the shipper through a trusted banking entity at export.

The shipper, being the seller or an agent acting on its behalf, can then assess whether the terms in the *L/C* match the agreed contract of sale, propose necessary amendments, and prepare the shipment of the goods. Then the shipper must present a valid *B/L* to the export financier to be forwarded to the import side. After this step the shipper can prove compliance with the contract of carriage and consider the payment as assured. Lastly, the cycle is completed with the receiver, being the buyer or an agent acting on his behalf, accessing through the export financier the shipping documents necessary to unlock and claim the goods. However, on top of the *L/C* mechanism used to settle the contract of sales between the seller and the buyer, there are additional contractual obligations within the physical logistics domain that need verification before the goods can be released.

Shown in Figure 2.8 are the relationships behind three conditions for goods release besides the buyer-seller *L/C* settlement: commercial release, customs release and discharge [61]. Customs release refers to the fulfilment of all declaration obligations, such as the *ENS*, while discharge is the physical availability of the goods at the port terminal. The commercial release can be more complex, and refers to the payment of the transport of the goods.

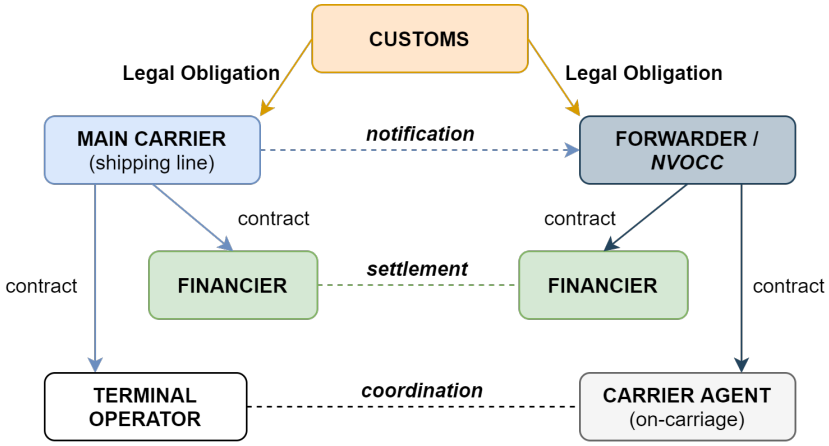


FIGURE 2.8: Contractual relations for terminal operations, adapted [61].

The payment guarantee for *B/L*'s issued between supply chain actors (see Figure 2.4) resembles the *L/C* cycle. It involves the exchange of information between the main carrier, the consignee's agents and their financiers, which has been simplified in the figure as *settlement* between the main carrier and the freight forwarder. This does not only apply to the final release at *POD*, but also to the on-carriage transshipment legs towards the *PLD* [61].

As explained in section 2.3, the exporter can rely on the *B/L* to transfer responsibility once a freight forwarder or shipping line has taken custody of the goods, and in doing so transferring liability downstream the physical supply chain. Similarly, by issuing a *L/C* the financiers become liable for any damage produced if the *L/C* was to be used by an unlawful recipient to conclude extra-official transactions [18]. As a result, trade financiers are directly interested in the improvement of *ICT* systems to enhance confidentiality while accelerating the exchange of information in the support of trade activities. This interest is also fueled by the fact that a *L/C* can be a source of international commercial fraud [13].

Combining the analysis of the *B/L* issuing process (Figure 2.4), the European *ENS* lodging procedure (Figure 2.5) and the *L/C* transaction cycle (Figure 2.7) reveals a hidden interaction between contracts of carriage, customs regulations and trade finance. One of the consequences is a very intricate web of contractual and regulatory responsibilities among traders, carriers, banks and institutions [17]. This poses a grand challenge for European institutions: allowing robust national security and organic economic development to coexist by establishing paths to collaboration and enabling transactions to occur in such complex ecosystem. The data exchange in trade activities covered in the previous sections also shows the legal and operational dependencies between the commercial releases throughout the cargo custody chain, the customs control at transport infrastructure and the financial due diligence of trade through the *L/C* transaction cycle. The impact of these dependencies transcends the commercial realm, affecting the efficiency of international transport itself and the safety of services offered by financial institutions. As a consequence, the role of data privacy goes beyond hiding sensitive sales information, and serves the much deeper purpose of avoiding a distrust ripple effect across the entities responsible for prosperous, stable, legal trade.

2.5 Digital Infrastructure & the Future of European Trade

The concept of *trusted trade lanes* (*TTL*) forms part of a regulatory supervision strategy that can be used by European customs to navigate the data transformation that the logistics sector is undergoing [81]. They represent collections of transport activities carried out by a group of enterprises connecting a source and a destination, in which supply chain actors implement internal control systems that allow the detection and reporting of suspicious events while complying with the requirements imposed by customs administrations [81]. The realisation of this new supervision strategy is envisioned as a potential solution to the policy dichotomy European institutions are facing: improving regulatory compliance while reducing the administrative burden to enterprises and facilitating trade [81].

There are two key figures identified within the new regulatory supervision model: unknown traders and trusted traders [128]. Unknown traders are those declarants and supply chain actors from which no real risk references can be extracted, being risk assessments completely based on itinerary and cargo descriptions. On the other hand, trusted traders are entities that voluntarily have subjected their internal business processes and *ICT* systems to an exhaustive audit in exchange of simplifications in their declaration duties. Such system is already used in the *EU* with the figure of *Authorized Economic Operator* (*AEO*), to which European customs ensure *"favourable treatment in respect of customs controls, such as fewer physical and document-based controls"* (pp. 5) [53].

In terms of regulatory supervision, a *TTL* falls under system-based regulation [81]. As opposed to more invasive regulatory supervision strategies - such as physical inspections in container terminals - system-based regulation analyses more complex organisational processes and certifies their adequacy towards a specific regulation [81]. In that regard, *TTL*'s and trusted traders follow similar supervision strategies. For instance, a key requirement to receive the *AEO* status is the demonstration by the applicant of sufficient control of his operations by means of secure and trustworthy commercial and transport transaction records [53]. A similar approach is applied to *TTL*'s, where the architecture supporting the exchange of information does not only verify the compliance of an individual stakeholder. Instead, the legitimacy of the commercial transactions and the movements of goods taking place in the ecosystems described in the previous sections is guaranteed.

To this end, it is interesting to cover the concept of *data pipeline*, which was originally introduced by *Her Majesty's Revenue and Customs* and the *Dutch Tax and Customs Administration* as an innovative solution to increase the visibility of the supply chains passing through their borders [155]. In a nutshell, a data pipeline is a form of digital trade infrastructure connecting public and private supply chain stakeholders [142]. It consists of different data sources and information systems connected through a serial *pipeline* with the ultimate purpose of enabling dynamic and secure inter-organisational information sharing [96]. A graphic representation in the research context is shown as the *Data & Document Layer* in Figure 2.9.

Data pipelines are also intended to eliminate data redundancies and inaccuracies during customs declarations [75]. This is achieved by ensuring that the information about companies, commercial transactions and movements of goods is gathered at the original source, as far upstream the supply chain as possible [155]. The main benefit is avoiding data discrepancies during audits performed by public institutions, making it easier for the latter to cross-validate declaration data [49]. For example, customs could detect risks earlier and with increased reliability [75] as data is shared through dedicated channels, which makes tracing data back from transaction and logistic event logs simpler. A scenario avoided by

trusted trade lanes is when a shipper deliberately undervalues transported goods. This is done because the freight rates charged by carriers are associated to the liability risk they are exposed to, which increases with the value of goods [75]. The shipper provides then a false declaration and enters a secondary insurance to cover the risk. This would be difficult to hide in a data pipeline connecting the main carrier, banks, insurers and customs.

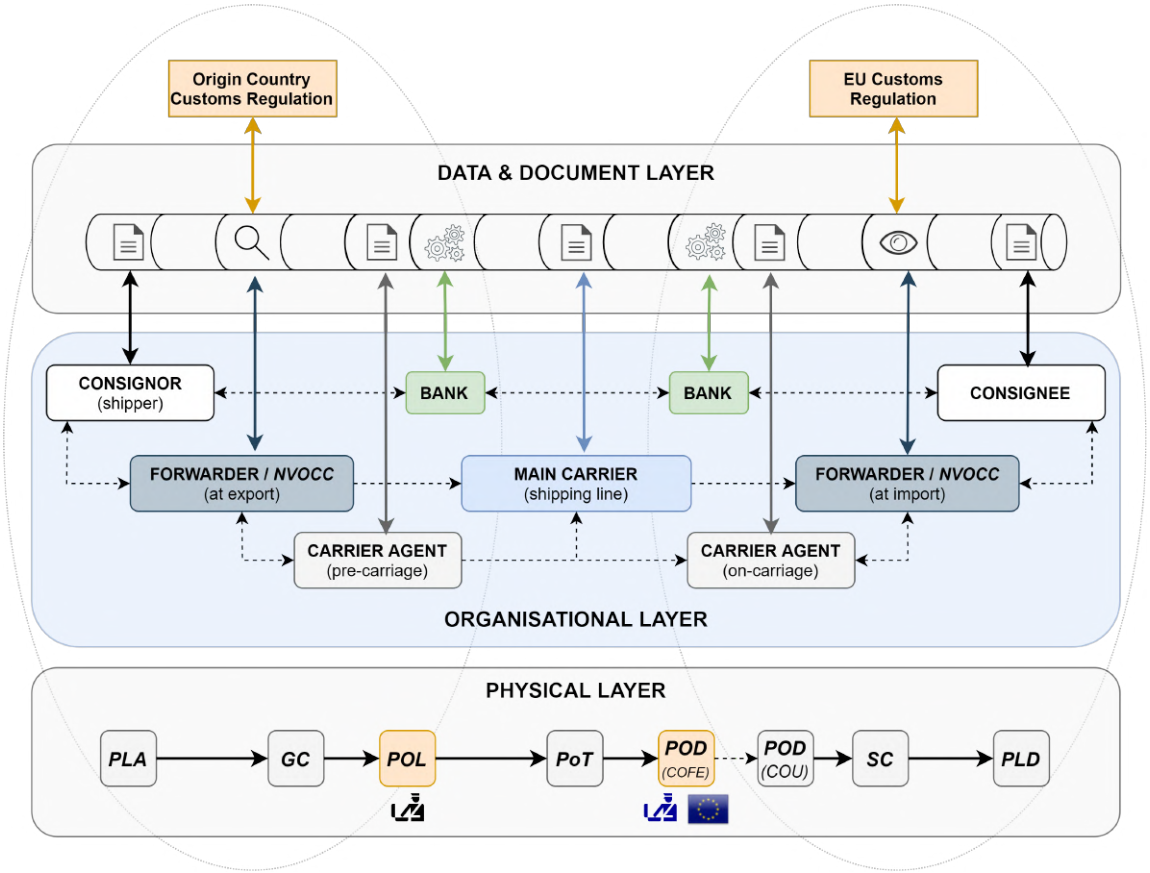


FIGURE 2.9: Data pipeline concept, adapted [155, 166].

European customs have historically used the physical layer to gather the information for risk assessments. This is due to the lack of sufficiently mature *ICT* systems able to differentiate between the three layers proposed in the data pipeline concept and their interactions [75]. The data pipeline approach has the potential to accelerate the transition towards a European system-based custom supervision by re-using supply chain data for other purposes than it was produced for, known as data piggybacking [162].

How traditional data-sharing procedures can be useful in today's trade scene - such as the issuing of a *B/L* and *L/C*-based transactions - have been called into question upon the difficulties for their integration in globally digitized commerce [13]. However, this perspective comes from a transaction-based supervision approach, in which individual transactions are inspected instead of the context in which they are carried out [81]. Following the new system-based supervision paradigm there is an on-going effort to develop *ICT* systems based on *BCT* to automate the processing of *L/C*s together with *B/L*s and other shipping documents [69], discussed in detail in subsection 2.6.3. This trend indicates that trade audit mechanisms and the control of physical logistic processes are destined to come closer in the near future, and that *BCT* will play a leading role in solving the new generation of challenges to feed global trade lanes with high quality data [105, 110, 168].

2.6 Commercial Blockchain Platforms

BCT is a very powerful tool to tackle the deployment, maintenance and governance of digital infrastructure - such as data pipelines - aimed at fostering the exchange of information between private businesses and government authorities, also known as business-to-government (*B2G*) communication [49]. It is specially useful for storing transactions securely in ecosystems where complete trust between actors is not provided and relying on a single authority to govern the infrastructure is not feasible or undesired [59, 168]. As discussed in section 2.3 to section 2.5, the digitization of the transport documents involved in the customs declaration process is one of these ecosystems.

2.6.1 Introduction to Blockchain

In general, a *ledger* can be defined as a distributed data structure that contains entries acting as digital records of actions [12]. Distributed ledgers are commonly used to enable the concurrent editing of a shared digital asset or transaction system while maintaining its state unicity [94]. A blockchain is a type of ledger-based database that creates a digital record of transactions, where each of the blocks building a sequential chain is associated to a timestamp updated simultaneously by all the participants in the network [47]. Each block contains the cryptographic hash of the previous block, from which the name *blockchain ledger* originates. This process is shown graphically in Figure 2.10.

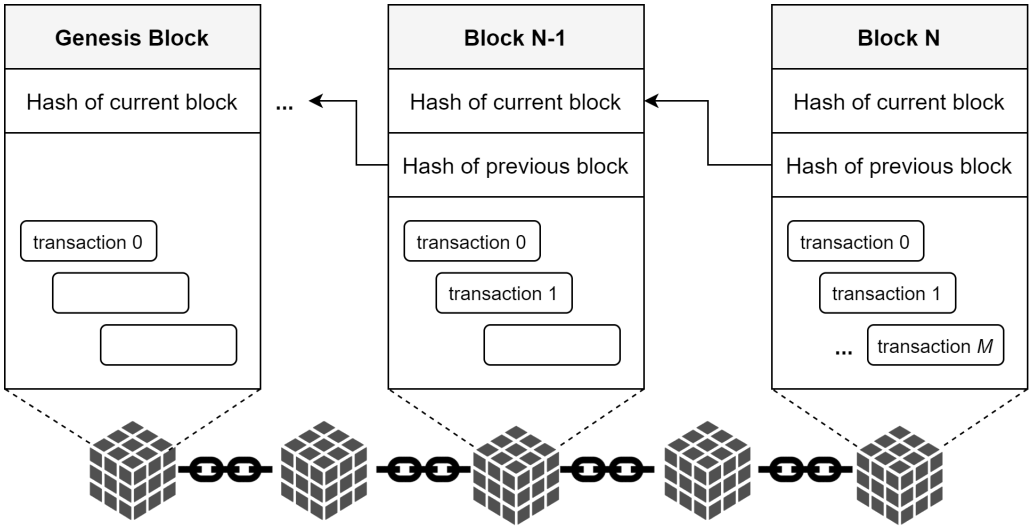


FIGURE 2.10: Data architecture in a blockchain network, adapted [194].

If this design is applied to a distributed computer network as shown in Figure 2.12, the result is a peer-to-peer (*P2P*) network, in which every node keeps an independent copy of the transactions executed [148], or at least a copy of the chain of their unique hashes, from which the order of execution can be audited [47]. One of its features is immutability: the use of cryptography makes it impossible for a single party to alter the history of transactions once stored in the ledger [181]. The transactions are not only timestamped but also digitally signed by the actors involved. Depending on the type of blockchain used, this can prevent users form hiding their authorship or ultimately neglecting the liabilities associated to their actions, also known as non-repudiability [164]. As blockchains function under strong non-repudiation and record irreversibility principles, they are considered very secure data storing systems [165] and particularly useful in regulation-driven processes [122].

Besides preventing changes in the transaction record, entering fraudulent transactions that mismatch the current state can also be avoided [164]. Every node in the network maintains its own copy of the transaction history, so all nodes need to reach consensus on the state of the ledger in order to achieve this. As opposed to a traditional centralised network like Figure 2.11, there is no governing reference available to all nodes, which makes consensus mechanisms a key element of blockchains [195]. The consensus mechanisms will dictate whether a new transaction is accepted or rejected by a sufficient number of nodes, and ultimately included into the ledger [148].

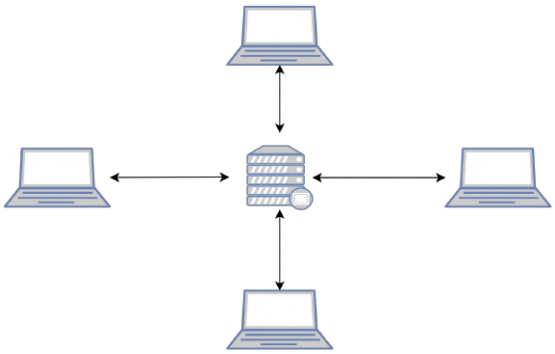


FIGURE 2.11: Centralised network.

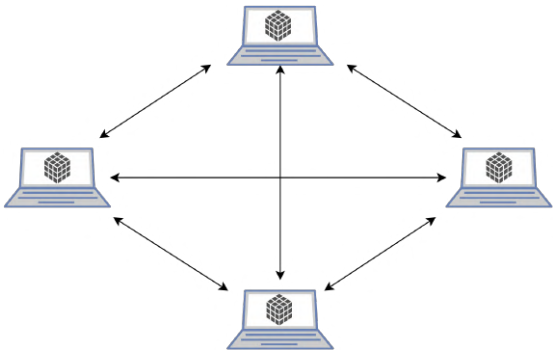


FIGURE 2.12: Distributed network.

There are three main type of blockchains. Public or permission-less blockchains allow any individual to become a node of the network [122], perform transactions and access the transaction history, which is mostly common for large-scale financial applications, such as Bitcoin and Etherum [59]. Public blockchains require a consensus protocol based on a transaction validation algorithm like the *proof of work*. These normally operate under reward policies for those involved in the verification of legitimate transactions by solving cryptographic puzzles [195] and turning malicious activities unattractive in the process [148, 168].

On the other hand, the nodes of a private or permissioned blokchain belong to a closed consortium of known users that control the addition of nodes to the network, which elevates the power of non-repudiation [59]. Private blockchains are commonly used in business-to-business (*B2B*) settings and tend to have more efficient consensus protocols, enhanced network throughput and reduced transaction latency [194]. There are also hybrid blockchains, in which a reduced number of nodes control the addition of new ones, as well as whether the latter can access the transaction history or their ability to execute transactions.

2.6.2 Automating the Execution of Contractual Obligations

The term *smart contract* was coined by Nick Szabo in 1996 as a reflection on how legal systems are destined to evolve in the *cyberspace* era, particularly the role of cryptography in the accomplishment of verifiable, private, digital transactions [115]. In the present, smart contracts refer to agreements between stakeholders that are programmed and stored in a blockchain to digitise assets or automate the execution of business transactions [144]. They allow to define conditions, obligations and rights between stakeholders that can be activated and enforced by information acquired through blockchain transactions [194]. A smart contract is an interesting tool for the implementation of system-based regulatory supervision as discussed in section 2.5, due to its vast potential to reduce fraud [148, 178, 194]. Some examples are certified origin solutions, dispute resolutions or smart documents to assist in the cross-validation of declaration information and accelerate trade transactions [92, 148, 168].

2.6.3 Electronic Transport Documents

As mentioned in section 2.5, *BCT* has the ability to reduce the complexity behind the *B/L* issuing process (Figure 2.4), the European *ENS* lodging procedure (Figure 2.5) and the *L/C* transaction cycle (Figure 2.7). Figure 2.13 shows a conceptual model for blockchain-based document sharing, where the compliance of contracts of sales between exporters and importers through a *L/C*, as well as the contracts of carriage between physical supply chain stakeholders through a *B/L* can be verified. This is an example of how *DLT*-based architectures are envisioned to enable trusted trade lanes by linking trusted data sets.

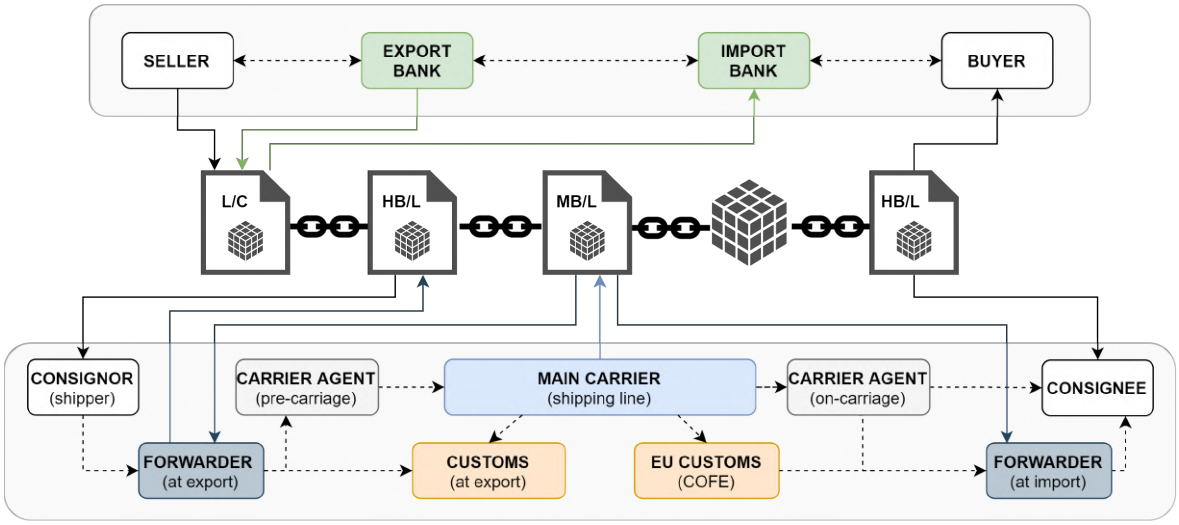


FIGURE 2.13: Conceptual blockchain-based document sharing.

This type of architectures for the support of digital transport documents, like electronic *B/L*s (*eB/L*s), are already being used, but their legal functionality is limited and fails at conserving their full *negotiability* (see section 2.3) across the complete logistics domain [69]. This creates an unbalanced exposure to compliance and fraud risks between trade stakeholders. As a result, currently available forms of digital assets integrated in blockchains can not yet be considered fully reliable vehicles for critical documentation throughout supply chains entering European territory nor for all the participants therein [121].

2.6.4 Blockchain-based Ecosystem for Supply Chain Data

Despite the benefits that new digital infrastructures may bring to an industry, it is difficult to abruptly redefine deep-rooted practices around which regulations have evolved. Further work on the standardisation and privacy of blockchain-based data objects is required before smart contracts for automated document processing are embraced as reliable solutions for inter-organisational data sharing in the shipping industry [194]. Initiatives such as the aforementioned *TradeLens* intend to close this gap by providing enterprises and institutions with a single source of shipping data [82, 121].

However, existing solutions focus on specific supply chain segments. They do not capture all the transactions executed throughout the complete logistics domain. For example, *TradeLens* supports an ample ecosystem of actors, but focuses on ocean shipping and excludes short-sea and air space activities (*e.g.*, pre-carriage and on-carriage multimodal transshipment [188]. This reduces the visibility between enterprises and institutions and is one of the reasons it is difficult to integrate information flows under the same blockchain architecture.

In addition, there is a lack of research on the use of smart contracts for the automation of customs processes, such as the lodging of *ENS* [148]. The data pipeline concept shown in Figure 2.14 emerges from this context, where the whole logistics domain is mapped into continuous information flows. This way, companies and institutions can coordinate their activities and expedite the verification of logistic milestones.

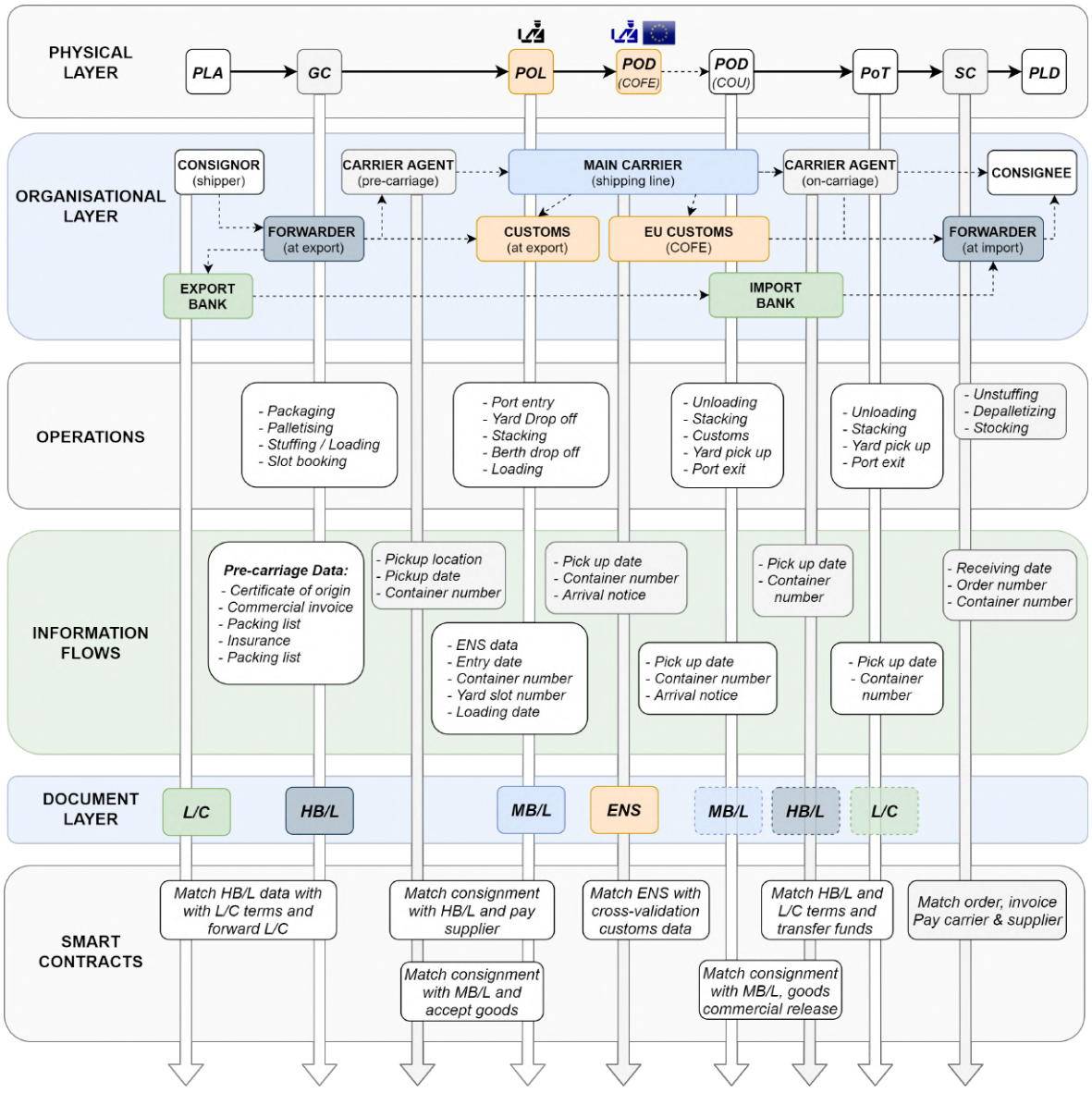


FIGURE 2.14: Proposed shipping data pipeline ecosystem, adapted [117].

Following the current trend for the development of supply chain data pipelines covering the logistics domain partially, the evaluation of their integration is a very interesting first step towards trusted trade lanes. Nevertheless, the data sovereignty and platform interoperability implications of this endeavour are still covered by uncertainty. The next phase of the research is an exhaustive assessment of the requirements for the design of a peer-to-peer data sharing architecture that enables the information flows between private business partners and public institutions as shown in Figure 2.14.

2.7 Conclusion

This chapter tried to answer how data is shared between supply chain stakeholders and customs administrations. The document flow containing information about the contracts of carriage governing the movement of goods has been analysed, and the role of blockchain technology in improving the efficiency and transparency of this process has been discussed. The following paragraphs provide answer to the research sub-questions and a concluding answer to *RQ1*.

How is data shared between supply chain stakeholders?

The term data is very broad. The research has focused on the data included in documents used for the development of commercial activities related to containerised shipping. Supply chain stakeholders make use of contracts of carriage to agree on the terms for the transport of goods. These agreements are laid down in bills of lading (*B/L*s), which offer a legally binding proof of the existence of the aforementioned contracts, and the subsequent responsibilities between the parties involved. This document can be considered the cornerstone of international trade. The concatenation of *B/L*s is used to form trusted cargo custody chains and take advantage of complex multimodal transport hubs. For this reason, a unique network of stakeholders with business models built around the exchange of bills of ladings has emerged in the last decades.

How is data shared between European customs administrations and supply chain stakeholders?

The *Entry Summary Declaration* (*ENS*) is the declaration procedure used by European customs administrations when cargo is imported into the *EU*. The main carrier is responsible for submitting a legitimate and complete *ENS* before cargo arrives to a European port. However, the information provided by the carrier in an *ENS* is nothing more than a selection of data already present in previously issued documents (mainly different types of *B/L*s) plus additional information about the itinerary followed by the vessel before its arrival to the European port.

Traditionally, cargo declarations were submitted directly to customs administrations as individual documents that aggregate all relevant information about the cargo transported by a specific vessel. However, European customs regulations have evolved in the recent past to allow carriers, or authorised entities acting on their behalf, to provide access to *ENS* data in the form of links to private information systems.

What is the role of blockchain technology in supply chain visibility?

Despite the evolution of European regulation and the advancement in information technologies, the transport industry has faced challenges for the digitization of *B/L*s. The highly trustless nature of the interactions between supply chain actors and the lack of reliable technical solutions able to combine operational improvements with legal certainty has been preventing transport processes from reaching its full efficiency potential.

Blockchain technology has proved its ability to reduce the friction of information sharing by expediting the issuing and processing of *B/L*s while guaranteeing the security of cargo ownership chains. This is the reason commercial blockchain platforms have gained popularity in the shipping industry and show the potential to become the new industry standard for *B/L* management.

The data pipeline concept emerges from this trend, which envisions dynamic and secure inter-organisational information sharing by leveraging the security and verifiability characteristics of these platforms. Therefore, blockchain technology, and *DLT* in general, offer data sharing incentives and promotes trust in the data provided by other entities. For this reasons, it is envisioned as an enabler of supply chain transparency.

Answer to Research Question

This chapter has provided insight in a number of data sharing processes within and between private and public entities. A deeper understanding of the exchange of contractual information between supply chain actors, the declaration process used by European customs administrations and the role of blockchain technology in the visibility of supply chain data has been gathered to ultimately answer RQ1: ***What is the relationship between supply chain visibility, import declarations and the risk assessments performed by customs administrations?***

Up to now, the only eyes customs administrations have counted with are customs entry declarations. Their visibility of supply chains has been limited to the information provided by the last hands cargo passes through before arriving to the *EU*, mainly shipping lines and other logistic service providers. In the recent past, European customs have been willing to trade-off the consistency of physical cargo inspections for the collection of high quality declaration data. Shipping companies have been incentivised to share more information with a relaxation of customs requirements and privileges in customs facilities.

Currently, the rise of commercial blockchain platforms is seen by European customs as an opportunity to not rely solely on economic incentives to increase their visibility over supply chains entering the *EU*. They envision an ecosystem in which a data infrastructure able to provide a constant stream of high quality declaration data can be deployed. This would have a direct impact on the reliability of customs risks assessments and the competitiveness of the busiest European transport hubs. However, European customs should not take for granted the effectiveness of interacting individually with this new generation of data sharing service providers.

Regardless of the sophistication level of the data infrastructure used by public institutions, there will be a limit to its effectiveness directly associated with private sector practices. Before data quality can be improved, an environment where entities are less reluctant to share their information must be fostered. This can only be achieved by giving the keys to data governance back to the producer of the data. It is such scenario that has the potential to trigger an increase in stakeholder interactions due to less friction in information sharing, which would eventually lead to a richer data landscape that has grown organically, and that customs administrations could then leverage.

The default participation of customs administrations in all blockchain platforms could decrease the administrative hurdle of customs declarations to a certain extent. However, this is not the answer for a long-term increase in terminal throughput and enhanced border control, as siloed information flows would still be destined to form. To avoid this, a different way to communicate between private and public entities is needed. The main driver is thus equipping the users of commercial blockchain platforms with the technical capabilities to generate verifiable links to trusted data, which can then be used by customs administrations to perform their institutional duties.

Chapter 3

Requirement Definition

In this chapter, the requirements of the researched artefact are identified. Sometimes referred to as *Design Principles Induction* [99], this is a crucial phase of the design science methodology in order to arrive at a sufficiently constrained design space. Using the *Problem Explication* covered in chapter 2 as input, this phase provides design guidelines to be used during the *Design & Development* phase covered in chapter 4. Using the research framework presented in Figure 1.2 as reference, the requirement definition can be interpreted as the set of design constraints emerging from design principles chosen by the researcher as a guide towards relevant and rigorous design decisions during the *Design & Development* phase [99]. These principles represent the boundary between theory and practice where the researchers and users of the artefact transfer knowledge between each other [66]. The following sections cover the process used to arrive at the design principles and requirements of the researched artefact and ultimately answer the following research questions:

RQ2: *What are the design requirements to preserve the data sovereignty of supply chain actors?*

RQ3: *What are the design requirements to allow multiple blockchain platforms to use the architecture?*

The design principles used in the research are presented in section 3.1. Then, section 3.2 covers the requirement categories included in the analysis. The functional and non-functional requirements elicited from each design principle are covered in section 3.3 to section 3.5. An overview of the generated requirements can be found in section 3.6.

3.1 Design Principles

Since the purpose of design science research is to produce artefacts able to solve real world problems, the design principles should reflect the core challenges that the research is addressing. Design principles can be derived inductively or deductively [99]: inductive principles emerge form the needs found in the real world to provide a firm source of relevance, while the deductive principles emerge from the supporting scientific work and facilitate the rigorous application of academic knowledge. An artefact can also belong to a larger class of artefact goals that share design principles and requirements, known as meta-requirements [21]. In this research, the term *design principle* is used to refer to an inductive principle, while *blockchain principles* will be used to refer to the meta-requirements of blockchain-based data-sharing artefacts.

The goal of the researched artefact is the preservation of data sovereignty and the enhancement of blockchain architecture interoperability while improving event visibility across the logistics domain of Figure 2.1. A special focus is given to data-sharing for customs purposes within the system-based regulatory supervision context introduced in section 2.5. Shown in Table 3.1 are the chosen design principles (*DP*) driving the utility of the artefact based on event visibility, data sovereignty and architecture interoperability.

TABLE 3.1: Overview of design principles (*DP's*).

code	**name**	**description**
DP1	Logistic Event Visibility	Provide features that allow European customs to maintain the visibility of logistic events across the logistics domain
DP2	Stakeholder Data Sovereignty	Provide features that ensure the preservation of the data sovereignty rights of supply chain stakeholders
DP3	Architecture Interoperability	Provide features that support the compatibility of data exchanges between European customs and different blockchain platforms

Limiting the design principles to these three categories does not imply that other artefact properties that *BCT* entails will be completely overlooked. For instance, any blockchain artefact requires considering scalability, storage and governance implications [148, 165]. However, these blockchain principles will be eventually identified throughout the requirement generation if they are found needed to complement the design principles.

3.2 Types of Requirements

Requirements can be divided into two main categories: functional and non-functional. Functional requirements cover the tasks to be executed and drive the application architecture, while non-functional requirements provide criteria to assess the operation of the artefact and define its technical architecture [5, 90]. This two categories are a good initial mutually exclusive and collectively exhaustive framework, but their parallel compliance is of utmost importance. In fact, insufficient monitoring of dependencies and compatibility assessments between functional and non-functional requirements is one of the largest failure drivers in information systems from a requirement engineering perspective [41].

Transferring requirement engineering practices from traditional scientific disciplines to information systems has systematically suffered from a lack of consensus among researchers and professionals [21, 22]. This situation has been addressed by the academic community during the last three decades, starting with the absence of requirement generation theories for *vigilant information systems* aimed at detecting *"discontinuities in the organisational environment relevant to emerging strategic threats and opportunities"* (pp. 37) [182]. Most of the first studies aimed at filling these knowledge gaps were focused on the optimisation of business processes, but the aforementioned description resembles to some extent the current definition of system-based regulatory supervision. As public administrations have gradually embraced their responsibility to lead this race, such studies have become more common in public initiatives like the PROFILE project. Interestingly, governments and companies are still facing similar difficulties when using design science research to define requirements for modern digital infrastructure [22], which is partly the motivation of this research.

Although there are currently numerous approaches on how to specify design principles and their accompanying artefact requirements [63, 66, 196], it is essential to ensure that the requirements generated can be mapped into at least one of the chosen design principles. This is necessary in order to effectively assess artefact functionalities during the *Demonstration & Evaluation* phase (covered in chapter 5) by tracing the consistency of specifications from design decisions to design principles [99].

3.3 Event Visibility Requirements

Most of the information used for customs risks assessments is produced upstream supply chains and included in documents created well in advance the initiation of any customs declaration process, such as commercial invoices, packing lists from grouping centers, *B/Ls* or certificates of origin [175]. However, this information tends to accumulate in silos inaccessible by customs administrations, or are at least not integrated in the customs procedures. To this end, the proposed artefact is expected to provide institutions timely access to information through piggybacking, as set out in the following requirement:

Functional Requirement 1	*FR1*
The artefact should allow European customs to access information produced upstream supply chains entering the European Union.	

The disassociation of products and freight cargo is an inevitable consequence of the high level of segmentation of supply chains, which implies decenteralised knowledge [180]. This is addressed in section 2.1 with the goods and container tracks of the logistics domain. The fragmentation of the logistics domnain is visualised in Figure 3.1. The unification of every single source of logistic event data into a common frame of reference is probably unattainable as the competitiveness of supply chains is linked to the disaggregation of functions to maintain scalability. However, creating robust links between supply chain segments to align diverse actor ecosystems is considered an interesting approach to increase supply chain visibility. In this context, the artefact is an instrument for the creation of these links and the reduction of supply chain risks for European customs.

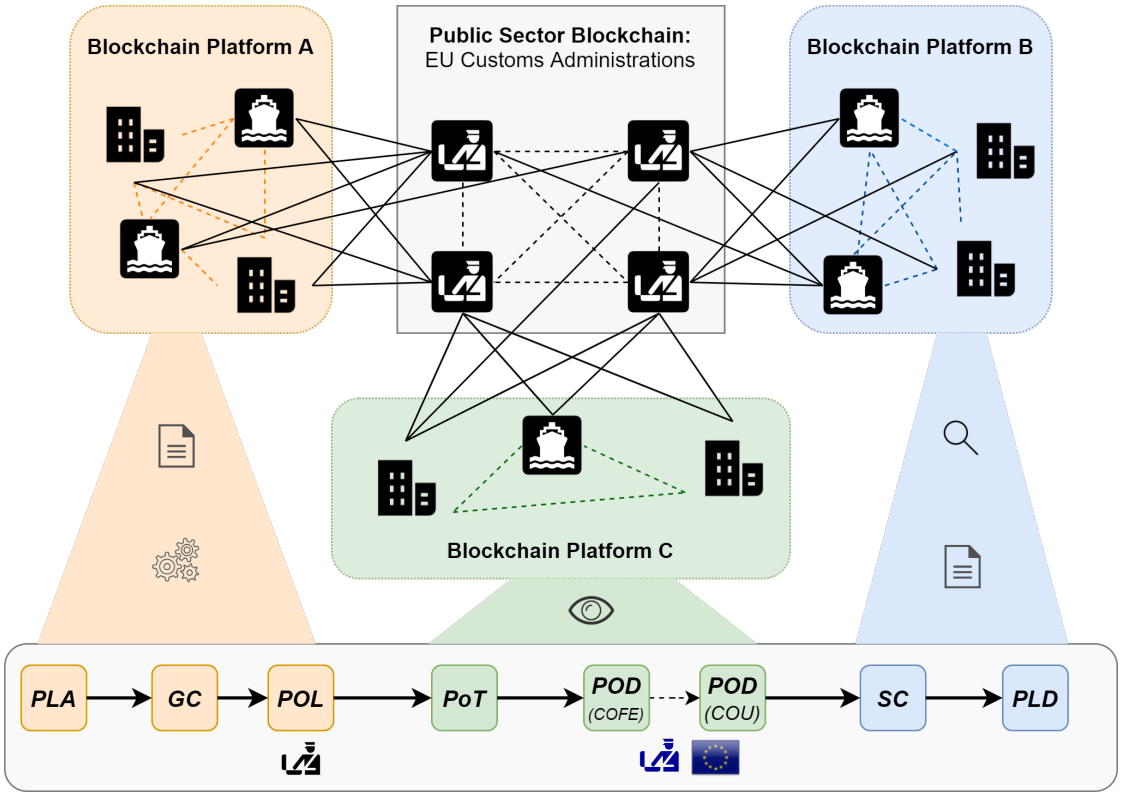


FIGURE 3.1: Fragmentation of the logistic domain by siloed data sources.

The segments of the logistics domain must be integrated by combining data sources. Based on the piggybacking principle and the fragmented nature of supply chain data landscapes, the artefact should be based on a *pull mechanism*, in which data is left at the source and links are created towards relevant sources. In contrast to the traditional *push mechanism* that disseminates copies of data in a network, the artefact would follow the once-only principle and foster single-window environments for communication between enterprises and governments [59]. Besides the practical advantages of the approach, data duplication is also avoided. Therefore, alignment with the data minimisation principle described in European regulation is also achieved [57, 59]. Additional information on the impact of European data protection regulations in the design of blockchain systems can be found in section 3.4.

Non-Functional Requirement 1	*NFR1*
Access to data should be provided by means of a pull mechanism based on links between data sources.	

Besides the access to information, its quality is key for the effectiveness of data pipelines for supply chain visibility [61]. The data requirements included in official deliverables of the PROFILE project [128] are directly linked to this research and should be included in the requirement analysis. These data requirements are formulated at the supply chain level and the logistics level. The supply chain level comprises the actors involved in goods shipping (see section 2.2), and the trade patterns that emerge form the interaction between these actors. These interactions can be described by framework agreements, referring to contract chains with a fixed structure, so there is special interest in the detection of anomalies in trade patterns from declaration data or any other form of documented contractual relations linked to activities carried out in the logistics domain (see Figure 2.1) [128]. This results in the next requirement regarding supply chain actor interaction:

Functional Requirement 2	*FR2*
The artefact should allow to monitor trade framework agreements and detect anomalies in the business transactions between supply chain actors.	

The PROFILE project differentiates between visibility requirements at the logistics level in terms of cargo flows, transport flows and the structure of logistics chains. Cargo flows are built on top of transport leg data, the container tracks and the goods tracks, which are defined as follows by the PROFILE terminology [128]:

- **Transport leg**: transport of particular cargo between two temporally adjacent locations with only one transport means (e.g. POL-POD).
- **Container track**: timed sequence of transport legs for a particular container. A container track can be derived from data provided by the individual transport legs.
- **Cargo track**: timed sequence of transport legs and/or container tracks associated to the transport goods.

Non-Functional Requirement 2	*NFR2*
Cargo flows should be monitored by processing data gathered by the artefact.	

The visibility of individual products and their properties - such as value, quantity and classification - are also considered part of cargo flow data. Individual products are not explicitly tracked in the intermediate segment of most large supply chains, such as the activities carried out by the main carrier in containerised deep sea shipping. However, they become the main reference in the extremes of supply chains, namely from the *PLA* and throughout pre-carriage, and far downstream during on-carriage in their way towards the *PLD*. From the perspective of customs administrations, the information about the provenance of an individual product or the commercial contract behind its shipping can act as clarifying boundary conditions during risk assessments, which can turn ambiguous otherwise.

Similarly to trade patterns, European customs is interested in the detection of anomalies in transport flows: itineraries and routes followed by the means of transport (such as a vessel) for the movement of goods [128]. An itinerary or voyage, is defined as the timed sequence of locations used by a means of transport to load and/or discharge cargo, also known as transshipment locations. A route refers to the description of the infrastructure use (sea, road, etc.) to move between two transshipment locations included in the itinerary.

Non-Functional Requirement 3	*NFR3*
Transport flows should be monitored by processing data gathered by the artefact.	

Besides the main piggybacking pipeline concept of *FR1*, the requirements presented in this section cover two goals: connecting data from different segments of the logistics domain and linking the logistics domain to business transactions. The structure of logistic chains can be built from these business transactions occurring in the logistics domain, which involves the visibility of commercial relations between providers and receivers of services linked to logistic activities, as described by *FR2*.

Customs must link information about the use of transport infrastructure and the execution of business transactions by gathering better physical supply chain data and interpreting contractual documents. This requirement is the core of the need to bridge the gap between supply chain milestones and the stakeholder ecosystems described in chapter 2. The need for an architecture based on links due to compartmentalised supply chains is expressed in *NFR1*. Finally, *NFR2* and *NFR3* cover the need to link cargo and goods tracks to actors by combining the business transactions where cargo and goods are mentioned, so the visibility of both products and freight is important.

3.4 Data Sovereignty Requirements

Blockchains are a double edge sword in terms of data privacy and confidentiality. If their design is centered around node anonymity, then those transactions aimed at verifying regulatory compliance might be ineffective from a legal standpoint, but if their design is driven by open identification and full transparency the infringement of data privacy rights becomes a considerable risk [137]. Additionally, as blockchain's power resides on its unique use of cryptography, European regulators face the arduous task of delimiting a line between data encryption as cyber-security shield and as cover for criminal practices [141].

It is accepted that *DLT* will inevitably transform legal spaces in the long-term [59] and that the work towards innovative legal instruments able to deal with new blockchain applications is valuable. However, due to the importance given to data sovereignty by European

initiatives for supply chain visibility [61] and to avoid speculation, an appropriate approach is the translation of current customs processes into the blockchain domain taking the European legal frames in force as reference. The main sources used for data sovereignty requirements are shown in Figure 3.2, which have been combined with the analysis of academic literature throughout the rest of the section.

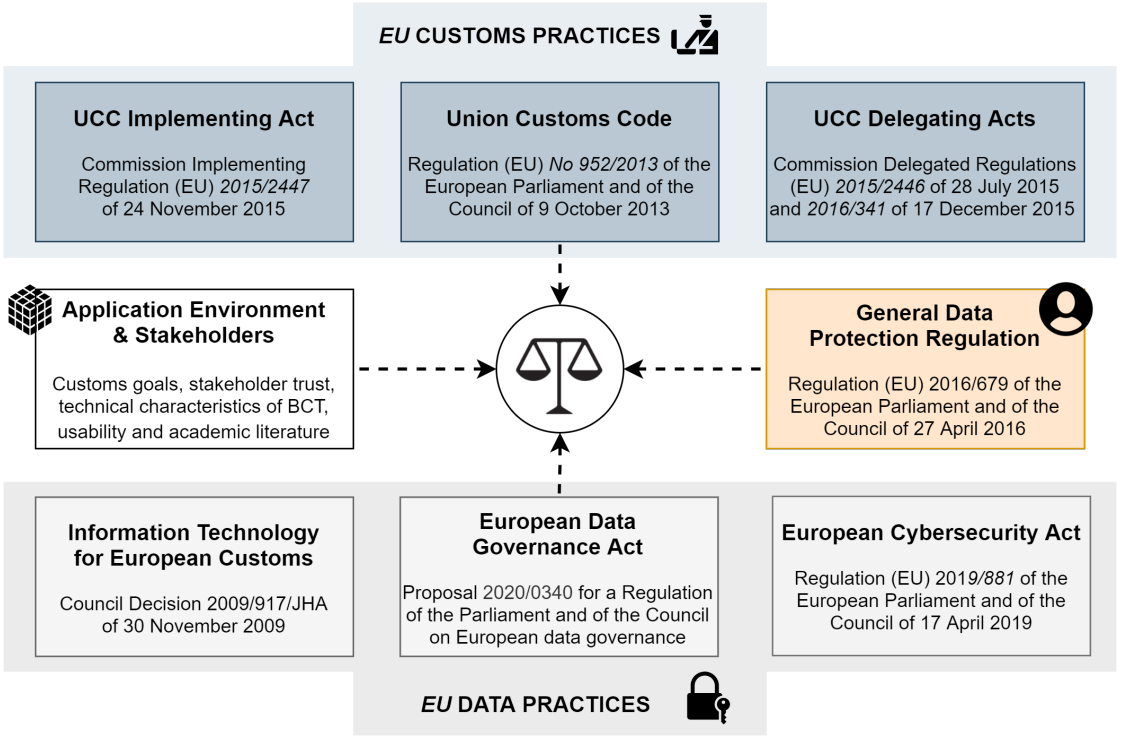


FIGURE 3.2: Sources for data sovereignty requirements.

Regulations to modernise the control of digital data in the *EU* have been recently approved. Some are truly disruptive at inter-industrial level, like the *General Data Protection Regulation* (*GDPR*) [57], others are guidelines for constrained legal spaces [39], while others are still proposals like the *Data Governance Act* [50]. The publication of some of these regulations is previous to the mass adoption of *BCT*, so they might have failed at capturing the peculiarities of data protection in distributed data systems [137]. Still, each of these documents provides a different perspective on data sovereignty, and although not all of them act in practice as enforceable legal constraints in supply chain applications - like the *GDPR* - they can provide valuable context to the research.

An approach to cope with the dual nature of anonymity in a regulatory environment is decoupling the concepts of identification and recognisablility. Anonymity refers to the certainty that action-actor relationships can not be built and thus a link between identity and accountability is unreachable, while recognisablility refers to an actor remaining unidentified unless there are socio-legal reasons to link a specific action to the actor [137]. Although the premise of recognisablility is the *status quo* in most legal spaces, it poses an obstacle in the conciliation of the practical and regulatory value of blockchain systems. Public developers must thus circumnavigate both technical and cultural hindrances behind the dogmatic view of anonymity in the blockchain community [73].

In view of the unclear extent to which new European legislation influences blockchain design, imputability, or the ability of a system to hold actors accountable for their actions through their observable behaviour [137], is still a necessary feature of blockchain systems for regulatory applications. This is partly due the absence of alternative legal strategies [59, 141], but also the relationship between enforceable accountability and the essential value proposition that *DLT* offers, particularly in relation to transport and commerce [61]: consolidated trust in complex inter-organisational settings [137]. The following requirement refers to the functionality of the artefact in terms of the division between anonymity and recognisablility:

Functional Requirement 3	*FR3*
The artefact should allow the enforcement of the recognisability principles implemented in all procedures carried out by European customs administrations.	

Joining blockchain platforms is voluntary and the regulated participation in the proposed data pipeline would resemble more a medium or long term thought experiment than a short term projection of European policy trends. Therefore, the real goal of European customs in this regard is, for now, to transmit the certainty that the enforcement of recognisability can be structurally restricted to customs administrations by the architecture of the artefact. In order to create consolidated trust, it is important that the recognisability requirement holds regardless of the perceived distribution of responsibilities around identification. That is, whether recognisability is interpreted as the obligation of trade actors to remain fairly identifiable or the right of institutions to execute legitimate identification.

Non-Functional Requirement 4	*NFR4*
The enforcement of recognisability must be structurally restricted to European customs administrations by the architecture of the artefact.	

Differentiating between transparency and accessibility is also important for the research context, as these two terms are the source of common misconceptions in the legal interpretation of information systems [137]. Transparency is the inherent ability of being perceived, as in being interpreted so that the purpose of information towards the recipient is fulfilled. Accessibility is on the other hand the ability to obtain information, whether or not the latter may represent any form of actionable intelligence to its holder. The key takeaway is that transparency and accessibility do not always coexist, each of them provides different security advantages and combining them without assessing their compatibility can decrease in practice the utility of information systems. As a result, although security and privacy by design are incorporated in European regulations [51, 57], no more than the necessary privacy and security features should be implemented.

Legislation indicates that European customs information systems must *"ensure that measures are in place for checking the source of data and for protecting data against the risk of unauthorised access, loss, alteration or destruction"* (*Article 3(1)*) [55]. There is also a generalised emphasis in the need for secure storage and transmission of data, particularly for activities related to public interest [39, 50, 58]. From the point of view of accessibility this means that the artefact should restrict the access to information to the person linked to the data or data subject [39], persons acting on behalf of the data subject and customs administrations.

Non-Functional Requirement 5	*NFR5*
The access to data must be exclusive to the data subjects, persons acting on their behalf and customs administrations.	

Also, and complementing *FR3*, the validity of the data to be used for identification must be protected. This means that the identities of users must be certified by trusted identify providers. This works as an incentive for trade actors to share information and for European customs administrations to trust the information gathered through the artefact.

Functional Requirement 4	*FR4*
The artefact should allow the certification of the user identities.	

Data must be intelligible by European customs for its analysis and the enforcement of accountability. Therefore, mechanisms to ensure customs can access useful information are needed regardless of the level of data transparency at different stages of the pipeline. However, measures to reduce transparency are not always effective, because behaviour patterns can always be reconstructed with the semantic analysis of transaction metadata and identities can be reverse-engineered from immutable ledger records [139].

The limitations of encryption are covered in European regulation. In fact, the *GDPR* classifies any type of encryption as pseudonymisation [57]. This includes both the hashing and salted-hashing used in blockchains [52]. Encryption keys can, in theory, always link data to actors, so it can not be considered an anonymisation technique. However, the security benefits that data hashing provides in practice should outweigh the infeasible *"possibility that may lead to a data subject's personal data on a blockchain to be linked to its identity"* (pp. 1239) [10], particularly in permissioned blockchains. This does not imply that European customs should not safeguard the security standards of the artefact with additional security measures. Therefore, the right combination of data access control and data transparency is required to ensure that data is reachable and interpretable only by its lawful recipients.

Non-Functional Requirement 6	*NFR6*
The interpretation of data must be exclusive to the data subjects, persons acting on their behalf and customs administrations.	

Information systems are used by European customs for status applications (see *AEO* in section 2.5), communicating decisions, amendments to lodged information, and processing notifications like *do not load* warnings [54, 55]. In these processes economic operators can exercise their right of appeal *"against any decision taken by the customs authorities relating to the application of the customs legislation which concerns him or her directly and individually"* (*Article 44(1)*) [53]. Unless there are reasons to believe that it would cause irreparable damage to the person concerned, appeals do not imply the suspension of the disputed decision in order to prevent the use of appeals to interfere in legitimate investigations [53].

Additionally, the *Union Customs Code Implementing Act* stipulates that *"each input, modification and deletion of data shall be recorded together with information giving the reason for, and exact time of, such processing and identifying the person who carried it out"* (*Article 3(2)*) [55]. This means that the artefact should also produce a record of amendments, their authors and rationale.

Non-Functional Requirement 7	*NFR7*
Amendments to data performed through the artefact, their motivation and authors must become traceable by customs administrations.	

Another barrier to overcome during the deployment of data pipelines connecting organisations and data sources is the right to erasure, also known as the *right to be forgotten* within the GDPR [57]. In the context of European customs, it represents the right to demand the revocation of data access to customs administrations if the data is not longer needed in relation to the original purposes for which it was collected and/or processed [39].

For example, the *Union Customs Code Implementing Act* specifies time limits for the handling of data records of inactive economic operators after which *"the competent authority of a beneficiary country or the customs authorities of the Member State shall delete the data"* (*Article 89(10)*) [55]. Technical solutions need to be contemplated to allow European customs to comply with the right to erasure while leveraging the advantages of immutable records stored in ledgers [137]:

Non-Functional Requirement 8	*NFR8*
Exercising the right to erasure must be enabled to the data subjects of the data collected, processed and stored by the artefact.	

3.5 Architecture Interoperability Requirements

The utility of most blockchain applications is fragmented, meaning they have been developed to serve very specific value chains and to create real impact in reduced business or industry contexts [120]. The rise in blockchain popularity has been followed by an increase in cross-chain transaction demand, and as a result, an increase in blockchain standard incompatibility and asymmetric access to information in general [88, 120]. Interoperability is thus a key element of any large-scale blockchain project, such as the data pipeline for *TTL's* envisioned by European institutions.

In general, architecture interoperability can be defined as the capacity to exchange and use information between computer systems and to share digital assets between blockchain networks while preserving the state of unicity of the assets [94, 187]. The goal of this research is not to solve the universal ledger compatibility problem, but it intends to throw some light on architecture interoperability for trade and supply chain applications and the interaction between businesses and governments. The proposed blockchain data pipeline enables customs to integrate commercial platforms that use different blockchain technologies in their declaration activities. This is expressed in the following requirement:

Functional Requirement 5	*FR5*
The artefact should allow European customs administrations to reuse information produced and/or stored by different blockchain protocols.	

As mentioned before, blockchain architecture interoperability can be interpreted in two ways: the exchange of digital assets and the exchange of arbitrary data. The exchange of digital assets refers to the ability to transfer a digital asset that originates from different blockchains [187]. A possible example for the research context is interpreting data from an *eB/L* created in an *Hyperledger* platform used at export by the shipper and carrier by a *Corda* platform used at import by European customs. This type of interoperability is a large driver of document piggybacking and can be summarised in the next requirement:

Functional Requirement 6	*FR6*
The artefact should support the exchange of digital assets between platforms.	

An attempt to foster piggybacking for customs procedures can be seen in the European regulation. As specified in the *Union Customs Code Implementing Act*, more than one data source can be used to submit, update or amend an *ENS* in order to *"take into account the cases where certain particulars of the entry summary declaration are to be submitted at an early stage in the transport of goods to allow for better protection against serious threats and also the cases where, in addition to the carrier, other persons submit particulars of the entry summary declaration to improve the effectiveness of risk analysis for security and safety purposes"* (pp. 561) [55]. What this implies for architecture is expressed in the following requirement:

Non-Functional Requirement 9	*NFR9*
The digital assets processed by the artefact should be accessible and editable by a number of users related to the data subject.	

On the other hand, the exchange of arbitrary data represents a deeper interaction between blockchain protocols, where the consensus of transaction events in a blockchain platform can be verified by smart contracts implemented in another platform without the need to transfer signed copies of the digital assets involved [120, 187]. An example for the research context is the automation of customs clearance based on the transfer of ownership of an *eB/L* across the logistics domain. The need for this is expressed in the following requirement:

Non-Functional Requirement 10	*NFR10*
The artefact must include features to exchange arbitrary events between platforms.	

These cross-chain transactions can be seen as a set of functionalities that are triggered by one blockchain platform so that another platform can execute an operation in its network [123]. An example of such arbitrary data exchange is the authenticity verification for information transferred between two different blockchains, which can be useful for European customs in terms of origin certificates and the decentralised cross-validation of trade data in general.

The key for interoperability is cross-chain communication while ensuring validity and verifiability of transactions [2], meaning consensus is kept in both ledgers and that event traceability is maintained. Efficient consensus algorithms for a single blockchain can be challenging, but it becomes a genuine wicked problem in the context of universal ledger compatibility. Nevertheless, interoperability frameworks is a growing research field [120], and the standardization of ledger protocols are desired by industry leaders and regulators [175,

200]. Consensus plays thus a main role in architecture interoperability [120], and the artefact is must provide European customs with consensus and verifiability control. This is summarised in the following requirement:

Non-Functional Requirement 11	*NFR11*
The cross-platform communication performed through the artefact must preserve consistency between the consensus protocols of different platforms.	

The architectures of different blockchain technologies act as the skeleton of blockchain platforms and are diverse, as well as their mechanisms to interact with external data sources. Additional functionalities can be achieved with the right interoperability infrastructure, but this should not entail the transformation of the architecture of its potential participants. Affecting their stand-alone operability can be perceived as a risk by users and operators, who will avoid putting in jeopardy the stability of their internal ecosystems.

Non-Functional Requirement 12	*NFR12*
Modifications to the internal specifications and functionalities of a commercial platform must not be required in order to interact with the artefact.	

3.6 Conclusion & Overview of Requirements

In this chapter, the design requirements that best suit the expected functionality of the data sharing architecture have been covered. This has been done by answering *RQ2*: *What are the design requirements to preserve the data sovereignty of supply chain actors when creating links to data stored in multiple ledgers?*, and *RQ3*: *What are the design requirements to allow multiple blockchain platforms to share interoperable links to their ledger states?*.

Three design principles are chosen for the requirement generation. The first principle, *logistic event visibility*, is the essence of the practical value of the design for customs administrations. The two remaining design principles are two of the accompanying implementation challenges addressed in the research: *stakeholder data sovereignty* and *architecture interoperability*. The requirements linked to *DP2* intend to answer *RQ2*, while those linked to *DP3* intend to answer *RQ3*.

The exploratory nature of the research implies that all research phases are performed at a conceptual level. The high level review of information sharing trends in the supply chain industry presented in chapter 2 has led to the generation of high level requirements. The latter are used in the next phase for the selection of components to describe a conceptual data sharing architecture. As a result, certain technical nuances of the application domain, such as detailed features of blockchain platforms and document formats, are misrepresented in this chapter. Additionally, the completeness of the requirements is limited by the fact that European customs administrations have been identified as main user. The functional requirements represent what is needed from the architecture by customs administration in order to successfully exercise their institutional duties. The user requirements of other stakeholders, such the blockchain-platforms themselves, have been implicitly included in the non-functional requirements. Therefore, the analysis of implementation drivers that capture the needs of all users when putting the conceptual architecture into practice (*e.g.*, development costs, maintenance or governance) have not been included in the requirement generation.

An overview of the functional requirements is shown in Table 3.2 and the Table 3.3. These requirements are used as input in chapter 4 during the design phase.

TABLE 3.2: Overview of functional requirements.

code	*description*
	DP1: Logistic Event Visibility
FR1	The artefact should allow European customs to access information produced upstream supply chains entering the European Union
FR2	The artefact should allow to monitor trade framework agreements and detect anomalies in the business transactions between supply chain actors
	DP2: Stakeholder Data Sovereignty
FR3	The artefact should allow the enforcement of the recognisability principles implemented in all procedures carried out by European customs administrations
FR4	The artefact should allow the certification of the user identities
	DP3: Architecture Interoperability
FR5	The artefact should allow European customs administrations to reuse information produced and/or stored by different blockchain protocols
FR6	The artefact should support the exchange of digital assets between platforms

TABLE 3.3: Overview of non-functional requirements.

code	*description*
	DP1: Logistic Event Visibility
NFR1	Access to data should be provided by means of a pull mechanism based on links between data sources
NFR2	Cargo flows should be monitored by processing data gathered by the artefact
NFR3	Transport flows should be monitored by processing data gathered by the artefact
	DP2: Stakeholder Data Sovereignty
NFR4	The enforcement of recognisability must be structurally restricted to European customs administrations by the architecture of the artefact
NFR5	The access to data must be exclusive to the data subjects, persons acting on their behalf and customs administrations
NFR6	The interpretation of data must be exclusive to the data subjects, persons acting on their behalf and customs administrations
NFR7	Amendments to data performed through the artefact, their motivation and authors must become traceable by customs administrations
NFR8	Exercising the right to erasure must be enabled to the data subjects of the data collected, processed and stored by the artefact
	DP3: Architecture Interoperability
NFR9	The digital assets processed by the artefact should be accessible and editable by a number of users related to the data subject
NFR10	The artefact must include features to exchange arbitrary events between platforms
NFR11	The cross-platform communication performed through the artefact must preserve consistency between the consensus protocols of different platforms
NFR12	Modifications to the internal specifications and functionalities of a commercial platform must not be required in order to interact with the artefact

Chapter 4

Design & Development

The purpose of this chapter is to generate prescriptive knowledge through the specification of artefact components, their functionalities and how they provide added value to the problem at hand [90]. The chapter covers the *Design & Development* phase of the design science approach, where the knowledge gathered during the previous chapters is combined with the contributions of this study in the form of a data sharing architecture. Overall, it includes the design rationale behind the decisions taken to design an architecture compliant with the requirements generated in chapter 3 and able to tackle the challenges described in chapter 2. The work presented aims to ultimately answer the following research question:

RQ4: *What architecture components can used by customs administrations to gather declaration data stored in multiple commercial blockchains?*

The structure of the chapter is as follows. An overview of the three layers included in the architecture - *cross-chain communication*, *credential management* and *event visibility* - is covered in section 4.1. This section also covers the conceptual design approach, the components of each layer and an explanation of the practical context in which the data sharing architecture will be implemented.

The detailed layer design is presented individually in the next three sections. The *event visibility* layer is covered in section 4.2. It includes an approach to help customs administrations leverage new *DLT*s to improve the visibility of the commercial and logistic patterns produced by the supply chains entering the *EU*. The *cross-chain communication* layer is covered in section 4.3 and focuses on the design of a network to exchange logistic event information. Here an interoperable approach to verify and exchange ledger states is included. Lastly, the *credential management* layer is covered in section 4.4. It discusses the link between the *cross-chain communication* layer and the chosen access and identity control components. This section focuses on an interpretation of confidential, secure, self-sovereign identities in the logistics domain and why they are needed to promote trust in the researched ecosystem of blockchain platforms.

The chapter ends with a summary of design decisions taken throughout the development of the architecture in section 4.5. The section also links each step of the design to the design principles and requirements that motivated their implementation.

4.1 Architecture Overview & Functional Context

The design and development activities start with a reflection on the answer given to *RQ3* in chapter 3. The relationship between the expected functionalities of the artefact and each design principle (see Table 3.1) are condensed in Table 4.1. The grey diagonal represents the functionalities generated by each design principle, while the white cells represent how they overlap. This research phase focuses on the technologies and components able to bridge the research gaps discussed in section 1.3 by exploring these overlap areas.

TABLE 4.1: Functionalities across design principles.

	Event Visibility	Data Sovereignty	Architecture Interoperability
Event Visibility	Mapping events into the logistic domain from commercial transaction data	Monitoring accountability while preserving commercial confidentiality	Exchanging ledger states between blockchain platforms
Data Sovereignty		Enabling entities to enforce rules for resource and data ownership	Authenticating the identities of platform participants
Architecture Interoperability			Establishing ledger-agnostic peer-to-peer connections

After analysing the research context and the design requirements from a bird's eye view, the problem is best described by the *service choreography* perspective:*"describing the sequence and conditions in which data is exchanged between two or more participants in order to meet some useful purpose"* [24]. Understanding a service as the logical representation of a repeatable task that is linked to a business activity, the *service choreography* concept arises from the need to model behavior in multi-actor systems based on data sharing patterns [158].

The information flows encountered in the transport industry are examples of such systems, in which the study of choreographed data sharing has produced architectures rooted on *service orchestration*: using standardised protocols and high-level languages to integrate architectures, and in doing so, achieve stronger relationships between businesses and their information systems [33, 71]. Figure 4.1 shows the essence behind the approach, which is finding links between the application and business logic that support these services. The design must capture the interpretation of design requirements from three process planes [70]: the orchestration plane (industry practices and process conventions), the value plane (stakeholder interactions) and a mechanical plane (technical protocols).

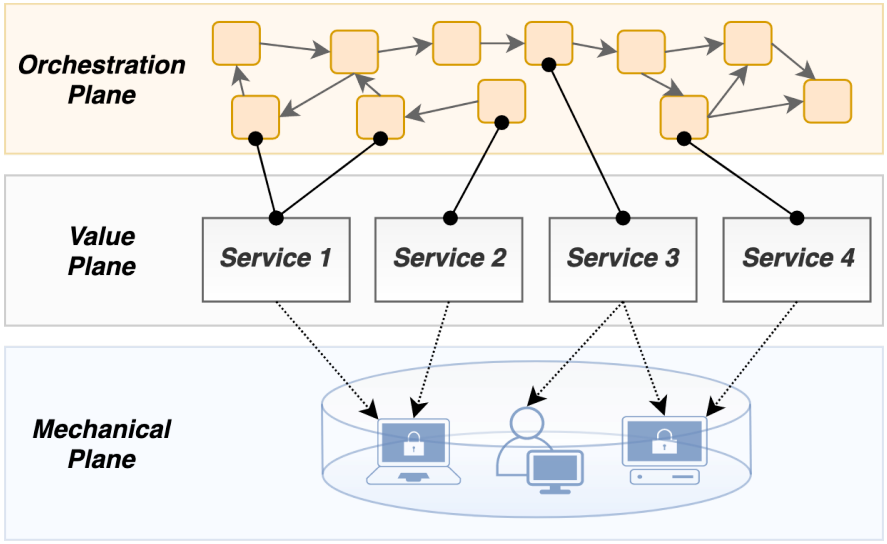


FIGURE 4.1: Orchestration of services, adapted [71].

From a process-oriented perspective orchestration can be very effective in reducing friction caused by semantic heterogeneity (see subsection 4.4.4) and making collaboration between supply chain actors easier. However, it is not sufficient to orchestrate services with a purely event-driven approach. The context in which services and transactions are executed must be also taken into consideration in the design, which can be done with an additional domain-based framework able to represent the underlying entity interactions in more detail.

Given the need to synchronise information flows between permissioned domains, the most appropriate design approach is that of a data sharing architecture in which transaction queries and index data are somehow captured [157]. The example shown in Figure 4.2 provides domain interoperability by using a common data repository. It stores links between pieces of private data so that information produced in one domain can propagate application logic in other domains. This way, supply chain actors can operate following internal domain policies while allowing their data to feed otherwise isolated services.

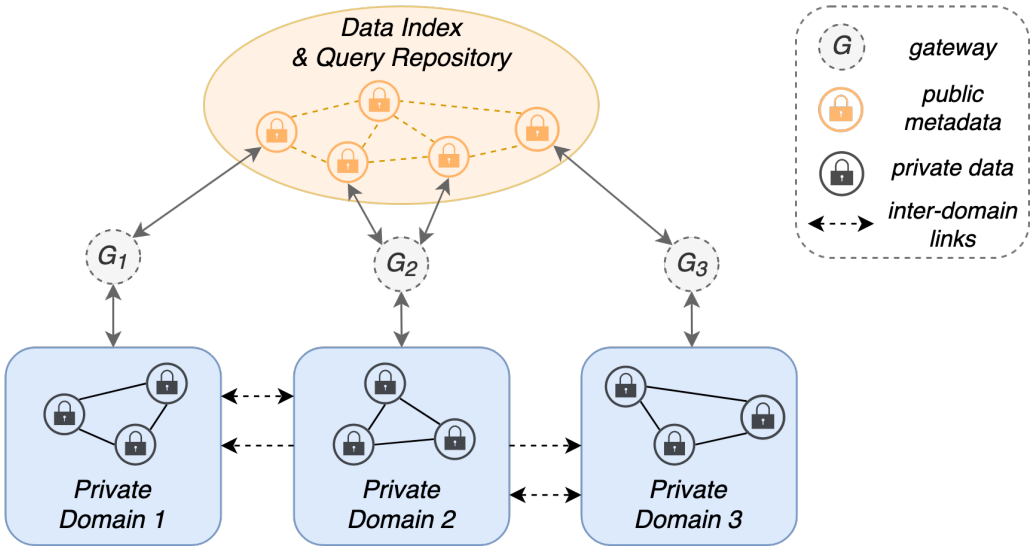


FIGURE 4.2: Conceptual data sharing model, adapted [157].

Similarly to the development of *vigilant information systems* [182] (see section 3.2), this data model was originally proposed in the context of ontology interoperability to link business concepts to network management data between organisations [156]. Different versions of this concept have been adapted to facilitate ontology integration [116], the implementation of context-aware data dissemination policies [157], *Linked Open Data* (*LOD*) for *B2G* communication [49, 77, 129] and blockchain-based supply chain monitoring [132].

Combining the two approaches is considered appropriate for the research due to the close relation between the use cases proposed in literature and the transition towards system-based regulatory strategies covered in section 2.5. However, it is needed to explore the characteristics of the researched application in detail to assess the need for additional functionalities. This includes defining a gateway design, a service orchestration mechanism, compatibility standards for the information produced in the private domains and the role digital identities play in complying with data sovereignty requirements. That work is reserved for the remaining sections of the chapter.

The framework of Figure 1.2 is applied to arrive at these design decisions. Scientific rigour is prescribed by knowledge extracted from literature, while relevance is ensured by business drivers and constraints captured in the requirements. The outcome of this process is summarised in Table 4.2 with an overview of design components. The grey cells are the type of design components needed by each design principle, and can be interpreted as an initial iteration of solutions found in literature. The white cells are specifications that emerge when the design principles are compared. They are a second iteration of more detailed solutions after combining the first iteration with the requirements of the researched problem.

TABLE 4.2: Link between components and design principles.

	Event Visibility	**Data Sovereignty**	**Architecture Interoperability**
Event Visibility	Directed acyclic graph ledger for distributed logic registry	Accumulators compatible with key rotation certificate revocation	Cross-chain resource protocol based on the Hash-lock paradigm
Data Sovereignty		Self-certifying identifiers based on the *DID* core data model	Pub-Sub model for gateway discovery and interaction
Architecture Interoperability			Overlay network for cross-platform gateway bridges

The analysis is translated into a conceptual design shown in Figure 4.3. A shared network infrastructure acts as bridge between nodes of a meshed ledger in which distributed applications are maintained. Each node mirror the participant of an external blockchain platform. They can propagate the logic of their original ledgers by sharing a specific view of the meshed ledger. An intermediate layer with accumulators supports key rotation and certificate revocation filters the access to certain parts of the distributed application.

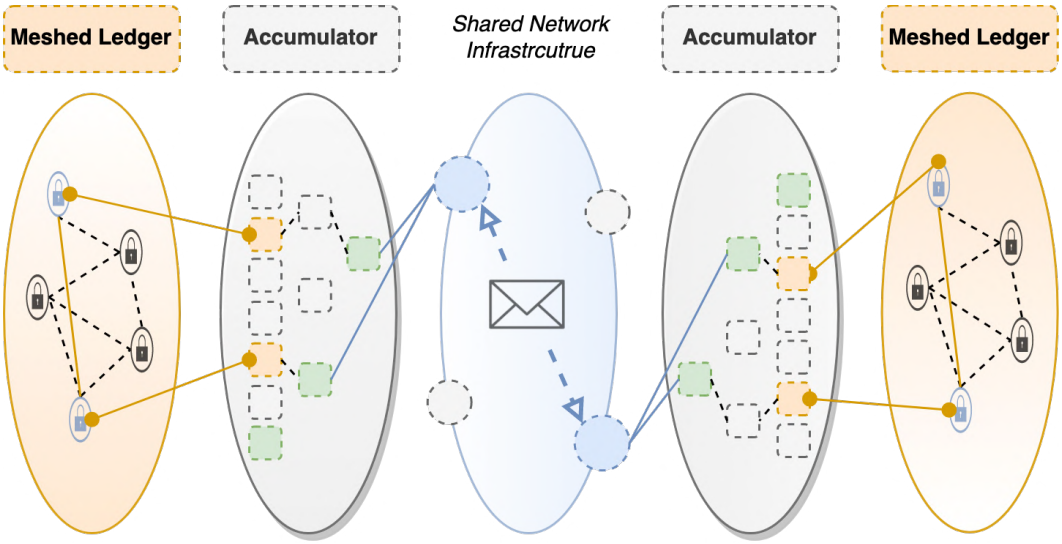


FIGURE 4.3: Conceptual architecture design.

The final iteration is shown in Figure 4.4. It is composed of three layers: communication, credential management and transaction visibility. An independent view of the two lower layers is maintained locally by each participant, while the communication layer represents the infrastructure used to bridge two isolated networks. The main functionality of the architecture is providing a digital space where combinations of processes particularly relevant for customs risk assessments can be partially captured to avoid information redundancy and foster data piggybacking. Instead of issuing documents with sequential duplication, links between digital resources belonging to the same logistic and commercial chains can be constructed. This new approach presents the junction between the dependencies of traditionally siloed supply chain data exchanges.

The application context for the architecture is shown in the lower segment of Figure 4.4: a group of supply chain actors, which can include both logistic partners and commerce regulators, taking part in independent data streams to generate logistic documents. Figure 3.1 already presented such a fragmented ecosystem of data sources, as groups of supply chain actors operate in an increasing number of digital platforms powered by *BCT*. These platforms have reduced the friction of data sharing processes in the private sector (see section 2.3 and subsection 2.4.3).

Following this transformation, the *EU* recognises the opportunity to gain competitive advantage in international trade and commerce by also reducing the institutional friction of import and export activities. In this context, the architecture presented in this chapter can help European customs update their digital infrastructure strategy to collect declaration data in a new era of information sharing.

The cross-chain communication layer is formed by an overlay network. Each participant can take part in the network through a dedicated gateway. A cross-chain resource transfer protocol based on the 2-phase commit paradigm [15, 70] is implemented to establish peer-to-peer connections through these gateways. It is the core representation of the data pipeline concept discussed in Figure 2.9 and represented as the *Data & Document Layer* in Figure 2.9. Its ultimate goal is to exchange ledger states containing verifiable presentations (see subsection 4.4.1) of non-fungible resources such an *eB/L* (see Figure 2.4) or a *L/C* (Figure 2.7).

The credential management layer handles the credential authentication functions for access control. In essence it acts as a *DPKI* (subsection 4.4.2) that leverages the advantages of asynchronous accumulators to ensure identity authentication between changing members of different blockchain platforms. The identity control is based on a self-certifying identifier model that allows entities to govern their data sovereignty and develop trust in the transactions powered by the resource transfer protocol. This way, confidentiality is maintained throughout the entire data collection process. Also, it enables data subjects to define effective data protection policies while complying with European legislation.

The transaction visibility layer enables cross-chain auditability by building directed graphs between the ledger states involved in decentralised applications. Since ledger states can represent a digital resource, arbitrary data or the confirmation of a process being completed, all type of associations between entities, assets and commercial information can be *orchestrated*. Therefore, customs administrations can gain visibility of end-to-end trade activities with a meshed registry of cross-chain transactions built around the single-window principle. Moreover, its tamper-evident design makes it possible to perform retroactive investigations when required and reduces the incompatibilities that arise between the *ENS* declaration level freedom and data dissemination between customs (see Figure 2.6).

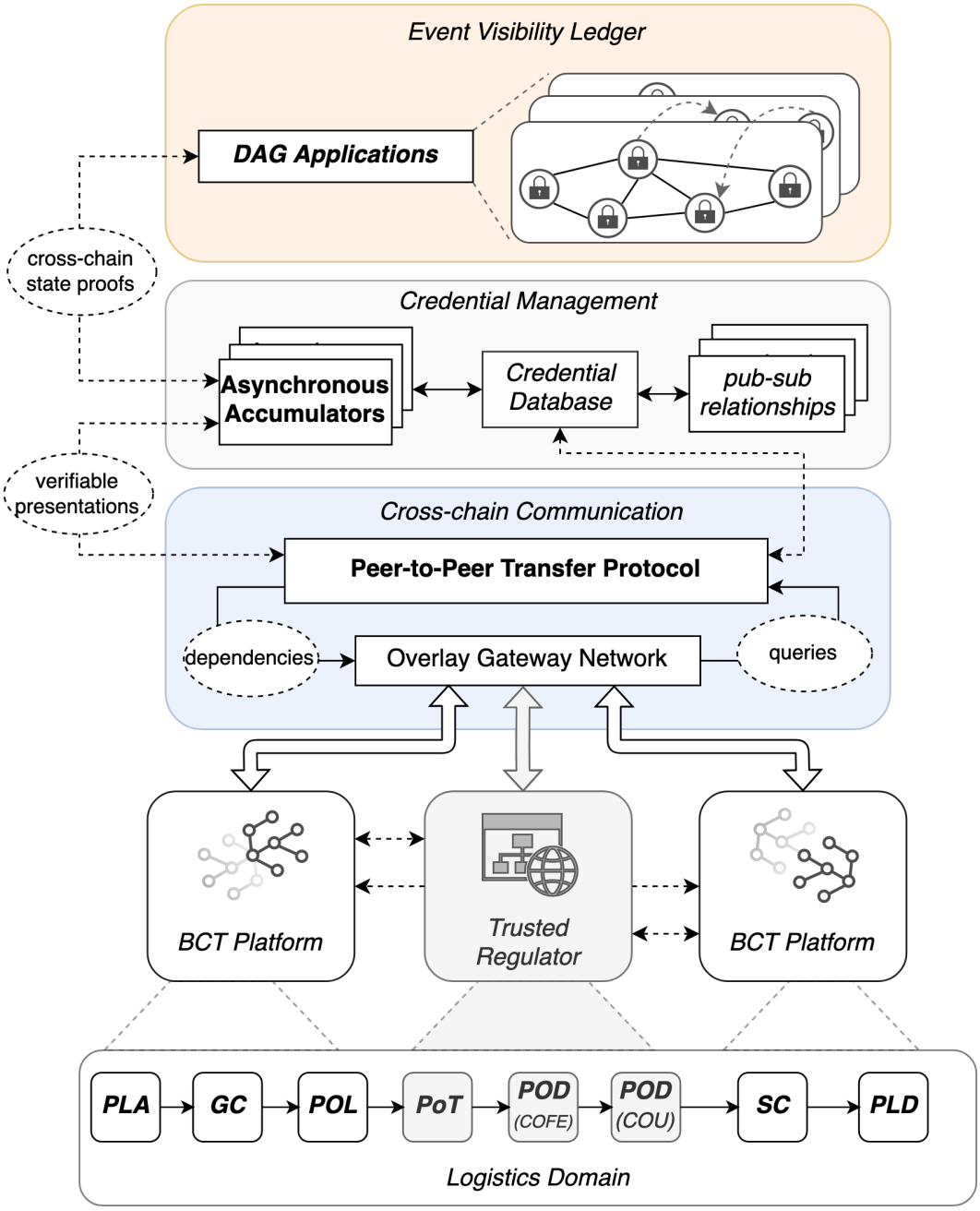


FIGURE 4.4: Detailed architecture design.

4.2 Transaction Visibility Layer

This section covers the components of the transaction visibility layer: a semantic event model for the application of verifiable claims and credentials in the logistics domain and a transaction registry to increase supply chain visibility. The goal of the registry is to ensure transaction visibility and accountability with tamper-evident registries of those transactions that include or refer to a logistic event.

4.2.1 Semantic Model for Event Claims

The semantic event model presented in this section will be applied to the verifiable credentials and the *RDF* claim model presented in subsection 4.4.3. It will classify credential properties and create links between their *subjects* and *values* for logistic visibility. By combining the data in the transactions with the links between them, customs administrations can reconstruct trade patterns. Therefore, the semantic model presented in this section is strongly linked to *DP1* and the logistic event visibility requirements.

The data exchange between supply chain stakeholders and customs administrations has been covered in chapter 2, where a constant reference to events has been made. The PROFILE project studies how to use data pipelines to transform external data sets into reusable data sources for customs data piggybacking (see section 2.5). To this end, a semantic framework that fulfils PROFILE's visibility requirements at logistics level, *i.e.*, *NFR2* and *NFR3* in section 3.3, is used. The main elements of this framework are shown in Figure 4.5. It was created by FEDeRATED [60] and is employed by other European initiatives, such as the Digital Transport & Logistics Forum [166].

Customs administrations can produce logistic event records by mapping acquired data into the semantic categories of the framework. Besides regulatory supervision, this categories can be used by the trade actors to construct traceable information paths necessary to verify the contractual obligations between other actors, such as the contracts of carriage behind a *B/L* or the underlying *L/C* and contracts of sales. The following paragraphs describe the categories of the semantic event model [60].

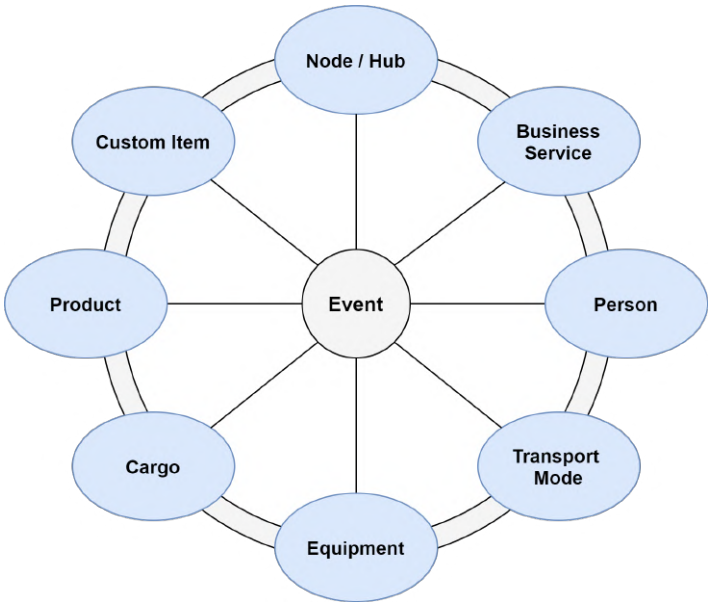


FIGURE 4.5: Semantic event framework [60].

Node / Hub: refers to the physical locations in which transport activities are performed. Cargo loading, customs inspections or temporary storage during pre-carriage must be linked to locations, which can be done through coordinates or a system of facility codes. These nodes are the discrete milestones of the logistics domain, such as the *POL* or the *PLA*.

Business Service: represents the commercial transactions between supply chain actors. The structure of business transactions should follow the rules of the organisational layer discussed in Figure 2.9, like the exchange of *MB/L's* between shipping lines and freight forwarders or the fund transfers during the commercial release of goods (see subsection 2.4.3).

Person: refers to any individual directly involved in the movement of goods, as in those liable for certain procedures throughout the cargo custody chain. Examples include the vessel captain usually responsible for signing a *MB/L* as agent of the carrier [174], or truck drivers taking part in the verification of container dispatch at port facilities [61].

Transport Mode: the vehicles that transports the goods. These can include trucks, vessels, ships or barges. The tracking of the vehicles has benefits in terms of operational monitoring, such as increased accuracy in arrival estimations. However, fleet tracking can provide useful information during customs investigations, the execution of smart contracts or in the resolution of disputes between stakeholders.

Equipment: represents the machinery or physical assets used during the handling of goods. They can be interpreted as sub-nodes inside nodes (the first category discussed) in which the internal node activities are carried out. For example, the cranes a container passes through in large port terminals. However, in more sophisticated representations of the cargo handling between nodes, a sequence of locations could be inferred from the information provided at equipment level.

Cargo: the goods being transported form the *PLA* to the *PLD*. These normally take the form of consolidated packaging in containers, meaning that individual items are not described at its lowest level throughout the whole logistics domain nor by all stakeholders.

Product: in contrast to consolidated cargo (combined parcels packed in larger units) that might be referenced in a *MB/L*, a product represents the actual object or the group of objects that are described in the original contract of sales between a buyer and a seller.

Custom Item: formally defined in the framework as *"a sort operation on HS-codes (or another applicable customs classification) of cargo (incoming/transit) or products (import/export)."* [60]. The term *HS-codes* refers to the Harmonised System codes. This system is used by authorities to assign tariff rules to different categories of products [54]. For example, agriculture commodities and computer hardware do not follow the same valuation rules. A similar code is also used to describe the packaging of the cargo, such as the shape and material of its containing recipient [56]. In combination with the HS-code and the type of cargo movement, such as import/export or incoming/transit, customs can perform appropriate risks assessments and the internal data-sharing covered in subsection 2.4.1.

As discussed in subsection 2.4.3, private stakeholders and customs administrations are interested in the correct categorization of information handled during supply chain data exchanges. In that sense, the framework offers a complete overview of the necessary information types that should be processed by the artefact to comply with *FR2*, *NFR2* and *NFR3*.

Additional categories could be included explicitly in the framework, such as cargo custody chains or the physical itinerary of goods. Specifying such categories is not necessary, as combinations of the other categories allows customs administrations to form the necessary semantic constructs for their internal analyses. For example, cargo custody chain claims can be generated combining *Node*, *Person* and *Cargo subjects* and *values*. A comprehensive set of relationships to form claims as proposed by FEDeRATED [61] is shown in Figure 4.6.

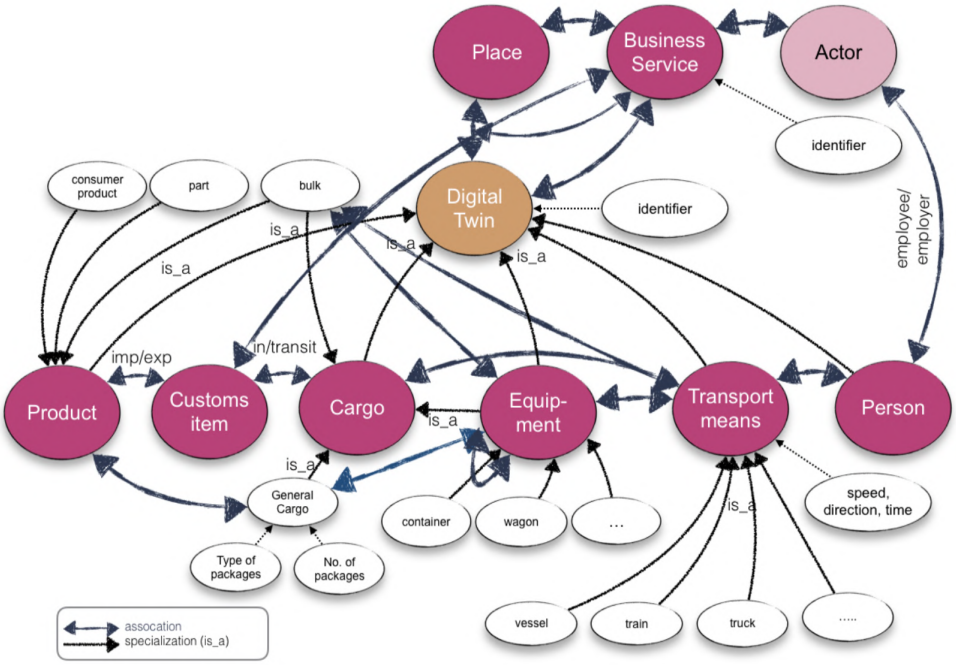


FIGURE 4.6: An ontology for supply chain visibility [61].

Different claims (events) can be generated. For the purpose of the research, a distinction can be made between commercial claims and logistic claims. Commercial claims represent coordination between supply chain actors, commonly protected through legal agreements such as the terms of a *L/C* for trade finance or a *B/L* for contracts of carriage. On the other hand, logistic claims represent physical milestones in the logistics domain (see Figure 2.1). The utility of expressing claims using triples (see subsection 4.4.3) is registering transaction attributes to extract the underlying business logic behind the movement of goods. However, this is not possible if the claims that form an end-to-end cargo custody chain are not linked, *i.e.*, if it is not possible to create a verifiable network of claims. This might be due to three reasons, which are avoided by combining the architecture layers.

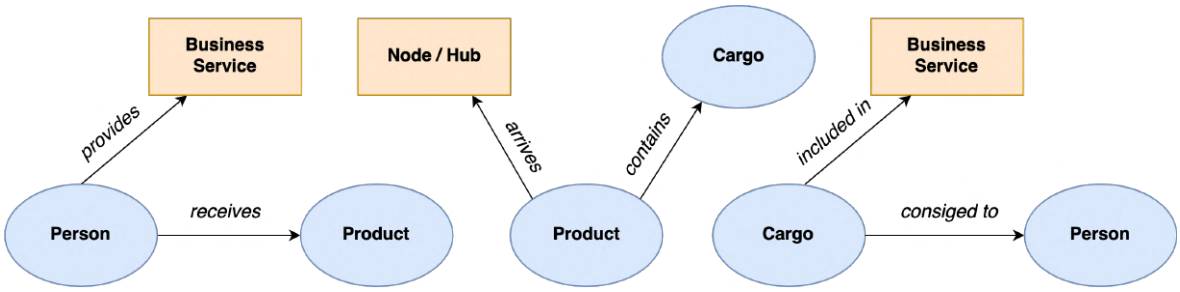


FIGURE 4.7: Unconnected commercial and logistic claims.

The first reason is the lack of claims in a logistics domain segment. Second, creating the required verifiable network might entail trespassing confidentiality rights between logistic service providers (data sovereignty barriers). Lastly, the inability to reach consensus due to the abundance of contradictory claims or incompatibility between claim semantics, meaning it is technically infeasible to connect claims although they exist (interoperability barriers). An example of these scenarios is shown by Figure 4.7, where three unconnected claims might describe individual portions of the movement of a product using different credential models. This has been discussed in chapter 2 to be hindering the effectiveness of risk assessments. Also, it is desirable to understand the connection between these claims in a timely and reliable manner from the point of view of public institutions . However, if this is achieved in practice the process used is inefficient.

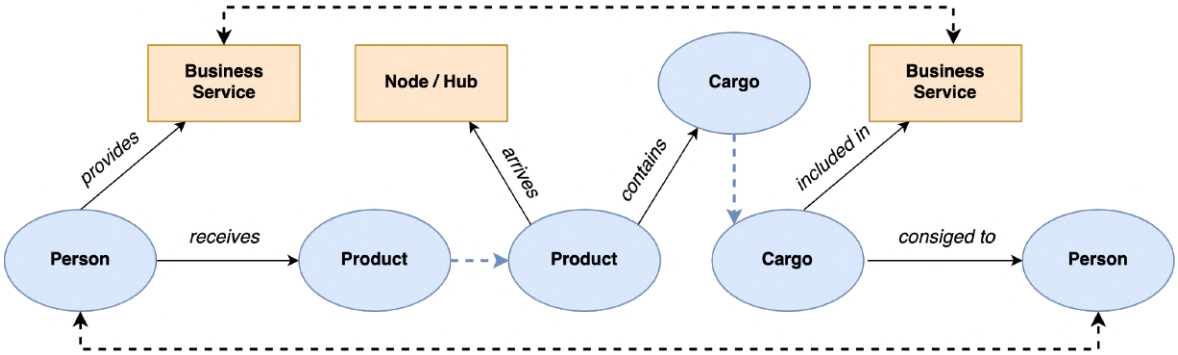


FIGURE 4.8: Improving links between commercial and logistic claims.

The customs declaration process can be simplified through data piggybacking if the links shown in Figure 4.8 as blue dashed arrows are built. This way customs administrations are able to construct better trade patterns shown in Figure 4.8 as black dashed arrows. In order to achieve this, section 4.2.2 introduces a presentation and credential registry that allows to construct these strategic links between logistic and commercial claims.

4.2.2 DAG Applications

This section presents the *DLT* selection for the decentralised transaction registries. It includes a conceptual *directed acyclic graph* (*DAG*) design to build parallel ledgers following the information graph approach of subsection 4.2.1 and a detailed graph model to illustrate the connection between the ledger state proofs generated by different blockchain platforms.

Selection of Ledger Technology

Mathematically, a *directed acyclic graph* is a finite set of nodes connected by unidirectional edges where no directed cycles exist [80], meaning that feedback loops cannot be generated. In this approach verifiable presentations are not bundled and stored in blocks, thus its nickname *block-less* ledger. Instead, they are linked directly between each other in a network of ledger states binned together by similar cryptographic techniques used in traditional blockchains. There has been a growing interest in modelling distributed information systems in terms of *DAG*s. This technology is considered the next iteration in *BCT* [80], being sometimes referred to as *Blockchain 3.0* [27], and is particularly promising for use in permissioned ledger interoperability and *Internet of Things* (*IoT*) [185, 104, 197]. Also, *DAG* technology is an interesting option to process and organise cross-chain transactions, acting as an independent reference to validate states between ledgers while allowing third parties to act as auditors (such as customs or any other regulator) [25].

There is a strong conceptual relation between *DAG* ledgers, asynchronous ledger consensus and the dynamic membership proofs required for credential management (see subsection 4.4.6). The use of *DAG* is an attempt to combine the benefits of these concepts to improve interoperability with decentralised logic. Moreover, it is ideal for the representation of logistic events of Figure 4.8 using *RDF* statements covered in subsection 4.4.1. The design choice is also motivated by the numerous *DAG* protocols [9, 11, 35, 197, 198] aimed at solving scalability and interoperability barriers [80, 185, 189, 104].

Currently, most *DAG* protocols are designed for public ledgers [79, 189], but the architecture will be used in the environment shown in Figure 3.1. This is an ecosystem of private, permissioned and consortium blockchain platforms. Therefore, the application of *DAG* for the interaction between permissioned ledger states is a novel design approach. *DAG* technology has limitations, mainly due to the lack of research on its application for large scale systems, although its potential to allow new use cases is large and is slowly becoming a prominent area of research [93].

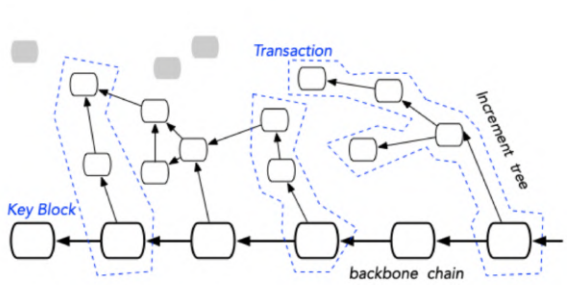


FIGURE 4.9: Convergent topology: Haootia protocol [163], from [185].

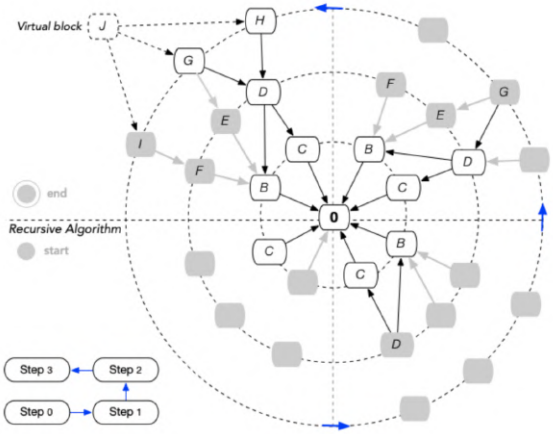


FIGURE 4.10: Divergent topology: Phantom protocol [151], from [185].

Literature on *DAG* taxonomies identifies three types of *DAG* ledgers based on their network topologies [185]: divergent, parallel and convergent. Divergent *DAG*s are sparse natural graphs spreading in unpredictable directions. Convergent *DAG*s are highly organised graphs with predetermined cluster sequences, often designed to converge towards a reference blockchain as part of multi-layer protocols that combine traditional blockchain architectures and *DAG* features [185]. The Phantom [151] (Figure 4.10) and Haootia [163] (Figure 4.9) protocols are two examples of these two types of *DAG* structures.

The research focuses on parallel *DAG*s: transaction clusters representing independent perspectives of a ledger while participating in consensus. Parallel *DAG*s are suitable for the proposed application. Supply chain stakeholder groups operating in different platforms can create trustworthy cross-platform references, adapt them to their internal process logic and validate transactions. At the same time, customs administrations can take part in applications related to declaration processes and oversee the aggregated perspective of applications relevant for audit purposes. The implementation of a *DAG* consensus service is outside the scope of the research, but supporting work on the need for asynchronous consensus and the consensus in the chosen *DAG* protocol is included in Appendix E.

Conceptual DAG Ledger Design

The architecture presented is based on the CAPER protocol: an asynchronous ledger where different applications run on a number of nodes known as *agents* [9]. An application refers to a private smart contract in which a specific logic is encoded as the rules to process internal transactions. These contracts only run in the nodes of the application. Additionally, rules to process cross-application transactions can be included in public contracts. Languages widely used to encode smart contracts, such as Solidity [95], can be used for both private and public contracts to ensure the deterministic execution of transactions [9].

Sensitive business logic can be kept confidential within an application while standardised procedures can be encoded as public contracts to facilitate the exchange of information to trigger smart contracts in other applications. There are no tested and reliable solutions that allow internal and cross-application transactions between untrusted applications within a ledger. The CAPER concept can overcome this barrier that is hindering the development of efficient, scalable and secure cross-application communication.

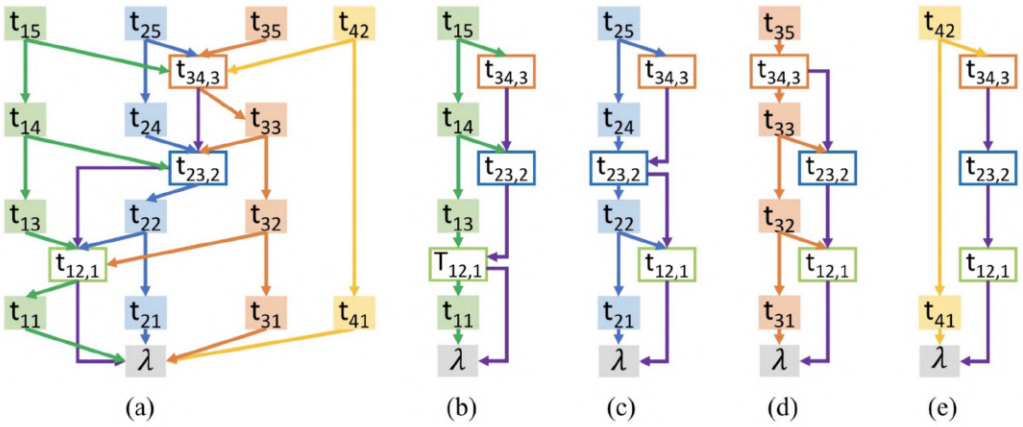


FIGURE 4.11: Example of distributed applications: main *DAG* ledger (a), consisting of four parallel applications (b, c, d, e) [9].

Its essential contribution is the distinction between trust at node level and at application level. While the nodes within an application, such as the participants in a *Hyperledger Fabric* channel, might not behave maliciously within the application, the application might behave maliciously when interacting with other applications [9]. CAPER allows coordinating concurrent transaction logic around permission-ed ledger environments. Its main function is *"considering both the confidentiality of internal states generated in each application and the interoperability of external states that come from the cross-application transactions"* [185]. An example of these parallel applications is shown in Figure 4.11.

The proposed separation of applications can be used against some data distribution problems faced by carriers and customs administrations, such as the *ENS* declaration level of Figure 2.6, discussed in subsection 2.4.2. A single record could act as the genesis record of multiple applications using the same public contract. This can be used to trigger additional private contracts in internal application logic to update a private perspective of a supply chain. Moreover, less frequent data duplication can help customs administrations detect dependencies between risk assessment data effectively. A detailed demonstration for the use of this design in the supply chain domain is presented in chapter 5.

Detailed Transaction Graph Model

The mathematical formulation of *DAG*'s has a lot in common with the formal representation of transaction semantics ruling blockchains. A good reference is the *TDAG* model developed by *IBM* [25], resulting from the generalisation of state transitions in blockchains. The *TDAG* model has been successfully applied to represent transaction structures and validity rules of well-known blockchain systems, including *Bitcoin*, *Ethereum* and *Hyperledger Fabric* (see Appendix D). Since the purpose of the artefact is to synchronise transactions from more than one blockchain protocol, *TDAG* is an interesting tool to formulate a conceptual model to migrate from siloed blockchains to a cross-chain transaction ecosystem.

A graphical overview of the type of transactions included in the model are shown in Figure 4.12. There are two kind of nodes: *states* and *witnesses*. A *state* is depicted as a circle. They can represent an individual digital asset, the link to an asset or their properties in the form of verifiable credentials, any variable included in a smart contract or cryptographic proofs of their completion. The genesis state is the initial state of the ledger and is depicted as two concentric circles. A *witness* is depicted as a rectangle. It represents the information necessary to validate a transaction according to a specific set of rules. Exactly one *witness* is required in order to formalise a transaction.

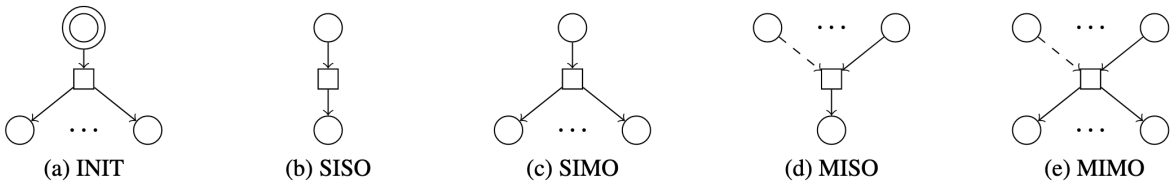


FIGURE 4.12: Graphical representation of transactions: (a) initialisation, (b) single-input single-output, (c) single-input multi-output, (d) multi-input single-output, (e) multi-input multi-output [25].

There are also three kind of edges: consuming, observing and producing. Consuming edges link a *state* to a *witness*, meaning that the *state* becomes permanently linked to the unique transaction the *witness* is part of. Once a *state* is *"consumed"* by a transaction no other consuming edges can be produced by the *state*. Observing edges provide visibility paths between *states* through *witnesses*, meaning that a *state* becomes part of the transaction linked to the *witness* while conserving the ability to produce consuming edges. These allow multiple transactions to *read* a state [25] and are depicted with dashed arrows to differentiate them from consuming edges. Lastly, producing edges link a *witness* to a *state*, and establishes a causal relationship between previous *states* through *witnesses*, which represents that a new *state* has been created.

Using this notation, a transaction is the transition from one *state* or a set of *states* towards a new *state* or set of *states*. The validity of the transaction comes from its unique *witness*. A transaction is formed by *input states* that the transaction consumes or observes and *output states* produced by the transaction. An example of this notation is shown in Figure 4.13 while the resulting transactions are shown in Figure 4.14. Now it is possible to link ledger states with the semantic model by expressing claims using the *TDAG* model.

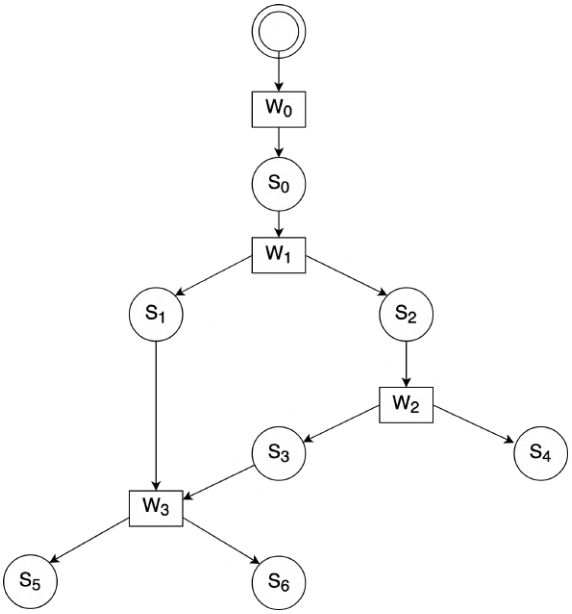


FIGURE 4.13: Example of a *TDAG*, adapted [25].

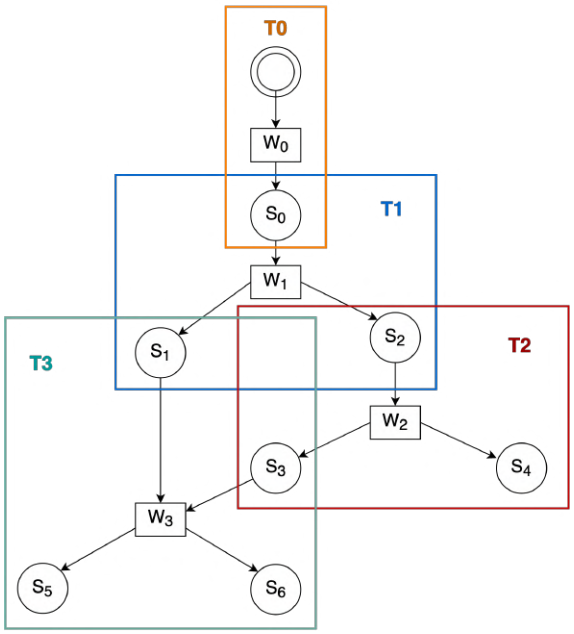


FIGURE 4.14: Mapping transactions within the *TDAG*.

A graph of transactions with shared *states* and validating *witnesses* can be converted into a *DAG* ledger. This is done exploiting the fact that the objects of two consecutive transactions in the *TDAG* domain overlap, as output states can be input states of another transaction. For the purpose of the research, describing the internal logic of a blockchain is not sufficient to model arbitrary events external to the blockchain itself. Therefore, it is useful to express an equivalent logic using each transaction as main unit. Representing graphically the links between the verified *states* of external ledgers can be done using their *witnesses* as part of the authentication of a verifiable presentation linked to a *DID* document. This is where the credential management layer and the event visibility layer interact. In order to validate a new record, the proof graphs attached have been authenticated by the asynchronous accumulators covered in subsection 4.4.6.

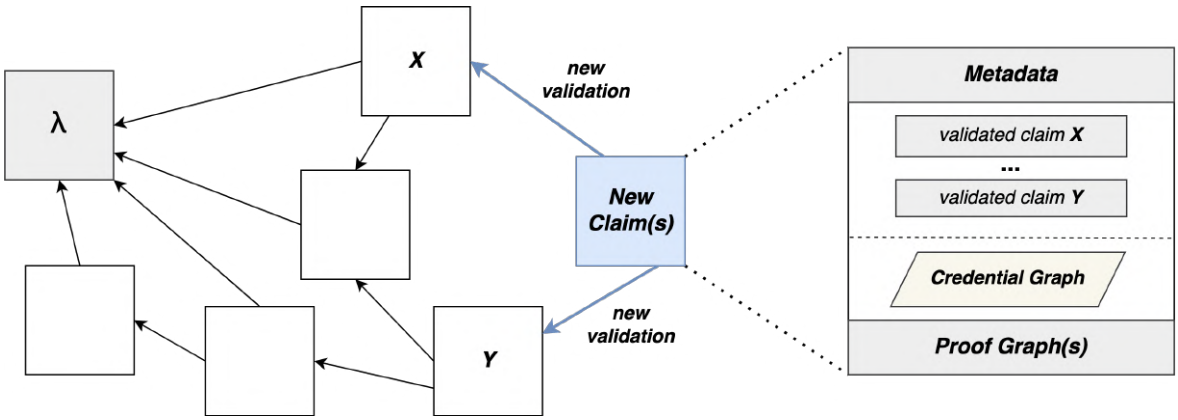


FIGURE 4.15: Record format and validation, adapted [197].

Understanding transactions as credentials, each claim is appended individually and left unconfirmed. To append a new claim, previous claims must be validated. This process is shown in Figure 4.15, where the direction of the edges does not indicate the flow of information but the order and direction of the validations starting at the genesis claim λ. The metadata of each credential contains fields such as a unique identifier, timestamps, the claims being validated and other properties. Each credential also includes one or more proof graphs, which are related to the presentation encoding used. Depending on the cryptographic system used to authenticate the credential, combinations of keys and certificates signatures might be included. Translating the transaction graph shown in Figure 4.13 using the event network model results in Figure 4.16.

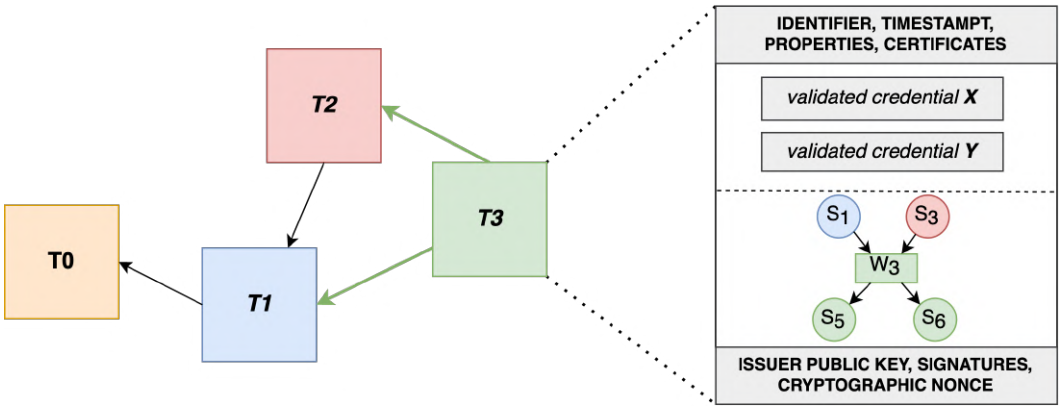


FIGURE 4.16: Example conversion from *TDAG* to event network.

This example uses the individual transactions of one ledger, but the same approach can be used to build event networks of cross-chain transactions. This is visualised in Figure 4.17. These cross-chain logs are the backbone of the piggybacking functionality of the design. The longer the *DAG* network, the more piggybacking chains are possible.

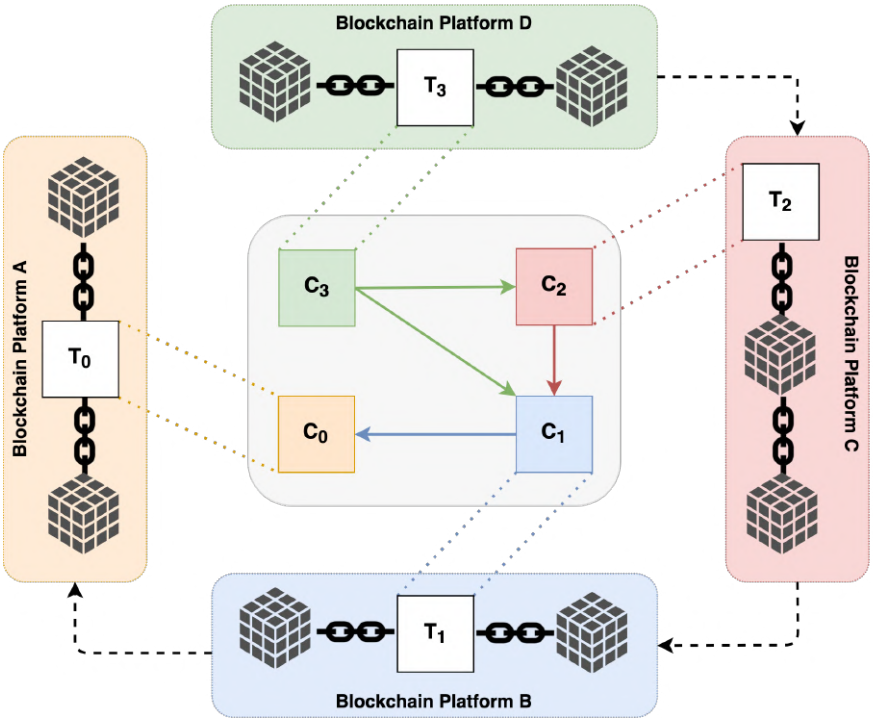


FIGURE 4.17: Event network for platform integration.

4.3 Cross-Chain Communication Layer

The presented *DAG* applications require a network infrastructure to support the peer-to-peer connections between blockchain platforms. This section presents such interoperable cross-chain communication solution. Cross-chain interoperability can be seen from a number of angles. There are challenges that can be addressed in each layer of blockchain architectures as shown in Figure 4.18. Based on this framework, different approaches have been explored to tackle the cross-chain communication problem. In practice, interoperability solutions must take into account the interaction with other layers, but this section focuses on the effect at network level of a design based on an overlay gateway network that supports an end-to-end resource transfer protocol. The following subsections cover the selection of components and their functionality within the cross-chain communication layer.

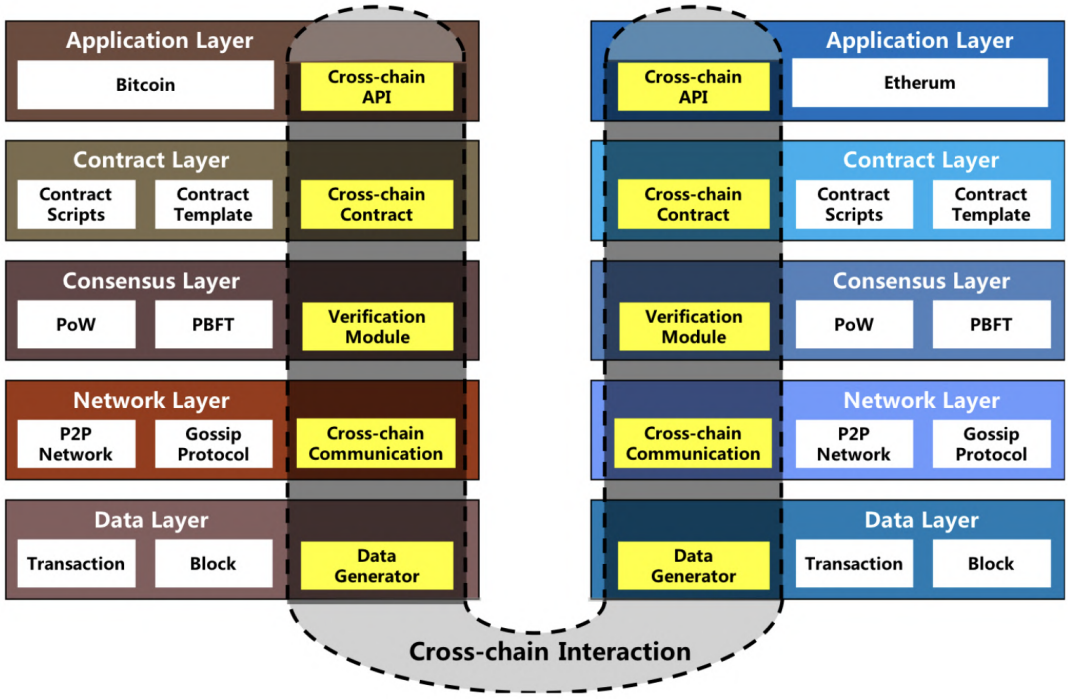


FIGURE 4.18: Cross-chain framework [89].

4.3.1 Comparison of Interoperability Solutions

There is no consensus on the classification of interoperability solutions. An approach is to differentiate between chain-based, bridge-based and *dApp*-based solutions [183], which can be alternatively referred to as public connectors, hybrid connectors and *blockchains of blockchains* respectively [131].

Chain-based solutions focus on the chain-to-chain interactions behind atomic swaps. Bridge-based solutions build connections between blockchains to reduce or remove large technical incompatibilities between layer components. The purpose of *dApp*-based solutions is to ease the implementation of and interaction between decentralised peer-to-peer applications, and represent a more holistic approach to interoperability linked to the emerging *Blockchain-as-a-Service* design paradigm (*BaaS*) [106, 152]. Among these categories, literature distinguishes four subcategories that are relevant for the research: sidechains, notary schemes, hash-locks and trusted relays [2, 45, 62, 70, 98, 119, 131, 160, 183, 193].

Sidechains act as complementary chains build around a mainchain [183]. They are used as buffers to delegate certain phases of resource transfer protocols, and can be designed as one-way or two-way systems [131]. However, the number of sidechains required in the blockchain environment described in Figure 3.1 would grow at an unsustainable rate, as well as adding unpredictable complexity to the maintenance of the design. Therefore, sidechains have been discarded the due to their limited scalability in the proposed application.

Notary schemes rely on a trusted third party to monitor multiple blockchains, witness the terms of cross-chain commitments and trigger the execution of contracts [45, 98]. A popular application of notary schemes are centralised cryptocurrency exchanges, where the security of token transfers is guaranteed by the platform provider [131, 183]. This option radically increases the centralisation of the system and is thus not a preferred design approach, particularly in terms of data sovereignty and decentralised trust.

The next subcategory are hash-locks. They can be described as decentralised escrow services that can alter the ownership of assets without relying on a trusted third party, unlike notary schemes [70, 98]. They can be chained after each other [119], which makes them particularly useful when transaction sequences want to be programmed between entities with no direct connections [131]. Hash-locks can be implemented as smart contracts triggered by arbitrary conditions, such as the time limits for the provision of cryptographic proofs used in hash time-lock contracts (*HTLC*) [183]. Also, when combined with the appropriate network configuration, hash-locks allow a group of entities operating in independent blockchains to exchange authenticated updates of the internal state of their ledgers [8, 42].

The last solution type are trusted relays, with a focus on trusted gateway bridges [70]. Also referred to as *relay services* [2] or *chain relays* [193], they handle requests to fetch ledger state proofs between remote networks and verify application logic [2, 132]. Relay services make it possible for an entity on a chain to verify events registered in other chains by building a bridge that provides smart contract services between platforms [183]. This means that a smart contract implemented in one chain can become a *client* of another chain [193]. For the research context, the main advantage is that they allow *clients* to define arbitrary business logic that can be fed with evidences of external data without a centralised entity [62, 131]. It is worth mentioning that the functionalities of trusted relays resemble the concept of service orchestration discussed in section 4.1.

Research on more advanced blockchain-agnostic protocols, consensus engines and security infrastructures able to aggregate complete architectures are being developed [131]. However, they fail to offer backward compatibility, implying that legacy systems would need to be heavily modified. In view of this, trusted relays are seen as the most realistic solution to link network layers between permisisoned blockchains while maintaining a design philosophy inspired in the *d-App* paradigm [70, 131].

4.3.2 Selection of Interoperability Solution

Most literature on blockchain interoperability focuses on technically connecting two or more ledgers, lacking organisational and value-driven assessments [98, 131]. This means that the maturity of interoperability solutions is low, and that implementation challenges are yet to be found when studying stakeholder-oriented views. Blockchain discoverability, privacy and governance are fields strongly attached to interoperability, yet they are not sufficiently researched [131]. This is evidenced by the fact that the solutions either focus on technical

nuances behind on-chain interactions or high-level descriptions of frameworks for the deployment of *d-Apps* with no immediate practical value. A good example is the addressed migration of current customs declaration processes to a ledger-based environment. Despite the technical possibility to physically connect information between ledgers, solutions to overcome barriers related to regulation and business incentives are needed.

Literature on interoperability suggests that hash-locks and trusted relays can reduce this gap. The former can help data subjects produce their own data governance rules while complying with regulations, and the latter enables trusted peer-to-peer interactions to unify business logic. Although these solutions are still on their infancy, they can complement each other and elevate the effectiveness of a cross-chain data-sharing architecture. Moreover, there are explicit suggestions for further research on the combined use to improve interoperability between permissioned environments [131] and to increase supply chain visibility [2]. Besides their technical suitability, blending the features of trusted relays and hash-locks is aligned with the core research goal in the value plane: choreographing interactions between supply chain stakeholders. Combining both technologies allows to proof the technical capabilities of the design and also demonstrate it can adapt to varying business logic.

4.3.3 Gateways & Overlay Network

Gateways are used to relay (connect) a *client* and a *source*. They are normally dedicated nodes, but can be implemented as an additional service layer within a permissioned network [70]. This depends on the level of centralisation in the architecture and consensus mechanism used. In any case, gateways that represent a group of nodes maintaining independent ledgers can be grouped. Assuming every node can interact with its gateway, peers can leverage the functionalities of their ledger to offer *ad hoc* services to external clients [112]. Creating a logical layer above the networks represented by each gateway to provide these services results in an overlay network, which is shown graphically in Figure 4.19.

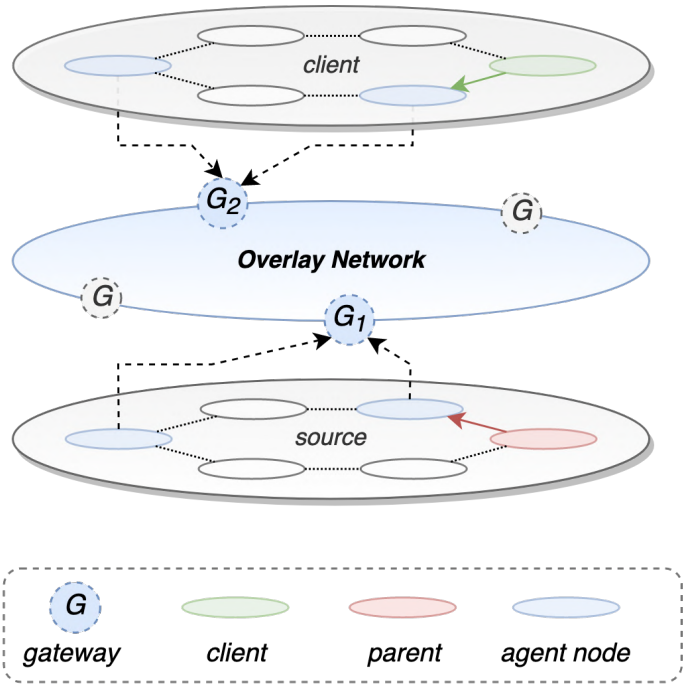


FIGURE 4.19: Overlay network.

The model includes a *client* and a *source*. This terminology should not be confused with the terms *sender* and *receiver* used in atomic swaps, *e.g.*, transfers of fungible tokens linked to account balances for financial applications. The goal is rather to describe an environment where information (or traces towards information) about ledger states becomes accessible under certain rules to legitimate entities. Depending on the use case, this may mean migrating the copy of an asset between ledgers or only forwarding transaction proofs to propagate application logic between independent ledgers.

In this context, the concept of *dependency* is introduced: a piece of data stored in a ledger that is required in order to initiate or complete a service (see section 4.1). As mentioned before, a dependency can take the form of a non-fungible asset (a *eB/L* or *L/C*), a zero-knowledge proof describing the properties of an asset or a proof of an asset being issued in other ledgers under certain conditions. Smart contracts can be used to rule the visibility or the transfer of ownership of the dependencies. The overlay network provides structural support to route the messages produced by the transfer protocol feeding these contracts. The cross-chain resource transfer protocol used in the design is covered in subsection 4.3.4.

The *client* contains a *client node*, which is a node that requires a dependency to execute a transaction. Similarly, the *source* is that where the dependency is registered. The *source* contains a *parent node*, which originally validated the dependency. *Agent nodes*, also understood as committee members in general [3], have increased visibility over a ledger´s activity due to governance privileges. They might be irrelevant for less sophisticated platform architectures, but they play a crucial role in the publication of internal ledger states for platforms using certain *BCT*s, such as peer agents in *Hyperledger Fabric* [3, 70] or oracle nodes in *R3 Corda* [111, 130]. Different agent node configurations and their potential effect on data sovereignty within a platform during ledger state publishing is discussed in subsection 4.3.5.

It should be noted, that a blockchain node might access its gateway in different ways depending on the architecture of its platform. Figure 4.19 shows the most common scenario, in which a dedicated node is in charge of managing the gateway. However, as shown in Figure 4.20, there are other alternatives. Figure 4.21 shows an example in which the upper platform uses a dedicated node while the lower platform uses a completely decentralised and meshed access with multiple gateways.

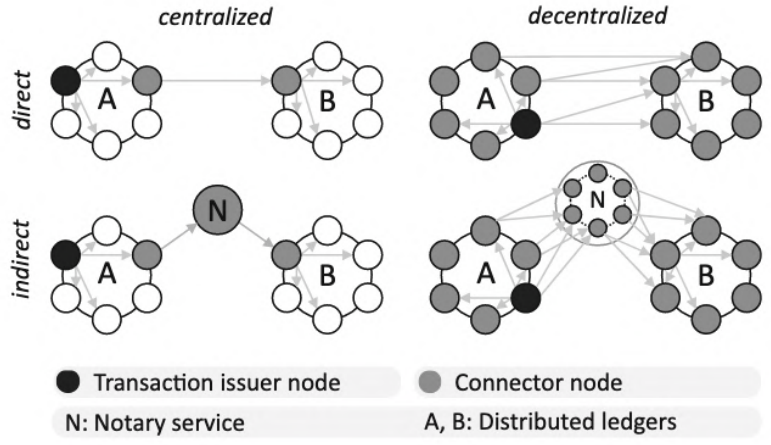


FIGURE 4.20: Cross-ledger interoperability patterns [160].

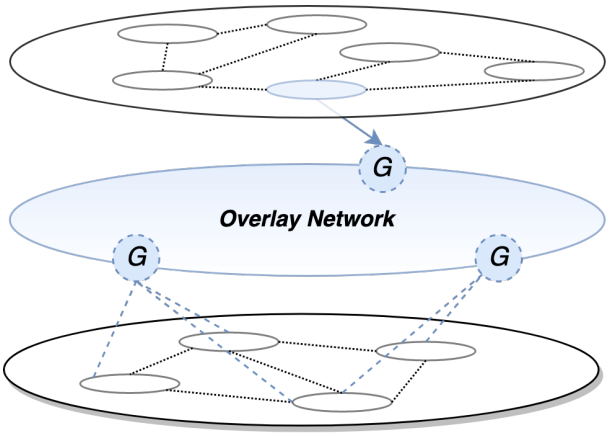


FIGURE 4.21: Gateway access alternatives.

Since a gateway belongs to a platform, multiple entities operating in that platform should be able to discover and authenticate it. The internal node communication has an effect on the publication of ledger states (see Figure 4.25), which makes it is challenging to assume a specify gateway access that is compatible with all platforms configurations. These considerations are included in the gateway identity model covered in subsection 4.4.5, until which trusted gateways are assumed, *i.e.*, gateways can authenticate each other´s identities.

4.3.4 Resource Transfer Protocol

The overlay network described in subsection 4.3.3 is used to maintain distributed cross-chain transaction logs. It is powered by a cross-chain protocol that allows platforms to exchange resources or links to resources. These exchanges are done between a parent node and a client node, who retrieve proofs from their ledgers and relay to the overlay network through their respective gateways. Figure 4.22 shows the conceptual design of the protocol. The verification between nodes is done by resolving *DID* documents. The bridges between gateways are authenticated using a publisher-subscriber system. If needed, resource transfers are executed directly between storage locations.

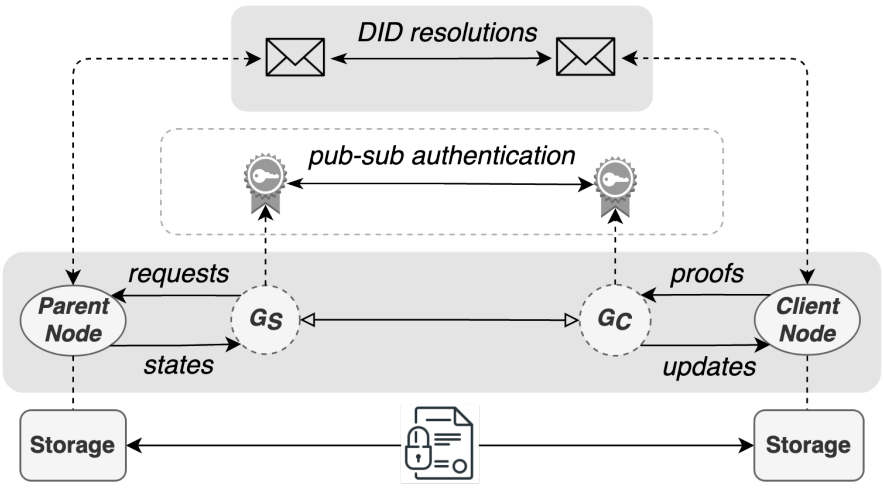


FIGURE 4.22: Conceptual protocol design.

Gateway Discoverability

One of the drawbacks of relayed networks is their static nature, meaning that their participants must know each other´s identities and configurations *a priori* [2, 131]. While this might not be an issue for permissioned environments with fixed participants, network discovery is important when participants are added and removed dynamically, which is the case for the research context. Modular designs are able to improve dynamic discoverability with a credential registry next to a publisher-subscriber (*pub-sup*) system [143, 147].

The logic of the latter is shown in Figure 4.23. It allows sources to share application logic and *clients* to receive verifiable updates via their gateways. The credential management layer controls access to the overlay infrastructure by processing self-certifying identities based on a decentralised identifier (*DID*) method (see subsection 4.4.4). It interacts with the transfer protocol through asynchronous accumulators, which provide dynamic membership proofs of publisher and subscriber credentials (see subsection 4.4.6). These proofs are used to regulate the access to cross-chain application logs covered in section 4.2. The transfer protocol is then used to distribute application updates among verified subscribers (clients), and the credential database shown in Figure 4.4 acts as a registry of the witnesses (see subsection 4.4.6) that certify said *pub-sub* relationships.

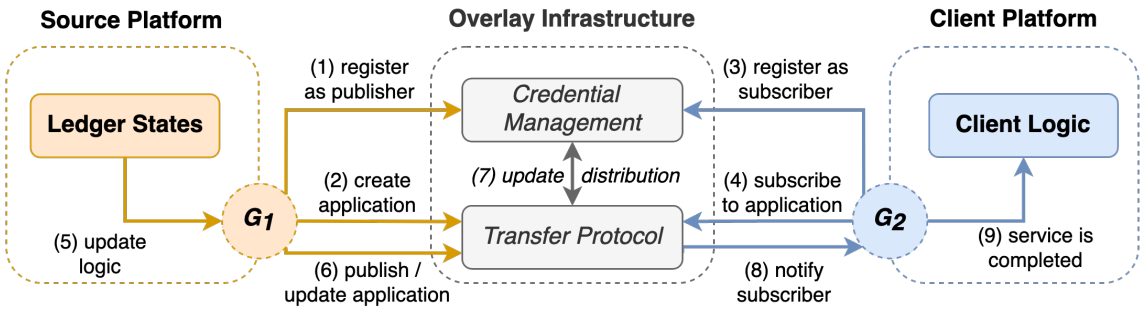


FIGURE 4.23: Publisher-subscriber system, adapted [143].

This *pub-sub* system can be used by European customs to enforce specific data dissemination rules based on the relationships between publishers and subscribers. These relationship rules can be programmed within a *DID* method (see subsection 4.4.4) regulated and maintained by European customs and integrated by the participants in the network.

Detailed Protocol Phases

An overview of the interactions between a parent node and a client node through the phases of the protocol is shown Figure 4.24: gateway validation, application commitment and resource exposing. There is a preliminary phase for *pub-sub* registration. A gateway validates its identity as covered in subsection 4.4.5. Then a parent node uses the gateway to register as the publisher of internal ledger states in an application. Similarly, a client node registers as the subscriber declaring a verifiable proof of their relationship to the publisher (see subject-holder relationships in subsection 4.4.3). Let us then assume that a node has taken part in a transaction within his platform. If the node wishes to share a verifiable prove of that transaction, *i.e.*, letting potential clients observe passively the state of his ledger, he can retrieve an on-chain proof from the platform´s blockchain. This is how the application commitment phase starts. The proof is then converted into a verifiable object subject to the core data model for *DIDs* [133] and processed by a decentralised application.

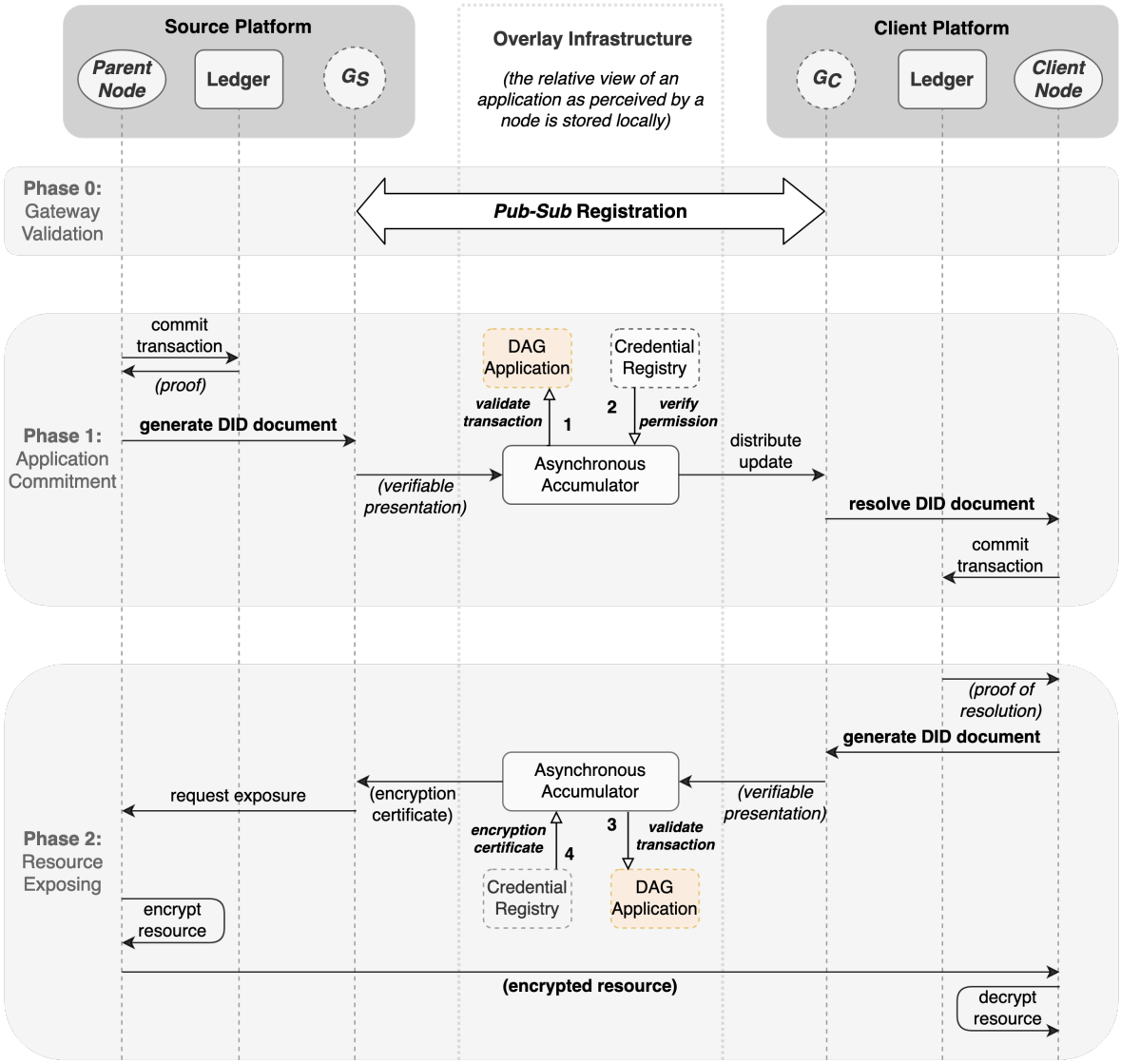


FIGURE 4.24: Resource transfer protocol.

Most asset transfer protocols define a clear originator and beneficiary from the beginning [70, 74]. In the proposed approach, at this stage the publisher can initiate a commitment before identifying a client. This highlights the asynchronous nature of the design. Routes towards resources stored in blockchain platforms are prepared locally in view of becoming dependencies for the services of other platforms. Afterwards, a client can activate this path after verifying his *pub-sub* identity and agreeing on the conditions of the *DID* method (see subsection 4.4.4) used by the parent node in a cross-chain application. As discussed in section 4.2, a node maintains a local view the application depending on his visibility of the application logic. A schematic representation of this process is shown in Figure 4.24 as an overlay infrastructure segment.

At this point, a tamper-evident proof of the source state is created. The state describes the completion of smart contracts or the existence of digital resources. Regardless, subscribers of the application containing the state receive an application update. If it includes chaincode proofs, the view can be used by a client as witness for contracts on his platform (section 4.2.2). If the state describes the storage of a resource, the client continues towards the resource exposing phase.

In order to proceed, the client must show a resolution proof for the transaction received from the decentralised application. This means that the client node must prove that he was able to resolve the *DID* document used to present the dependency and consume it in its local application logic. This is how the asymmetric cryptography concept (see subsection 4.4.4) is applied in the protocol. After using his ledger to generate this proof, the source gateway will forward an exposure request to the parent node, who will start preparing the resource to be transferred from his off-chain storage to the clients off-chain storage. The routing for the storage-to-storage transfer is left outside of the research scope. However, this can be achieved by fetching endpoint data from the *DID URLs* (see subsection 4.4.4) included in the exposure requests.

This approach is inspired by the *Distributed Trust Backbone* model [129], where a distinction is made between an *Identification & Authentication* protocol and an *Information Exposing* method. The main motivation to separate the aggregation of cross-ledger proofs in the decentralised applications and the end-to-end resource transfer is the fact that the latter is not always needed. For example, lodging documents dedicated to customs declarations is not required by European customs anymore. As discussed in subsection 2.4.2, the most recent updates of the *Union Customs Code* indicate that carriers are allowed to share declaration data by granting customs access to their private computer systems [53]. Therefore, customs could subscribe to the decentralised applications published by carriers as long as the *DID* resolutions and overlay access conditions comply with the customs procedures.

Smart Contracts

The overlay access conditions governing the protocol can be programmed as hash-lock contracts embedded to the properties of the *DID* documents published in decentralised applications. In the same way entities are currently responsible for lodging declaration data or granting access to the declaration data in a timely manner, publishers of the system are responsible for stipulating the *HTLCs* that respect regulated time frames. However, the data to be handled are references to non-fungible resources, so encumbrance is not strictly required. Therefore, the main difference with hash-locks for on-chain transfers is that what is being locked is a key-pair, an identity certificate or whatever property is chosen to maintain the discoverability of the resource. This design implies that the credential management must be able to control key rotation and certificate revocation, which is covered subsection 4.4.4.

A direct application are time limits after which European customs must remove data records of inactive economic operators [55] (see *NFR8* in section 3.4). Instead, the economic operators can share a discoverable presentation of their data, which will be active during the time limit. Although economic operators are still allowing other entities to access their data, they are not giving away their data sovereignty rights. This approach allows entities to take full control of both the visibility and ownership of their resources.

4.3.5 Publication of Ledger States

Until now, it has been assumed that platforms have the ability to generate verifiable proofs of their internal states and that there is a single gateway for each platform. However, the platform participants might not want to share all the information with their peers. It is possible that the configuration of a platform leads to the creation of multiple independent ledgers in which the interactions between nodes is restricted. Groups of nodes integrating an application might be reluctant to sharing a node in charge of the publishing service. Therefore, using a unique agent node for all the applications within a platform is not realistic.

To solve this issue, platforms include dedicated services to manage the publication of ledger states or transaction bundles, such *event-hubs* and the *peer channel-based event service* in *Hyperledger Fabric* [83], or the *oracle service* in *R3 Corda* [130]. These services allow groups of nodes to control how the information stored in their ledger is shared. The overlay network must thus interact with these event services. This is done via a state view board, from which ledger states that want to be published by parent nodes are retrieved.

The approach is based on the bulletin board protocol developed by [3] for verifiable state observations on permissioned ledgers. The design is intended to be plug compatible with commercial permissioned blockchain stacks, and has been implemented in *Hyperledger Fabric*. Its bulletin concept is similar to blockchain: it is append only and new entries are signed by its publisher after producing a rolling hash.

Depending on the platform configuration, it can be used by multiple parent nodes simultaneously, which would mean that previously published views are validated with the addition of new views. In that case, implicit verification of state agreement could be added when required (see subsection 4.4.6). The bulletin board model can also improve the effectiveness of the asynchronous accumulators covered in subsection 4.4.6. However, assessing these functionalities in detail and the implementation of the bulletin board is outside the scope of the research. Refer to the work of *Abebe et. al* [3] for further reading. An overview of the resulting cross-chain communication concept is shown in Figure 4.26.

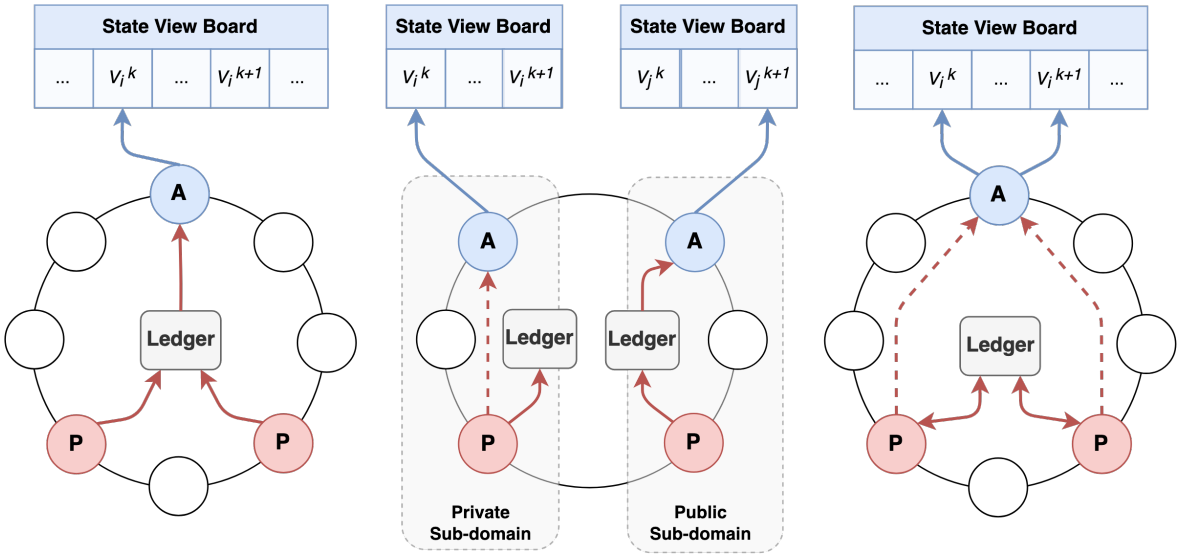


FIGURE 4.25: Platform layouts and types of state publishing: public (*left*), decentralised (*center*) and private (*right*).

As shown in Figure 4.25, nodes can interact with agent nodes in multiple ways. In the case of permissioned platforms with trusted nodes maintaining a single ledger, the state view board is directly fed by validated ledger updates without compromising the confidentiality between nodes. This is referred to as public publishing. Another scenario is a similar single ledger but with trustless nodes. Although a ledger is shared, public publishing can compromise data confidentiality between the nodes. Causes might include parent nodes not wanting to aggregate their publishing registries or their will to attach chaincode information to the board item in addition to the ledger commitment proof [83, 130].

Moreover, modular platforms with multiple ledgers maintained by groups of nodes can adopt a decentralised publishing style. The data flows between these groups of nodes create isolated sub-domains, such as the *channels* found in *Hyperledger Fabric* [83]. Here, multiple state view boards are maintained for each sub-domain. Also, public and private publishing can be combined in the decentralised publishing.

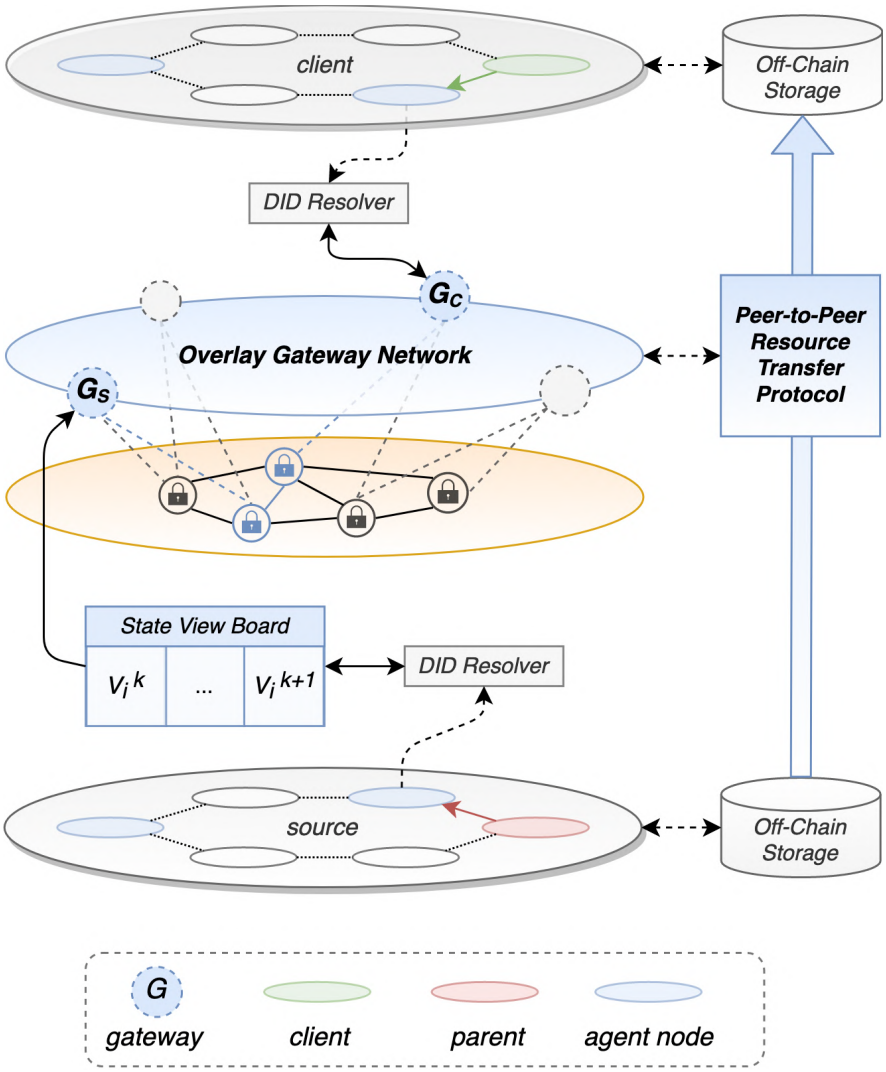


FIGURE 4.26: Cross-chain communication concept.

4.4 Credential Management Layer

This section covers the credential management layer. First, a general model to express verifiable credential claims is presented, followed by an introduction to the self-sovereign identity paradigm and a decentralised credential management framework. These elements are the foundation of the access control of the architecture, which uses asynchronous accumulators for certificate management and cross-chain state authentication. The section also covers the link between these components and the protection of data sovereignty.

4.4.1 Information Graphs & Verifiable Credentials

The content of this subsection is based on the specifications of the verifiable credential data model of the 2019 *W3C* Recommendation [154]. The verifiable credential model has its roots in the abstract model of the Resource Description Framework (*RDF*), which was originally conceived as a metadata encoding standard [190].

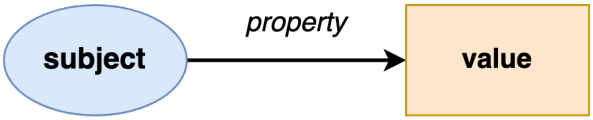


FIGURE 4.27: Standard abstract *RDF* claim [154].

An entity, such an individual or an organisation, can generate data containing claims about one or more data subjects (see section 3.4). These claims, also called *triples*, can be expressed as abstract *RFD* statements following the *subject-property-value* structure shown in Figure 4.27. Claims about a *subject* can be merged with other claims creating an information graph as shown in Figure 4.28. Large networks of information expressing the relationships between subjects, resources or processes can be constructed using these networks.

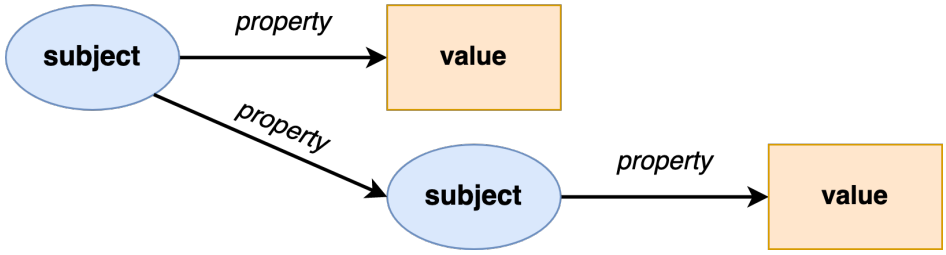


FIGURE 4.28: Information graph formed by claims [154].

A *subject* refers to a data subject. A *value*, also called *object* [190], can represent both tangible resources (*e.g.*, transport infrastructure) and abstract concepts (*e.g.*, risk category codes). A *property*, also called *predicate* [190], describes how the latter are related and is application specific. It can for instance represent a process status update (*e.g.*, customs release of cargo) or a legal bind between a *subject* and *value* (*e.g.*, assignment of liabilities). In the research context, this allows mapping logistic processes (*e.g.*, cargo custody chain) and commercial agreements (*e.g.*, contracts of carriage) using the same language.

When abstract *RDF* statements are grouped and stored as tamper-proof information graphs with cryptographically verifiable metadata, the term *verifiable credential* is used. The metadata describes credential properties like its issuer (see subsection 4.4.3) or revocation mechanisms [154]. Verifiable credentials are useful for access control in data exchange transactions,

as they can improve the understanding of cross-chain interactions when combined with the appropriate semantic model (see subsection 4.2.1) and technology that supports logic fed by cross-chain state proofs (section 4.2.2). Their strong link with digital identities and data confidentiality is also very relevant and is further discussed in section 4.4.

When (partial) data is extracted from a verifiable credential and is encoded so that its origin and authorship can be processed by a verifier, the term *verifiable presentation* is used. These normally include data synthesized from multiple original verifiable credentials in the form of zero-knowledge proofs for authentication purposes. Figure 4.29 shows the basic architecture of verifiable credentials and presentations. Color-coded format examples of complete information graphs can be found in Appendix B.

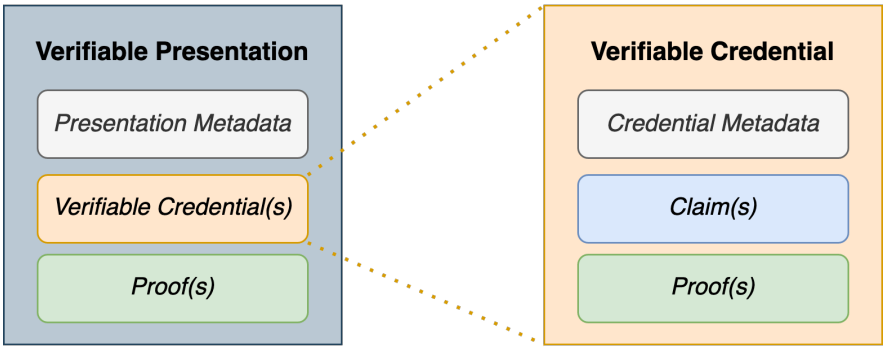


FIGURE 4.29: Verifiable presentation and credential basic architecture [154].

The model includes proofs used for provenance verification or the authentication of access rights. The format and validation mechanism of the proofs can be encoded in the credential itself to be processed following the rules of the environment the issuer, holder or expected recipient of the credential operates in. Besides technical nuances, the main take should be that verifiable credentials can be used to avoid unauthorised data dissemination, access and interpretation when combined with self-sovereign privacy design [6, 37].

4.4.2 Introduction to the Self-Sovereign Identity Paradigm

Unlike in the physical realm, where identities are backed up by associations to tangible objects, it is important to specify a clear structure when it comes to defining identities in the digital realm. Large reference networks between abstract objects make it possible to organise and connect natural persons, organisations and digital assets. Such powerful yet unintuitive space requires a standardised framework to deal with digital identities by differentiating between the real entities, their identities and their attributes. A graphical representation of the links between these three concepts is shown in Figure 4.30.

Besides natural individuals and organisations, it is possible to associate a digital identity with an information system (a private communication network) or a digital resource (digital wallets, documents or any digital asset). Digital identities defined by verifiable attributes foster the interactions between entities as they promote trust. However, entities might have more than one identity depending on the digital interaction they take part in and the point in time a transaction is executed [91]. Since individuals, organisations and systems interact constantly with each other, links between their identities are achieved in practice through attribute aggregation due to their overlapping nature [31, 91]. This overlap has been leveraged traditionally with coreferencing [116], although alternatives more compatible with the once-only principle sought after by European institutions are preferred [59].

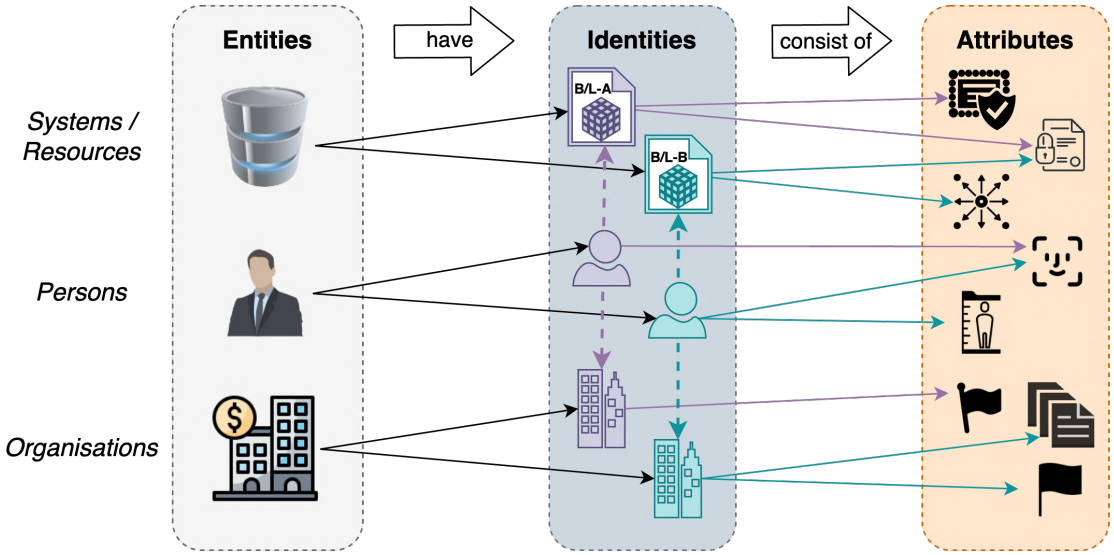


FIGURE 4.30: Entities, identities and attributes, adapted [91].

Following this trend, digital identities have experienced a transformation, starting at centralised identities, moving towards federated identities and arriving to modern decentralised identities [127]. Centralised identities (Figure 4.31) would link a data subject to a digital environment through dedicated credentials. These credentials are managed under internal environment rules and are not recognised by other digital environments.

The figure of *identity provider* (*IDP*) was introduced later in order to reduce the limitations of centralised identities. An *IDP* acts as intermediary between subjects and digital environments enabling reusable credentials to be trusted by more than one digital environment, which is the core concept behind federated identities (Figure 4.32). Although the level of centralisation is reduced, data subjects still rely on an external party to control the legitimacy of their digital identities and not all digital environments might use the same *IDP*.

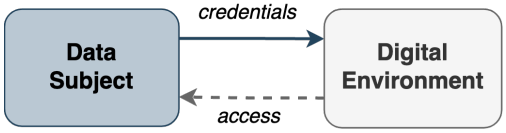


FIGURE 4.31: Centralised identity model.

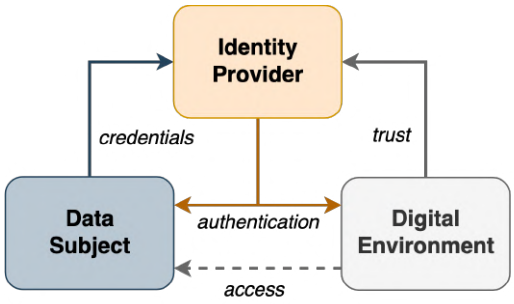


FIGURE 4.32: Federated identity model.

The next step in the evolution of digital identities are decentralised identities (Figure 4.34), which remove the need to rely on intermediaries to engage in digital transactions by fully operating one's digital identity. The characteristic attribute of decentralised identities is the swift from account-based access control towards trusted links between digital peers, whether data subjects represent a natural person, an organisation or a digital resource [127]. In this context, the term *self-sovereign identity* represents not depending on any organisation to make use of your digital identity in legitimate digital transactions.

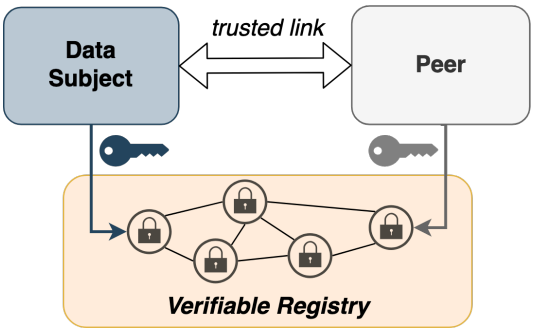


FIGURE 4.33: Decentralised identity model, adapted [127].

The use of the decentralised identities to combat the weaknesses of conventional identity management has been strongly influenced by the adoption of *DLT*, *BCT* in particular [37, 6]. Innovative public-key authentication and verifiable data registries can provide the certainty that an entity is linked to the public key being used in transactions [154, 6]. This is known in general as *public key infrastructure* (*PKI*), which has traditionally depended on the aforementioned *IDP*s [186]. To overcome this, *DLT* can be the backbone technology of trusted verifiable registries to achieve functional decentralised identities without the need for identity certificate authorities. In that case, a more appropriate term for the use of *DLT* for *PKI* applications is *distributed public key infrastructure* (*DPKI*) [135].

4.4.3 Decentralised Credential Management

Using a *DPKI* to manage credentials entails using a framework based on the *SSI* paradigm. An example is the credential ecosystem of Figure 4.34. It is based on the specifications of the verifiable credential data model of the 2019 *W3C* Recommendation [154]. A distinction is made between the subject and the credential holder(s), the latter being the credential possessor(s) with the ability to generate verifiable presentations and therefore engage in transactions. This leads to a redistribution of power relationships as shown in Figure 4.35.

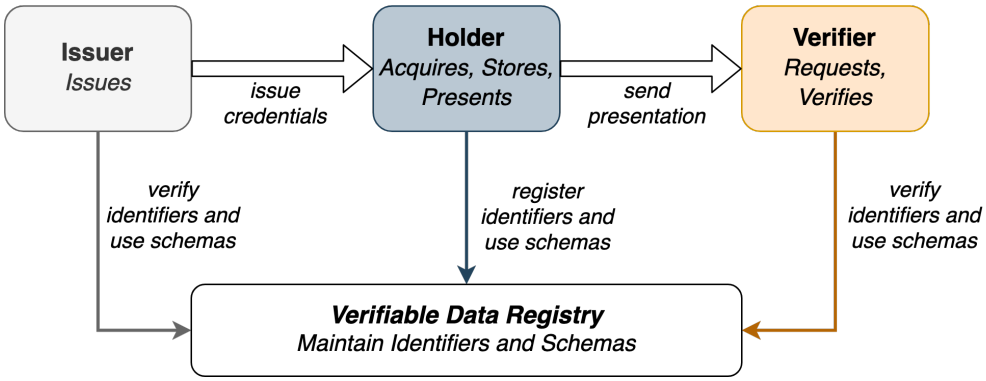


FIGURE 4.34: Verifiable credential ecosystem [154].

There exist many possible subject-holder relationships as shown in Figure 4.36. The simplest use case is when the subject and the holder represent the same entity. It may also be possible that no specific data subject is included in the credential (bearer credential). In the research context this can be the case for the so-called *Bearer Bill of Lading*, which grants access to the transported goods to the entity able to prove possession of the bill [18].

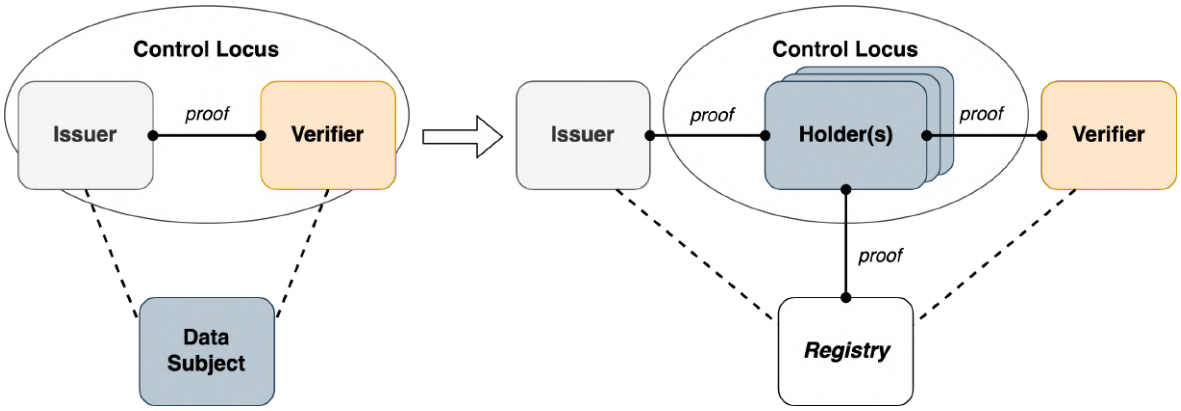


FIGURE 4.35: From traditional control (left) to self-sovereign control (right). Power relationships as circular links, adapted [127].

Even if the holder is not the subject, it is possible to identify the latter by processing unambiguous properties attached to the credential. Some use cases might require the verifier to validate the relationship between the subject and the holder [154]. A subject might also require to issue a second credential to a holder who is able to interact with a verifier. The format and rules for issuing this type of nested credentials are application-specific and can not be generalised. Another scenario is an issuer requiring an entity to become the holder of a credential, while not holding any known relationship with its subject. In that case the issuer can specify a relationship between this entity and itself issuing the credential.

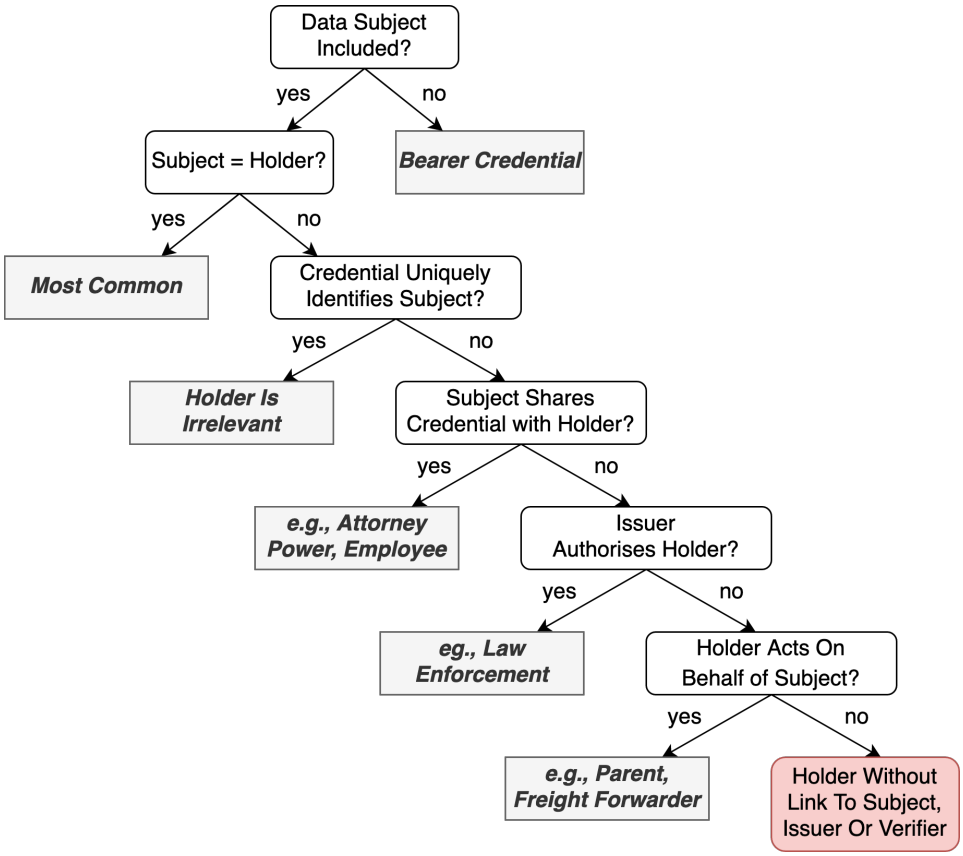


FIGURE 4.36: Subject-holder relationships, adapted [154].

Similarly, the holder can operate the credential when acting on behalf of the subject. This is accomplished in three ways: the issuer can specify the subject-holder relationship directly via the credential, the issuer can provide the holder with a new credential specifying the relationship, and the subject can provide the holder with a new credential specifying their relationship. The two last options require the holder to combine both credentials to form verifiable presentations when executing transactions. Lastly, cases unsupported are those in which the holder acts on behalf of the verifier or when the holder cannot prove any relationship with the issuer nor the subject. The role of subject-holder relationships is used to model and regulate the publisher-subscriber relationships processed during the gateway validation phase of the resource transfer protocol of subsection 4.3.4

Combining decentralised credential management with the right *DLT* for tamper-evident registries can prevent the miss-use of credentials without the holder's consent. Moreover, it increases the visibility of verifiable presentations and their dissemination patterns for audit purposes. Its application for the problem at hand is to transform verifiable presentations of trade and logistic data into reusable and legally binding information that can be safely shared between supply chain stakeholders and observed by customs administrations. Deciding with whom to exchange verifiable presentations depending on the subject-holder relationships encountered in the logistics domain allows supply chain stakeholders to have *self-sovereign* control over their data. Another benefit is that the extent to which credentials can be linked to verifiable presentations issued in the future is also controlled better.

4.4.4 Decentralised Identifiers, Documents & Methods

The content of this subsection is based on the specifications of the core data model for decentralised identifiers (*DID*) of the *W3C* Recommendation Draft of 2021 [133]. It covers how *DID* documents and methods can be used in decentralised credential management to avoid the weaknesses of traditional *PKI*s. Figure 4.37 shows *IDP* certificates for federated identities (see subsection 4.4.2) using asymmetric cryptography.

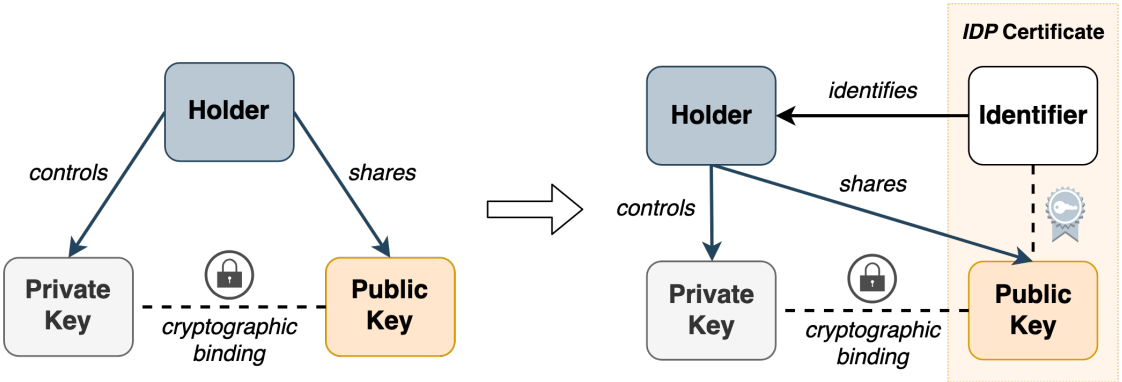


FIGURE 4.37: *IDP* certificates in *PKI*, adapted [127].

It uses a key pair formed by a public and a private key with computationally infeasible mutual reconstruction [100]. There are many alternatives for how to combine these keys for encrypting and signing credentials, which can add many security benefits for information sharing [1, 97, 100]. In the context of constantly changing keys and identifiers, traditional *PKI*s add considerable implementation and maintenance costs, are to some extent still centralised and their availability depends on a single point of failure [127]. Certificate and key rotation improves security. However, it adds complexity to the authentication of credentials in asynchronous environments, such as the cross-chain data exchanges in supply chains.

Unlike discrete atomic transactions with a unique sender and receiver, key rotation might become challenging when a number of entities wish to share information with each other indirectly. Key and certificate rotation should thus support the dependencies of transaction networks without interrupting the verifiability of credentials and their successive verifiable presentations. This means that the information flows must not be restricted to unidirectional sequences, but also allow arbitrary combinations of queries to authenticate credential at any point in time. This way, credential holders are able to engage in uncoordinated transactions without comprising their trust or usability. Figure 4.38 shows a scenario where an original credential is used in a chain of transactions.

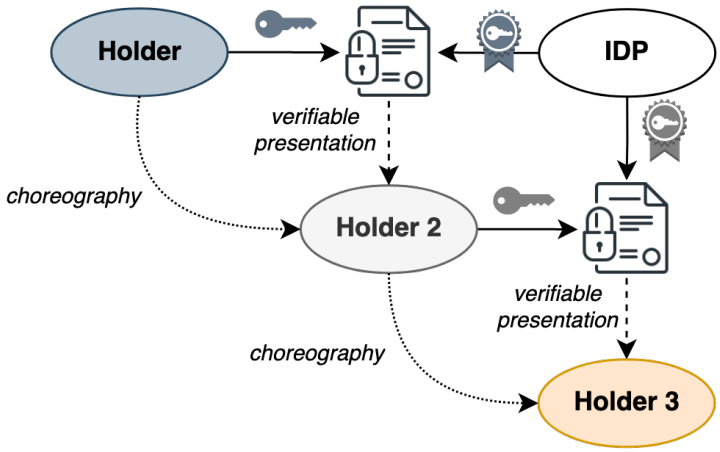


FIGURE 4.38: Credential choreography.

Initially, a holder produces a verifiable presentation for an original credential, which is then reproduced in further transactions. Along the way, keys and certificates generate unique credential properties used for authentication purposes. However, it is common to encounter issues in traditional *PKI*s regarding credential piggybacking and the assurance of auditability, which is another essential aspect for the research. For instance, in Figure 4.39 a service client or regulator is unable to produce a verifiable presentation by merging the information of two transactions due to the lack of visibility of previous credential properties. This is an example of how rotation and certificate revocation can interfere in service choreographies and reduce the effectiveness of piggybacking in *B2G* communication.

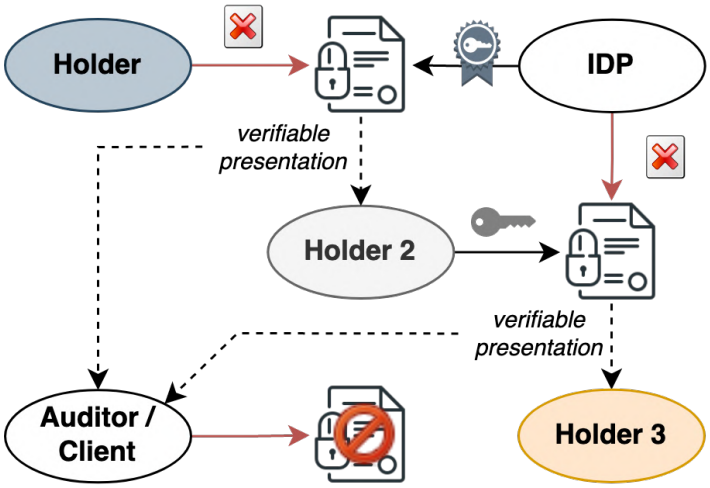


FIGURE 4.39: Effect of key rotation and certificate revocation in credential piggybacking.

This issue can be avoided using the *HTLC*s powering the cross-chain transfer protocol (see subsection 4.3.4) to define key rotation and certificate revocation rules. However, implementing such rules when interacting with traditional *PKI*s is difficult. In that case, an *IDP* would adopt a role very similar to the notary schemes discussed in subsection 4.3.1: issuing the certificates while also coordinating arbitrary verification and revocation rules. The problem becomes even worse when more than one *IDP* is involved. Therefore, an alternative that does not depend on *IDP*s is needed.

In the absence of *IDP*s, self-certifying identifiers can solve the problem. They fulfil the two main duties of *IDP*s: binding a public key to an identifier and binding an identifier to a credential holder. This model is shown in Figure 4.40. Trust in identifiers is achieved by creating cryptographic bonds with public keys. This is effective as the cryptosystems used to authenticate these links are the same trusted methods used to bind private and public keys, such as Rivest–Shamir–Adleman (*RSA*) algorithms, elliptic curves or the Diffie-Hellman algorithm [1, 97, 100]. *IDP*s are not needed as long as the authentication is trusted, because a legitimate holder is the only entity able to generate pairs of identifiers and public keys compliant with the aforementioned cryptosystems.

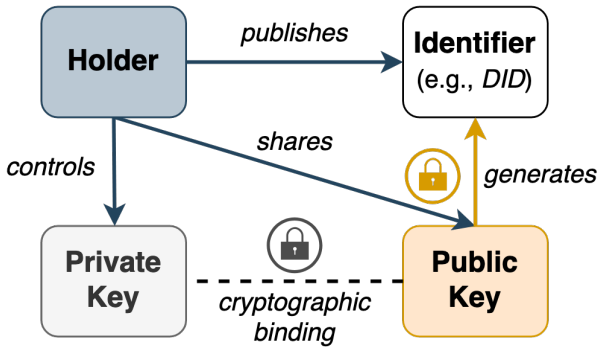


FIGURE 4.40: Self-certifying identifier model, adapted [127].

For this research the term identifier indicates a uniform resource identifier (*URI*), which is a string of characters used to represent physical resources that are not network-accessible (*i.e.*, persons, locations, etc.), as well as logical representations of objects retrieved from an information system, such as electronic documents or any other digital asset [190]. A *DID* is a *URI* management schema (see Figure 4.34) meant to be a component of larger information systems built around the verifiable credential paradigm presented in subsection 4.4.1. Although the design of *DID*s is driven by wider goals (see Table C.1), they represent a solution to enable key rotation and recovery when using self-certifying identifiers [127, 133].

From a design perspective this is specially relevant for data piggybacking. Traditionally, copies of digital resources have been used for data sharing, which implies coordinating numerous entities on how to generate credentials, authenticate identities and regulate access control to digital environments where resource versions are stored.

The result are networks of networks of co-referenced resources with complex resolution requirements [116]. This means it is becoming difficult to ensure the visibility of information stored in varying formats and authenticated through different methods. Also, public institutions face the risk of failing at monitoring the behavior of entities operating with data in these networks. An example for the research context is the visibility of records stored in blockchains representing the state of logistic processes and agreements digitally, such as the *eB/L* covered in subsection 2.6.3. If both trusted identifiers and discoverable resources could

be implemented through a universal schema, the data sovereignty and interoperability barriers set by semantic heterogeneity could be reduced [64, 98]. Thus, credential authentication and access control based on *DID*s can help public institutions piggyback on digital resources operated by logistic service providers in different digital platforms.

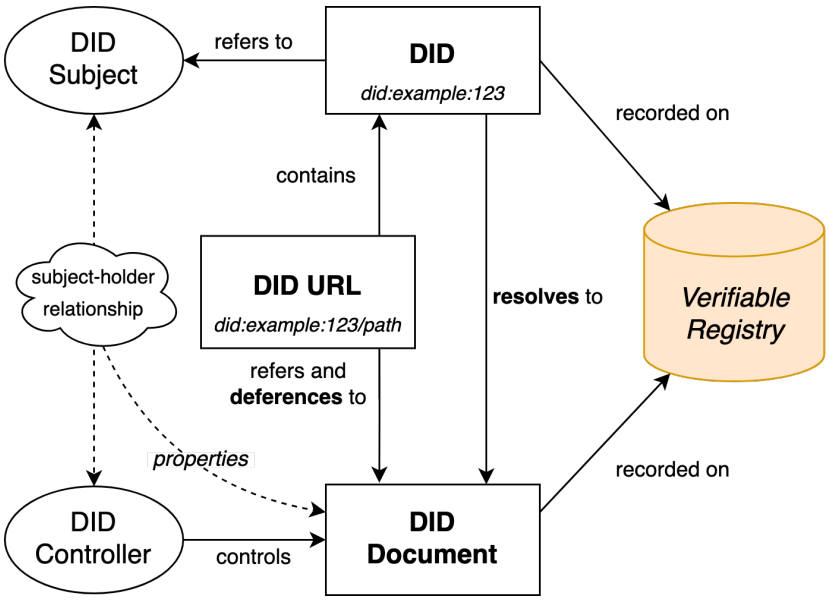


FIGURE 4.41: *DID* architecture accounting for subject-holder relationship, adapted [133].

The *DID* architecture is shown in Figure 4.41. A *DID* controller and subject are equivalent to a credential holder and data subject respectively. There can be more than one controller, who can be the subject at the same time. *DID*s are said to resolve to a *DID* document describign how to securely interact with the owner of a *DID* without revealing information about the entitiy that operates it. This process consists on producing a verifiable presentation of a credential by defining verification methods (such as the aforementioned methods for public-key authentication) and services needed to interact with the *DID* subject. Depending on the application use case, these services can include endpoint discovery, communication routing, access control and resource storage. This way, each *DID* document is equivalent to an identity certificate issued by a traditional *PKI* with added functionalities [133].

*DID*s also support the *DID* delegates, which are entities to whom a controller has granted permission to execute a verification method associated with its *DID* via a *DID* document. This is an additional subject-holder relationship useful for supply chain data piggybacking, because supply chain actors can delegate the verification of each other´s credentials once they have been shared to other parties in other platforms.

For the researched application, participant gateways commit to a *DID*-based credential format. The *DID* documents produced throughout the resource transfer protocol are resolvable by a set of standardised services supported by the overlay infrastructure presented in subsection 4.3.3. Discoverability is achieved with the *pub-sub* system of Figure 4.23, while its associated credential registry is responsible for access control. The protocol messaging routing is maintained by the gateway network. Off-chain resource storage is platform specific, although the cross-chain logs presented in section 4.2 store the verifiable presentations of transactions that feed the cross-chain application logic.

In order to retrieve resource information a uniform resource locator (*URL*) is used, which is a type of *URI* to define the network location of a resource [190]. Dereferencing is used to extract *DID* subjects, verification methods and any component of a *DID* document using *URL*s as input. A *URL* locates a resource, which will be accessed following resource transfer protocol (see subsection 4.3.4) and authentication the properties of *DID* documents. This process is depicted in Figure 4.42 and can be interpreted as a third instance of the trust triangle shown in Figure 4.37, although in this case every step interacts with a distributed information registry.

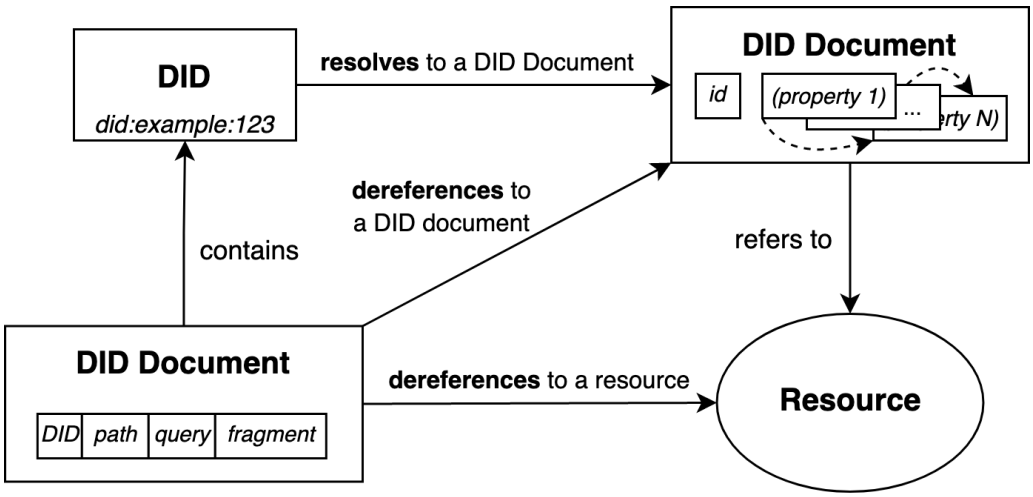


FIGURE 4.42: Overview of *DID URL* deference, adapted [133].

The syntax of *URL*s and the dereferencing specification is provided by the *DID* method used. A method is a description of how *DID*s and *DID* documents are published, resolved and maintained, which is directly linked to the verifiable data registry used. Independently of the technology behind the registry, and as long as some general requirements are met (see Appendix C), a *DID* method can become compatible with the computing infrastructure trusted by a group of entities. This includes *DLT*s, blockchain protocols, peer-to-peer networks or any distributed network or database [133].

Another advantage of implementing *DID*-based trust for information sharing, is that the model does not depend on a particular cryptography for the interpretation of *DID*s. Therefore, the trust model can be implemented as an additional layer on top of legacy systems that rely on identifiers based on the centralised or federated paradigm. Appendix C includes a detailed overview of *DID* architecture.

4.4.5 Gateway Signatures

The messages exchanged between gateways when executing the protocol require their digital signatures, not only to identify, but also to trust that the operators of the gateway will follow the prescribed protocol [70]. The identity authentication of the gateways composing the overlay network must be common between all gateways using the infrastructure. An example of the models used in literature for gateway authentication is shown in Figure 4.43.

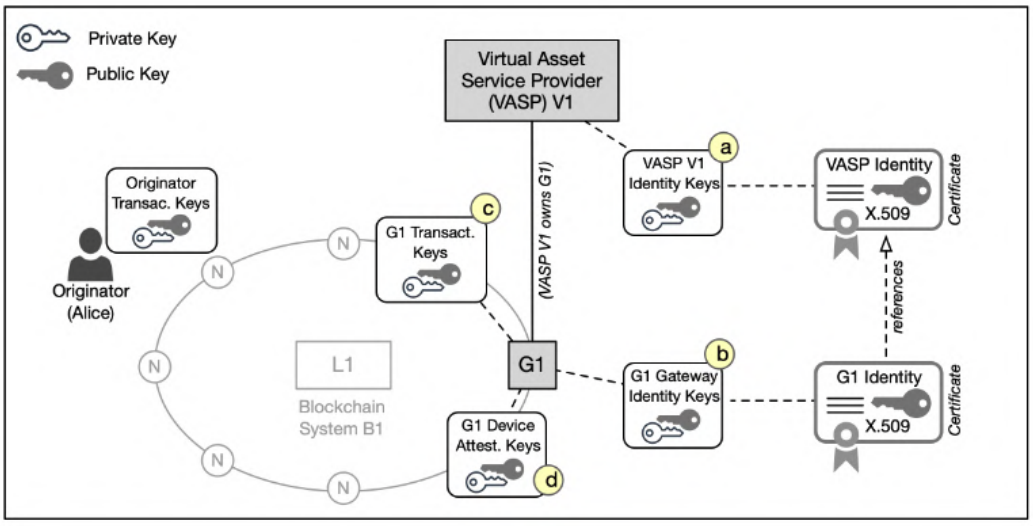


FIGURE 4.43: Gateway identity, key-pairs & certificates [70].

This approach interprets a gateway as an extension of an entity completely independent to the rest of the nodes (*VASP*). It differentiates between an identity certificate for the entity as legal person and another certificate to interact with other gateways. Hardware-bound keys are also used to certify the physical device used by the gateways, decreasing the vulnerability to attacks [167]. Assessing the implementation of these keys is beyond the research scope.

As mentioned in subsection 4.3.3, it is possible for a gateway to represent one or multiple nodes. That is why, in addition to the key-pairs and certificates used to interact with other gateways, internal key-pairs must be used to interact with internal nodes [70]. This is implemented in their design as *G1 transaction keys*, which are used when the gateway interacts directly with the ledger. This approach is useful when public publishing is used (see Figure 4.25). However, since public publishing can not be assumed for all platforms, a general certificate model compatible with all state view boards is shown in Figure 4.44.

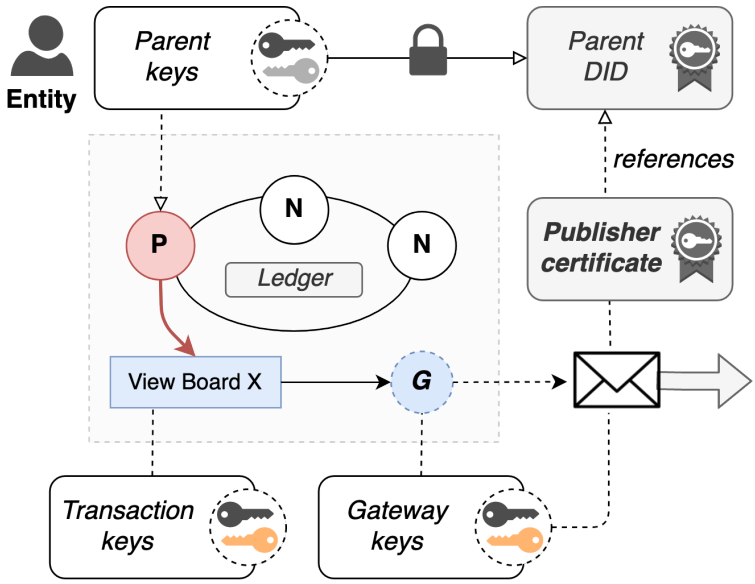


FIGURE 4.44: Gateway identity model.

In the model proposed for the architecture a new publisher/subscriber certificate is produced when the gateway is used. This certificate references the *DID* of the parent node. At the same time, the parent node produced a *DID* document of the ledger state to be shared using his *DID*, which would be sufficient to create a verifiable presentation. Also, the rest of signatures contained in the ledger state being shared is included in the view board item retrieved by the gateway, which can be accessed after the successful resolution of the *DID* document. This approach allows the parent node to define the verification method used, both for his identity and for his view of the ledger, while ensuring visibility of the ledger state and respecting the platform configuration.

4.4.6 Dynamic Cross-chain Authentication

The decentralised credential management framework presented in the previous subsections has assumed cryptographic properties that ensure the authentication of ledger states and identities. In order to take part in a decentralised application and provide trusted state views of a ledger, an entity must be able to show verifiable attributes. The *DID* model allows to define these attributes in the most convenient format for its implementation context. For the research, these attributes should follow the same authentication approach used in blockchains: cryptographic accumulators that generate membership proofs to validate transactions while controlling the visibility of the data and entities involved [184].

A membership proof is a type of zero-knowledge proof of knowledge: "*a prover convinces a verifier that some statement holds without revealing any information about why it holds. A prover can for example convince a verifier that a confidential transaction is valid without revealing why that is the case, i.e., without leaking the transacted values*" (pp. 319) [23]. Cryptographic accumulators achieve this by converting a finite set of values into a unique accumulator value Acc_X and a membership witness wit_x for each element x of the accumulated set [108]. Membership can be certified by providing a valid tuple (x, wit_x, Acc_X).

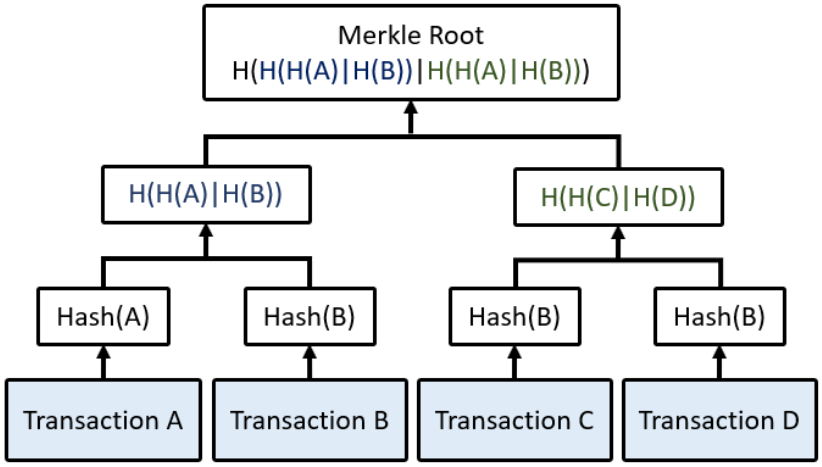


FIGURE 4.45: Structure of a merkle tree [34].

The simplest accumulator is the Merkle tree shown in Figure 4.45, which is a hierarchical data structure with the shape of a binary tree [20]. It uses hash functions to combine data of an arbitrary length into values of fixed length [34], so that no function can be used to reverse engineer the process and obtain the original data [29]. Merkle trees are able to aggregate membership proofs for a batch of transactions into a single constant-size proof (merkle root) [20]. They are used in blockchains to store tamper-evident information about the order of transactions contained by each block, which is shown in Figure 4.46.

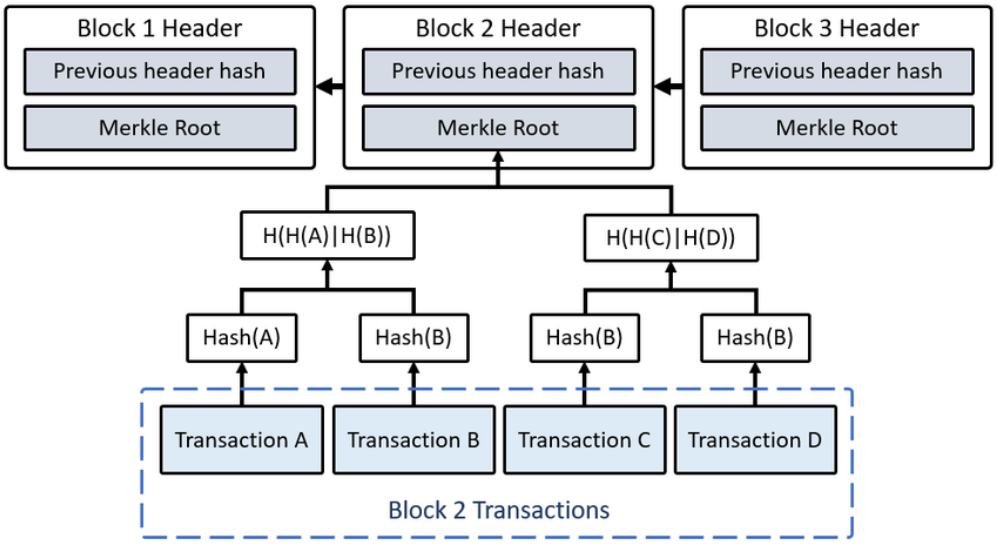


FIGURE 4.46: Use of a merkle tree in a blockchain [34].

A disadvantages of merkle trees is that they are static [20], *i.e.*, it is not possible to add or remove values from the accumulated set without recomputing the membership witnesses for the remaining values. Dynamic accumulators offer features to overcome this challenge. An accumulator is dynamic if "*membership proofs can be updated efficiently as elements are added or removed from the set, at unit cost independent of the number of accumulated elements*" (pp. 561) [20]. This means that it is possible to add or remove values from the accumulated set without having to recompute the witness values of each remaining element [46]. Dynamic accumulators are used for the revocation of anonymous credentials [16, 26] and the design of more efficient and scalable *PKIs* [135, 199]. They are also used in blockchain systems to generate interoperable validity and consensus proofs [20, 192].

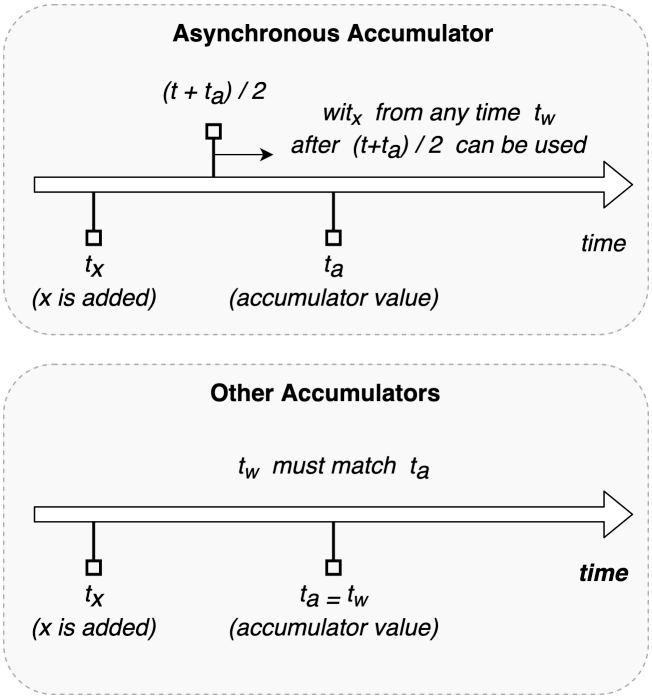


FIGURE 4.47: Comparison of membership witness constraints, adapted [135].

There are two additional accumulator characteristics that are very relevant for the research: low update frequency and old-accumulator compatibility. A dynamic accumulator with these properties is introduced by [135] as an asynchronous accumulator. Unlike other accumulators, the asynchronous accumulator does not require to perfectly synchronise membership witnesses with the accumulator. This is depicted in Figure 4.47.

Low update frequency means that verifying a value against a witness older than the current accumulator value is possible due to laxer witness time constraints. On the contrary, old-accumulator compatibility means that membership witnesses can be used against an outdated accumulator value (given that the latter already contains the element being verified). The low instantaneity requirements of these features allow to overcome the aforementioned limitations of static accumulators when used in decentralised environments [199], because prior knowledge about the accumulated set is not needed to execute element additions [135].

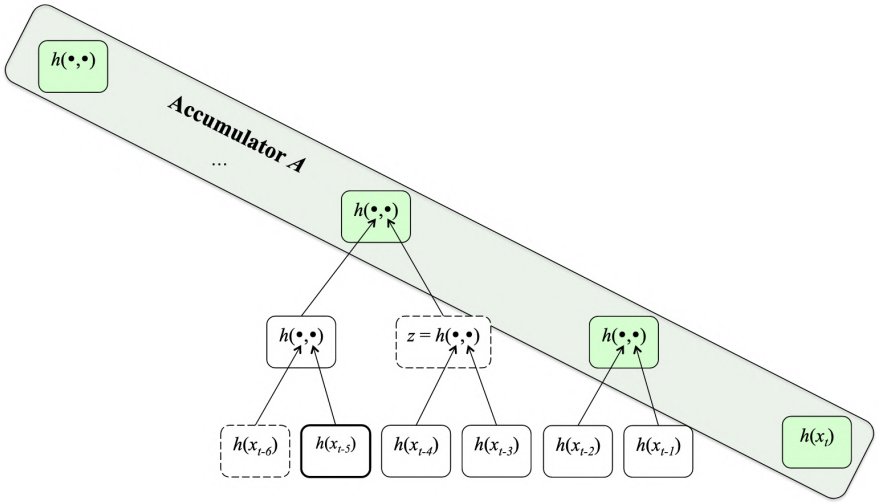


FIGURE 4.48: Structure of a merkle mountain range [135].

An example of an *MMR* is shown in Figure 4.48. The main difference between traditional merkle trees and the ones used in asynchronous accumulators is the hierarchy for the creation of hash nodes and leaf nodes (bottom nodes). The sequence in which nodes are added is shown in Figure 4.49 with the peak coloured in green. Instead of building a complete binary tree with a single root from a predefined group of leaf nodes, the peaks of multiple incomplete trees are used to form a structure known as *Merkle Mountain Range* (*MMR*) [199].

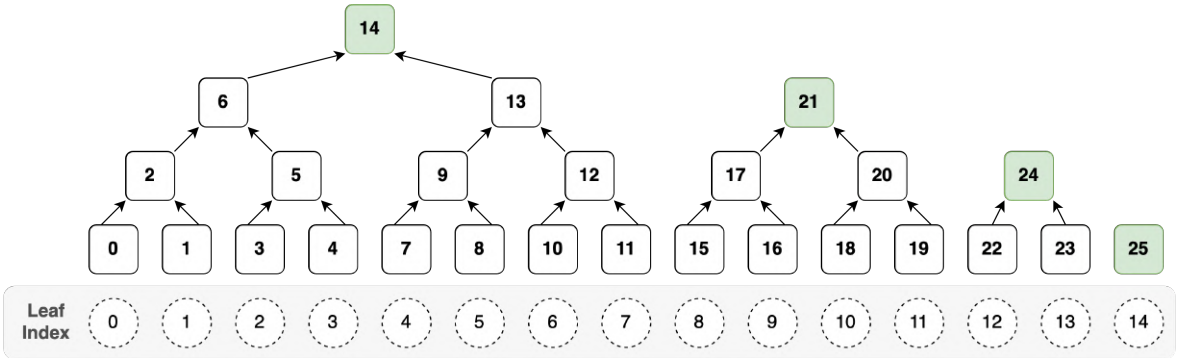


FIGURE 4.49: Merkle mountain range hierarchy, adapted [192].

The accumulated set is the peaks of the incomplete trees being concatenated. This way, the authentication paths for nodes in lower levels remain accessible because it is sufficient to show proof of a peak in addition to the merkle proof of a specific incomplete tree. Multiple incomplete trees can be added by different entities without compromising the verifiability of the information included in existing trees.

Asynchronous accumulators based on *MMR* are thus very convenient for the design in two ways. First, ledger state proofs between platforms can be aggregated using and arbitrary order. Cross-chain transactions can be validated by different entities at different points in time, while still allowing them to add information to the *MMR* tree dynamically. This opens the possibility of using ledger designs beyond linear block sequences to validate decentralised application logic. An example is the *DAG* ledger implemented in the event visibility layer and presented in section 4.2.2. Also, data piggybacking becomes less constrained, because entities can use witnesses that have been updated with varying frequencies to trigger the *HTLCs* discussed in subsection 4.3.4.

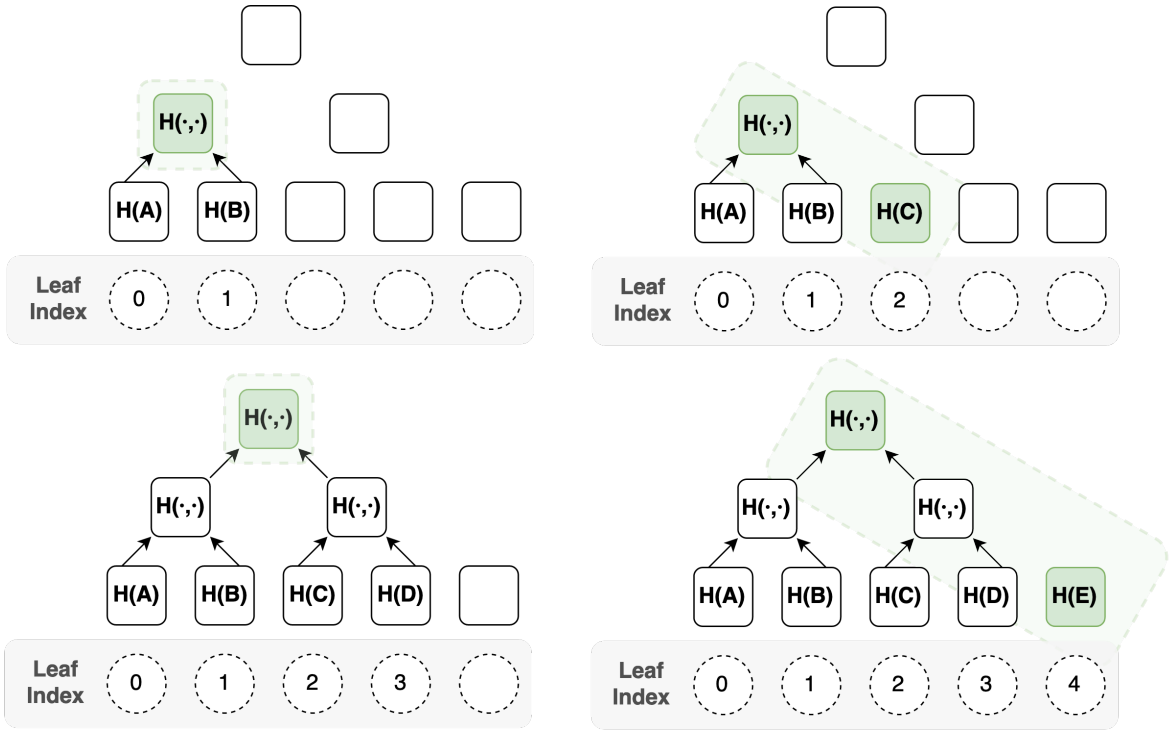


FIGURE 4.50: Sequential change in accumulated set (green area).

Second, the problems caused by key rotation and certificate revocation shown in Figure 4.39 can be avoided. Given a group of entities sharing access via a *pub-sub* subscription, their permissions can be updated individually without affecting the rest of the group. Such dynamic access rules that ensure backwards membership compatibility can be implemented as properties of the *DID* model used for credential management. This means that an entity can generate a credential linked to a *DID* that remains functional to the specified verifier(s), while updating or issuing more credentials.

Despite the advantages of using asynchronous accumulators to aggregate proofs between platforms, it is necessary to define the verification class required by the application and clarify the verification class attainable with the proposed approach. There are four types of cross-chain verification, or verification classes: verification of state, verification of state

agreement, verification of state evolution and verification of state validity [193]. The relationship between these classes is shown in Figure 4.51. These classes are rooted on the idea that a party P on platform X might hide information from a party Q on platform Y, but should not be able to trick Q into validating an incorrect state in ledger L_X [193].

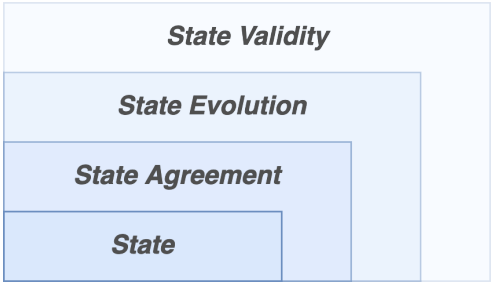


FIGURE 4.51: Cross-chain verification classes, adapted [193].

This can be ensured at different levels by the verification classes. Verification of state consists on certifying that a specific state exists, although it might not have been accepted by consensus. Depending on the consensus mechanism used to maintain a ledger, it can be proved that consensus on a state has been reached. This is the verification of state agreement. Verification of state evolution certifies that a transaction is part of a ledger state. This is achieved by providing a merkle tree path to the leaf node containing the transaction in question. Finally, the verification of state validity ensures that a badge of transactions complies with the internal rules of the source ledger, *e.g.*, there are no conflicting transactions.

Trusted relays based on the exchange of protected ledger views can only verify that a transaction was executed on an external blockchain. This level of verification is considered sufficient for the proposed application for two reasons: the *pub-sub* model and pseudo-proofs for state validity in smart contracts. The *pub-sub* system implies certain trust on the validity of transactions from publishers during the registration phase. Also, pseudo-proofs for state validity can be generated using the concepts of ledger *persistence* and *liveness* defined by [193] and shown in Figure 4.52 and Figure 4.53 respectively:

(Persistence). *Consider two honest parties P, Q of a ledger L and a persistence (or "depth") parameter $k \in \mathbb{N}$. If a transaction TX appears in the ledger of party P at time t, then it will eventually appear in the ledger of party Q at a time $t' > t$ ("stable" transaction). Concretely, for all honest parties P and Q, we have that $\forall t \in \mathbb{N} : \forall t' \geq t + k : \mathsf{L}^P[t] \preccurlyeq \mathsf{L}^Q[t']$, where $\mathsf{L}^P[t] \preccurlyeq \mathsf{L}^Q[t']$ denotes that L^P at time t is a (not necessarily proper) prefix of $\mathsf{L}^Q[t']$ at time t'.*

FIGURE 4.52: Definition of ledger persistence [193].

(Liveness). *Consider an honest party P of a ledger L and a* liveness delay parameter *u. If P attempts to write a transaction TX to its ledger at time $t \in \mathbb{N}$, then TX will appear in its ledger at time t', i.e., $\exists t' \in \mathbb{N} : t' \geq t \wedge \mathrm{TX} \in \mathsf{L}^P[t']$. The interval $t' - t$ is* upper bound *by u.*

FIGURE 4.53: Definition of ledger liveness [193].

The persistence depth k and liveness delay u can be included as variables of the smart contracts used in the cross-chain protocol presented of Figure 4.24. When chosen by the contract participants, if the conditions set by these parameters are not verifiable via the state view board, the publisher-subscriber updates will not be executed. Depending on the decentralised application that a state view will be part of, persistence and liveness proofs can be submitted by parent nodes (see subsection 4.3.3) if required. The publisher is thus responsible for adapting to the publishing configuration required by his platform (see Figure 4.25), and communicate with peer and agent nodes to ensure that the required proofs are present in the state view board that feeds the cross-chain application.

4.5 Design Conclusion

This chapter was aimed at answering RQ4: *What architecture components can used by customs administrations to gather declaration data stored in multiple commercial blockchains?* To do so, the previous sections offered an extensive description of a data sharing architecture to choreograph information sharing between the participants of different blockchain platforms. A summary of the design is shown in Figure 4.54. The architecture includes three layers, each of which provides support to the requirements specified in chapter 3. An overview of the components and their main functionalities can be found in Table 4.3.

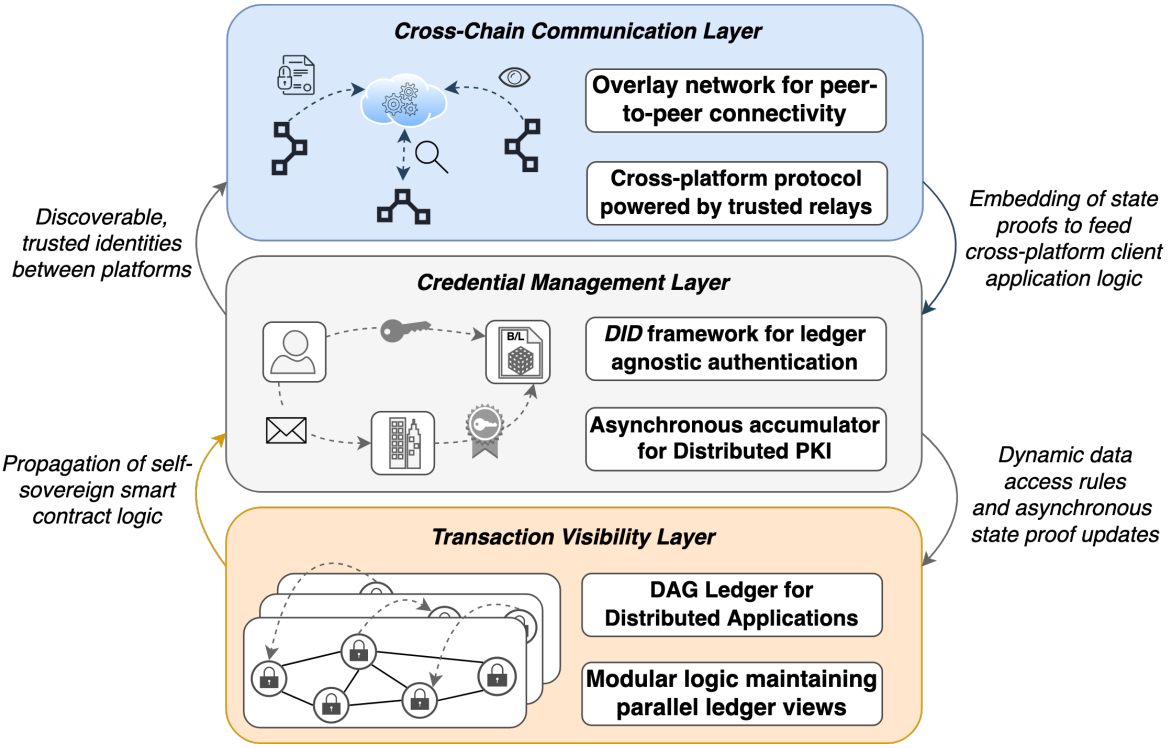


FIGURE 4.54: Architecture design summary.

TABLE 4.3: Summary of design choices and driving functions.

architecture component	*driving function*
Overlay Network	Cross-platform peer-to-peer communication
Trusted Gateway Protocol	Exchange of ledger states proofs to feed the application logic of cross-platform clients
Hash Time-lock Contracts	Propagation of self-sovereign smart contract logic
State View Board	Plug-compatible permissioned ledger state observation
Decentralised Identifiers	Self-sovereign identity management
Asynchronous Accumulators	Dynamic data access rules and asynchronous ledger state proof updates
Directed Acyclic Graph Ledger	Modular distributed application logic while maintaining parallel ledger views

The relationship between the design principles that guided the development of the architecture is described by the framework shown in Figure 4.55. Each pair of design principles has been found to bring different benefits in the context of supply chain data sharing. Event visibility entails that a larger number of actors across all supply chain segments keep track of the logistic events associated with their economic activities. Interoperability makes it technically possible to exchange data between the information systems used in each supply chain segment. Lastly, data sovereignty is an actor´s ability to control the level of exposure of his own digital identities and business information.

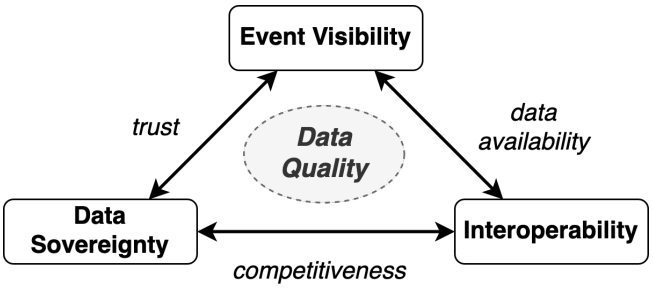


FIGURE 4.55: Design value framework.

Combining event visibility and interoperability produces more potential connections between information systems. Event visibility and data sovereignty elevate trust and incentivise supply chain actors to use externally validated data for business purposes. Finally, interoperability and data sovereignty foster friction-less collaboration by building joint business activities around these trusted data exchanges. The design is an example of the achievement of these three relationships to increase overall supply chain data quality.

Enhanced data quality, and therefore increased reliability during risk assessments, is the ultimate goal sought by European customs. However, it would be naive to take competitiveness incentives and consolidated trust between trade stakeholders for granted. This means that the quality of the declaration data does not only depend on the data infrastructure used by European customs to interact with carriers to lodge an *ENS*, but also on the previous data exchanges between carriers and other logistic service providers. Even if an interface able to automatically collect declaration data based on the operations of a carrier was used, data quality still relies on the information provided to the carrier by freight forwarders and other intermediaries. The design has addressed the challenge of taking into account these data exchanges preceding the generation of the *ENS*.

Chapter 5

Demonstration

The goal of this chapter is to showcase how the data sharing architecture presented in chapter 4 can address the researched problem. As part of the *Demonstration* phase of the design science approach, it is demonstrated that the design can indeed be applied successfully beyond the conceptual plane. By means of a use case aimed at highlighting different aspects of the design, the next sections will try to answer the following research question:

RQ5: *How would the current import declaration procedure be implemented using the specified peer-to-peer data sharing architecture?*

The structure of the chapter is as follows. First, an appropriate demonstration strategy is analysed in section 5.1. Here, the components of a practical scenario required to test the functionalities and effectiveness of the design from all relevant perspectives are discussed. This is very important in order to grasp the real utility of the architecture and address the sub-problems linked to the every design principles and requirements. The context of the use case is covered in section 5.2. It represents a scenario, in which data piggybacking simplifies the *ENS* declaration process and helps European customs administrations to cope with the complexity of *ENS* declaration levels. This is useful when carrier lodge data using irregular *ENS* levels or when vessel deviations occur. The next three sections present the configuration and use of the architecture layers for the use case. The cross-chain network configuration, the application of the credential management framework and the partial distributed application views of each supply chain actor involved are presented in section 5.3, section 5.4 and section 5.5 respectively. Lastly, section 5.6 includes a chapter conclusion.

5.1 Design Demonstration Approach

The components of the architecture have been selected based on requirements generated from three different design principles (see Table 3.1). Each principle represents a functionality that enables the design to provide value to customs administrations. It is therefore possible to accidentally over-represent a specific aspect of the design during the demonstration. However, the goal is to provide a complete picture of how European customs administrations can take advantage of the architecture. Therefore, a design demonstration strategy that equally covers the three design principles with a level of detail matching that of the architecture technical specification is preferred.

Each design principle should be instantiated differently. Logistic event visibility is expressed as sequences of logistic and commercial events, and represent the processes found in supply chains and that European customs is part of. Stakeholder data sovereignty is expressed as networks of subject-holder relationships, and represent the structure of the information flows that support the aforementioned event sequences. Lastly, architecture interoperability is expressed as a group of blockchain platforms involved in these information flows, and represent the connection between stakeholders and their information systems.

A use case can be put together after instantiating each design principle following these guidelines. The next step is translating the proposed scenario to a configuration of the architecture and show how it would function. The configuration is presented at a layer level. For the cross-chain communication layer, the network specification and publication of states will be covered. For the credential management layer, subject-holder relationships will be converted to *pub-sub* relationships governed by smart contract services. Finally, the configuration of the event visibility layer includes an overview of the decentralised applications that can be implemented following the *CAPER* protocol. The application view perceived by each stakeholder will be shown, as well as how the addition of data elements and/or participants is achieved by interacting with the credential management layer.

5.2 Use Case Context

The goal is to aggregate *B/Ls* and generate a modular *ENS* based on paths towards distributed data. Instead of dedicated documents based on the content of previously issued documents, European customs is provided with access to data that is scattered among multiple storage locations. The logistic background of this scenario is shown in Figure 5.1. In the same way a freight forwarder handles multiple products, the carrier will interact with multiple forwarders (dashed arrows), and customs will interact with multiple carriers. Although the benefits of the design are more pronounced with more interactions at each level, a simplified approach with a single forwarder and carrier is used to demonstrate the design.

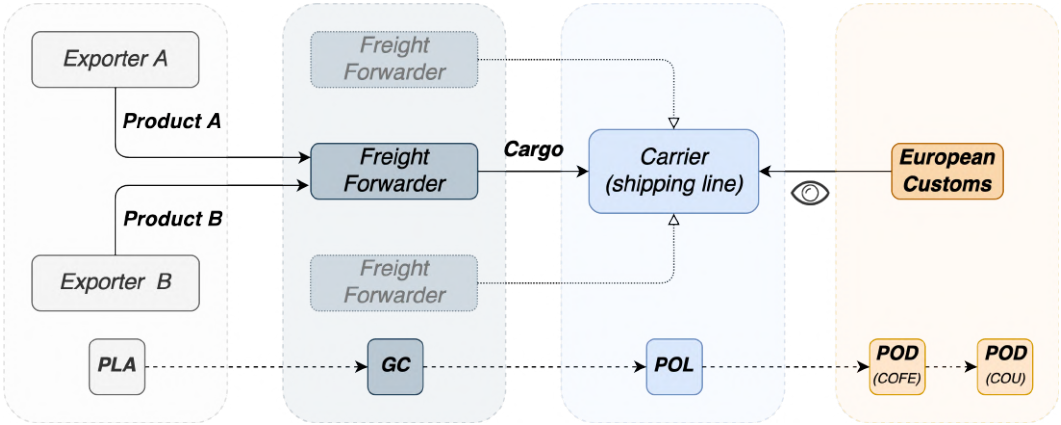


FIGURE 5.1: Logistic context.

Two entities export products from a *PLA* via a freight forwarder, who is in charge of consolidating the products in a *GC*. The resulting cargo will be shipped by a shipping line that acts as carrier from a *POL*, who is also responsible for lodging the entry summary declaration data for the European customs administrations (*COFE*) at a *POD*.

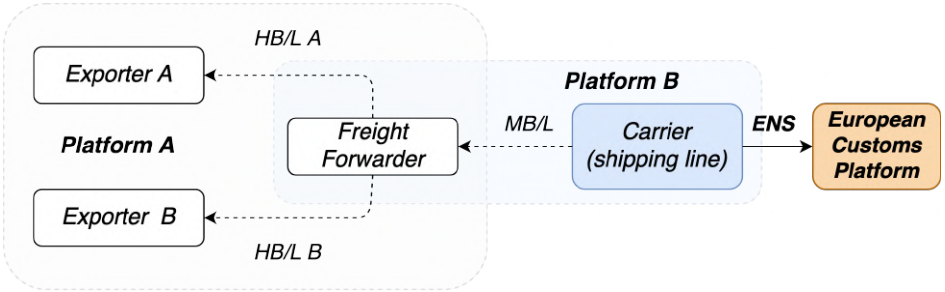


FIGURE 5.2: Platform context.

The exporters use the same platform used by the freight forwarder. This platform handles electronic *HB/Ls* between the freight forwarder and the exporters (consignors). The latter also shares platform with the shipping line, and is used to exchange information about electronic *MB/Ls*. This is depicted in Figure 5.2. It is assumed that European customs operate in a blockchain platform for internal use, where custom offices operate node. The credential management relationships shown in Table 5.1 indicate the role each entity as towards other stakeholders for each transaction.

TABLE 5.1: Credential management relationships.

	Exporter A	**Exporter B**	**Freight Forwarder**	**Shipping Line**	**Customs**
HB/L A	Data Subject		Credential Issuer		
HB/L B		Data Subject	Credential Issuer	Credential Verifier	
MB/L	Indirect Subject	Indirect Subject	Data Subject	Credential Issuer	Credential Verifier
ENS	Indirect Subject	Indirect Subject	Indirect Subject	Data Subject	

The challenge faced in this use case is allowing the shipping line to provide customs administrations with the regulated declaration data in the form of links to information contained in the three previously issued bills of lading. The bills are stored in different platforms. The *MB/L* should be created from links towards *HB/L A* and *HB/L B*. Then, a document equivalent to the traditional *ENS* should be made available to customs.

The information included in the documents must be filtered so that in the case of non-negotiable bills, the necessary data fields are included as a verifiable presentation while omitting any confidential data. Moreover, the data pipeline principles, such as data efficiency, should be followed. Besides sensitive commercial information, data fields not explicitly required downstream the cargo custody chain should be removed. This has two direct benefits. The processing of information can become faster and more efficient at the aggregate level by omitting data duplication in some transactions. This practice can also reduce the risk of data inconsistencies in the event of amendments, which are more difficult to track and correct if data is sparsely duplicated.

5.3 Overlay Network Configuration

The configuration of the cross-chain communication layer is shown in Figure 5.3. It includes the storage of each stakeholder and the link to their platforms nodes. In the case of the freight forwarder, he operates one node in two platforms. This will be shown useful in order to gain access to more than one view of the distributed *DAG* application maintained by all participants. As mentioned before, it is assumed that customs administrations operate a private platform connecting customs offices.

The reason to include a platform dedicated to customs is the nature of the credential management framework. Verifiable presentations require an immutable verifiable data registry to produce or authenticate *DID* documents. The internal configuration of this platform will not be discussed in the demonstration, because customs nodes do not act as a parent node in the proposed use case for *ENS* data collection.

The customs platform has been included for illustrative completeness, as its main role is to maintain a copy of the *DAG* ledger view held by a customs office. It represents the environment where customs offices exchange information with each other, in the same way each commercial platform is used by groups of enterprises to coordinate their business activities.

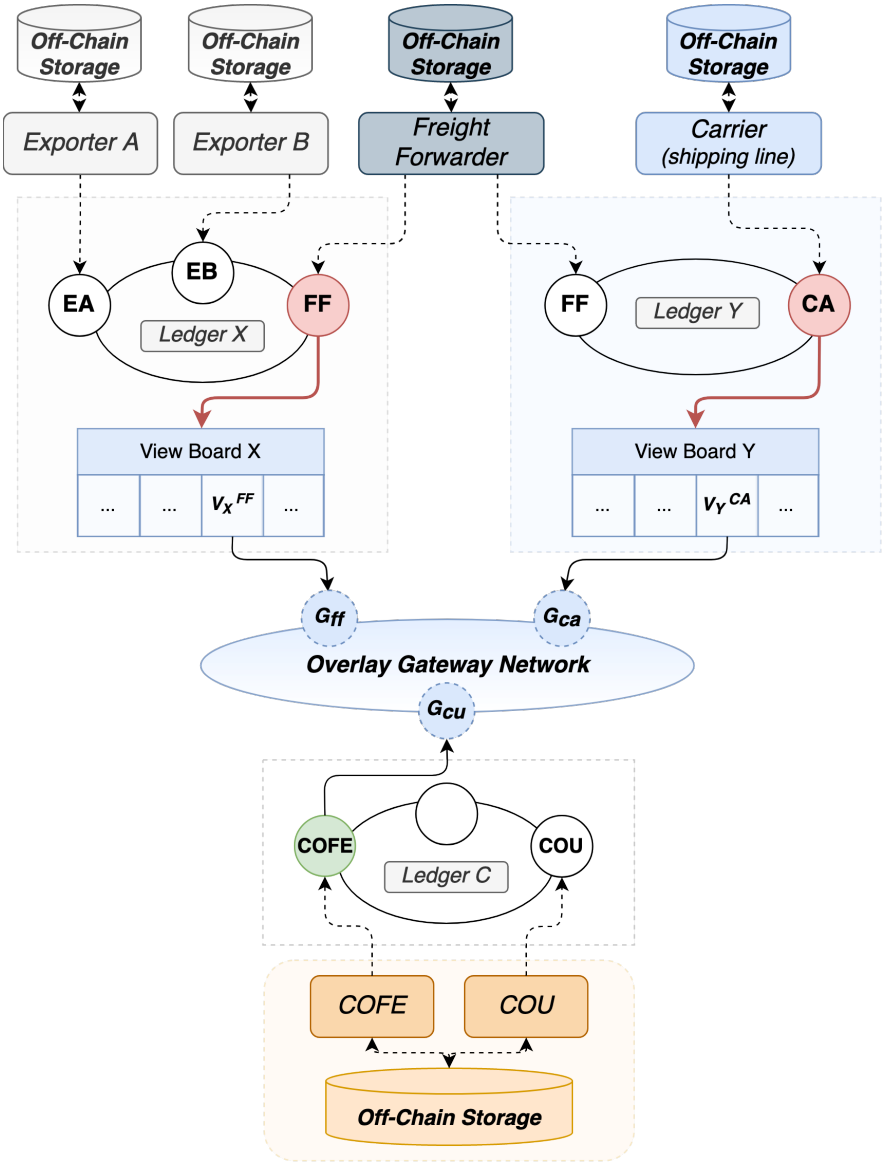


FIGURE 5.3: Applied cross-chain communication layer.

Using the network configuration, the transfer protocol is applied to express the data flow between supply chain actors and customs for the proposed architecture. This means that the freight forwarder will first generate *HB/L A* and *HB/L B*. From the data included, confidential fields will be omitted and the necessary fields will be shared with the shipping line, who will then generate a *MB/L*. The shipping line can then generate *ENS* data based on the data of the *MB/L*s produced for a specific vessel and allow customs to access it.

The protocol is applied in three phases: two application commitments followed by a resource exposure. First, the freight forwarder shares verifiable presentations of relevant *HB/L* transactions from platform *A* and registers them in the *DAG* ledger (Figure 5.4).

Once that the carrier has piggybacked on the verifiable presentations published by the freight forwarder, these presentations will be attached to the verifiable presentation of the *MB/L* and left accessible for customs (Figure 5.5). Lastly, customs administrations will use this verifiable presentation to execute the resource exposure, obtain the information on how to interact with the storage of the carrier and pull all relevant *MB/L* data (Figure 5.6).

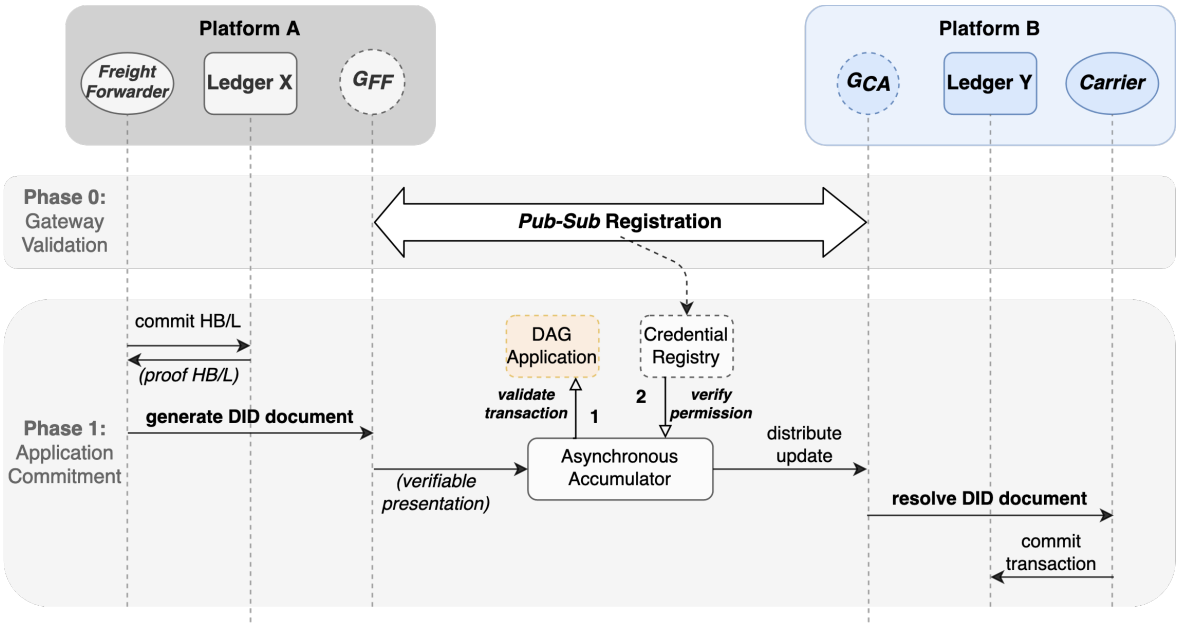


FIGURE 5.4: Application commitment between freight forwarder and carrier.

Note how after the application commitment between freight forwarder and carrier, instead of directly moving towards the resource exposure phase, the protocol proceeds with the second application commitment between carrier and customs. The last transaction commitment in Figure 5.4 activates the issue of the *MB/L* in platform *B*. This is how the logic of different ledgers is propagated. In this case, the carrier is using ledger state proofs the generated by the freight forwarder in platform *A* to trigger a response in platform *B*.This is an example of the smart contract chains discussed in subsection 4.3.4.

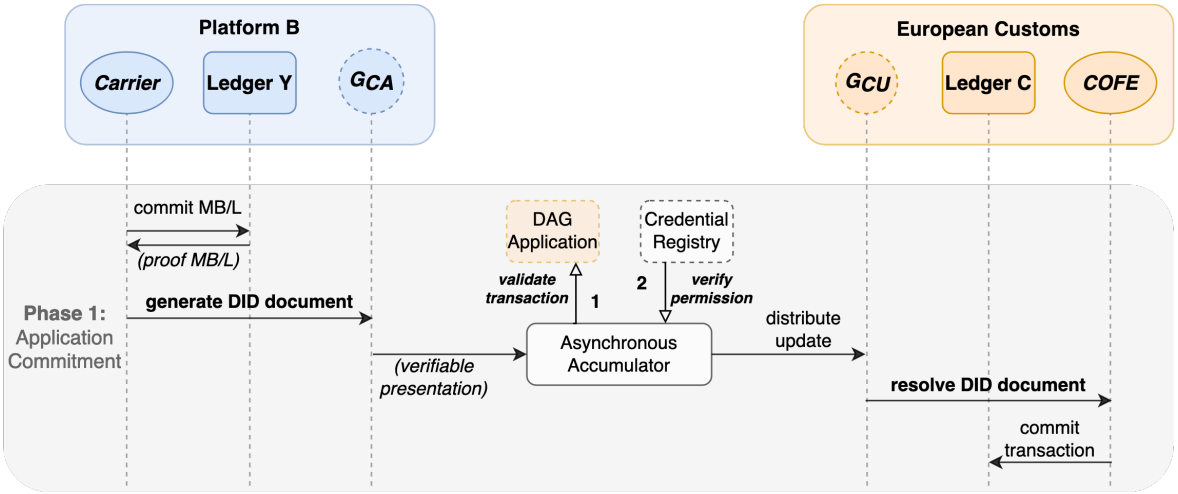


FIGURE 5.5: Application commitment between carrier and customs.

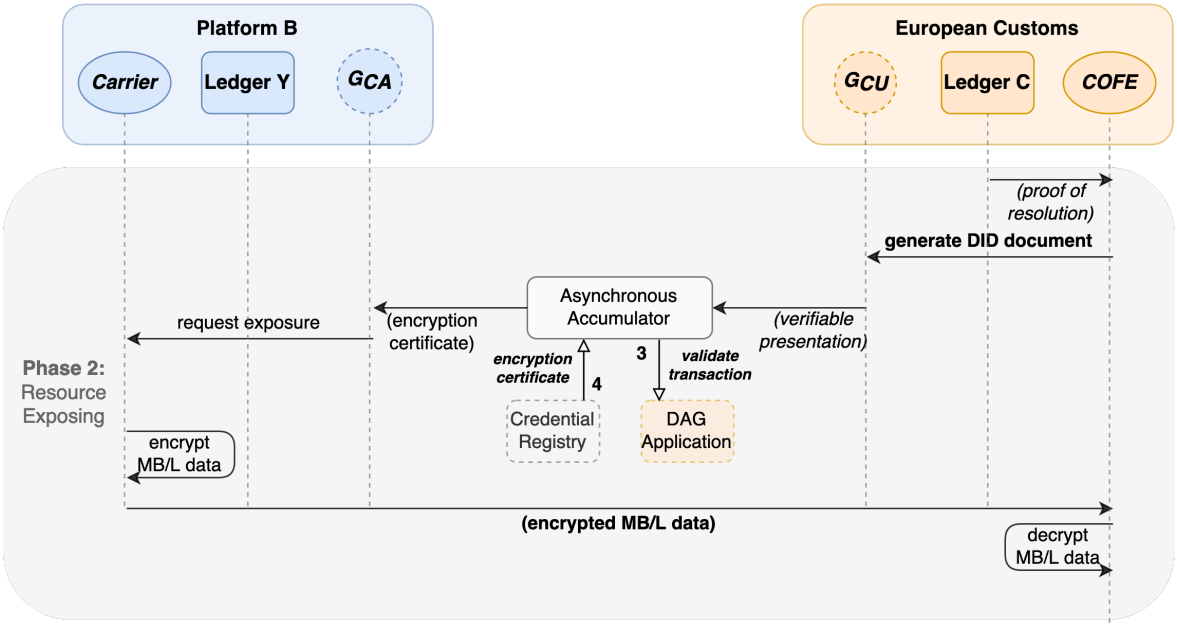


FIGURE 5.6: Customs resource exposure.

Additional resource exposure phases can be added depending on the length of the data exchange chain. In the proposed use case, the latter consists of only two data intermediaries between European customs and the original cargo custodian. However, in case that investigations are performed after the *ENS* lodging, European customs might require the carrier to provide additional information. Also, European legislation offers the carrier legal protection against previous cargo custodians withholding this data (see subsection 2.4.2), which would allow the carrier to reveal the issuers of previous *B/L*s linked to the cargo under investigation, who would eventually be obliged to provide the additional information.

This is another scenario in which the proposed protocol becomes a useful resource. As shown in Figure 5.7, instead of enforcing this legislation upstream a supply chain and relying on the carrier to provide this data, the freight forwarders have already provided verifiable links towards more detailed cargo information. The carrier has forwarded these links within the *ENS* enabling European customs to access additional information.

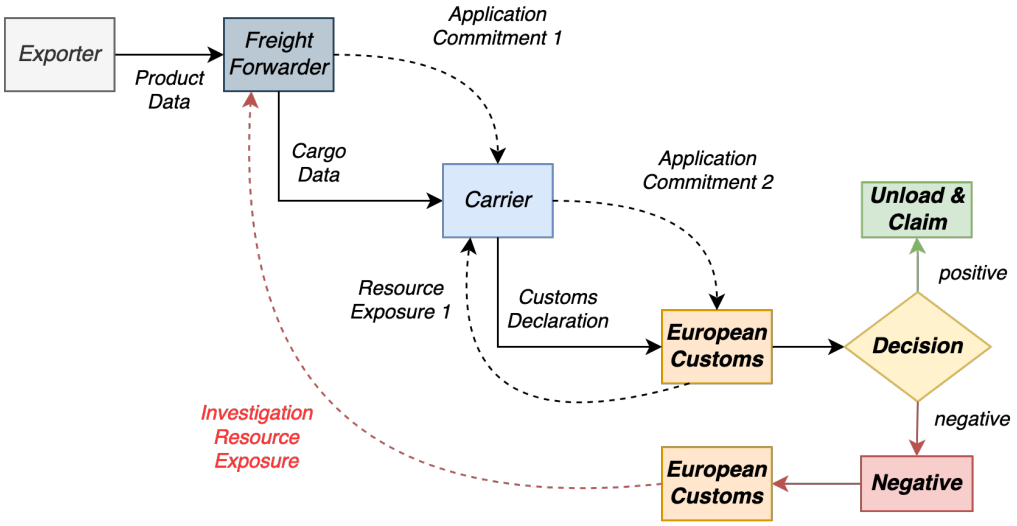


FIGURE 5.7: Investigation with additional resource exposure.

From a bureaucratic prospective, this process is more efficient. The technical characteristics of the architecture make the declaration process compatible with potential investigations *a priori*, which accelerates the execution of customs procedures that tend to delay the release of cargo in busy port terminals. Evaluating how the previous cargo custodians can ensure the validity of these links is outside the scope of this demonstration.

5.4 Credential Management

The messages exchanged in the protocol take the verifiable presentation format. Figure 5.8 shows an example for the ledger state published by the freight forwarder. The verifiable presentation includes two relevant transactions of the ledger state in the form credentials. In this case, it attaches a proof graph of the rolling hash from the state view board accessed by the gateway to certify the consensus on the state.

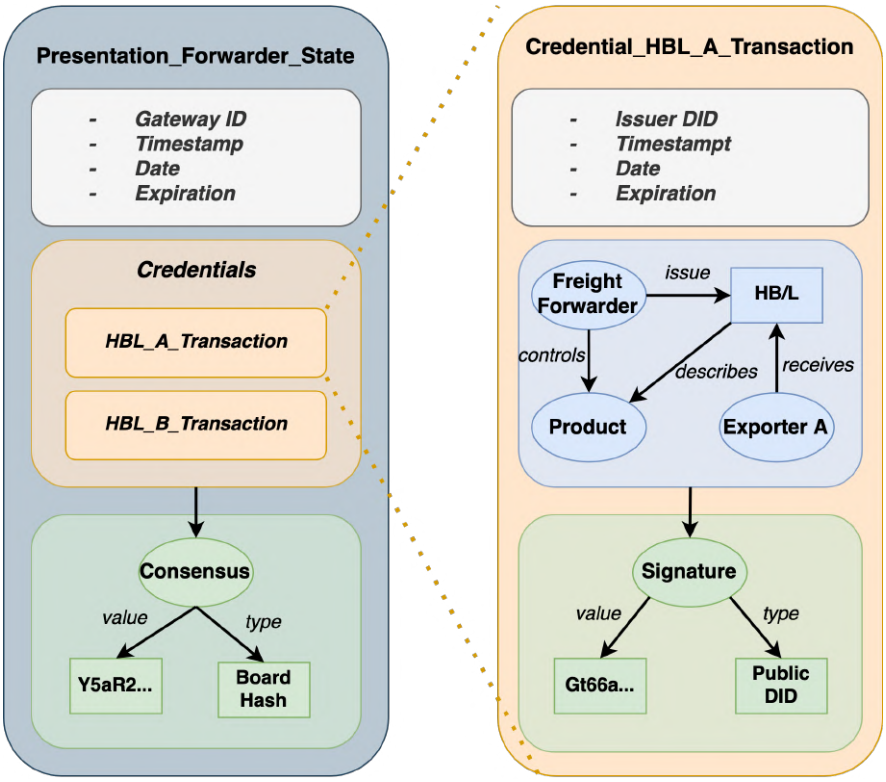


FIGURE 5.8: Verifiable presentation freight forwarder.

Each credential includes a *RDF* statement describing the event that the transaction represents. An illustrative example of a *RDF* statement about the issuing of a *B/L* has been added. It is complemented by the public *DID* signature used by the freight forwarder to sign transactions. This signature will be used by future client nodes, who must proof their intention or ability to use the credential as a dependency. In this case, such client is the carrier, who will use the credential and attach it to the verifiable presentation of the *MB/L* as shown in Figure 5.9. This verifiable presentation includes all the necessary information to describe the cargo custody chain from the current carrier to the original exporters.

The reason why these two verifiable presentations can be used to complement each other is the architecture of the gateway identity model. In the same way that the freight forwarder published the internal state of ledger X including the proofs for a bundle of transactions he

took part in, the carrier is publishing the internal state of ledger Y including the proof for a transaction that the freight forwarder took part in. A trace of the forwarder´s *DID* signature was left in this transaction.

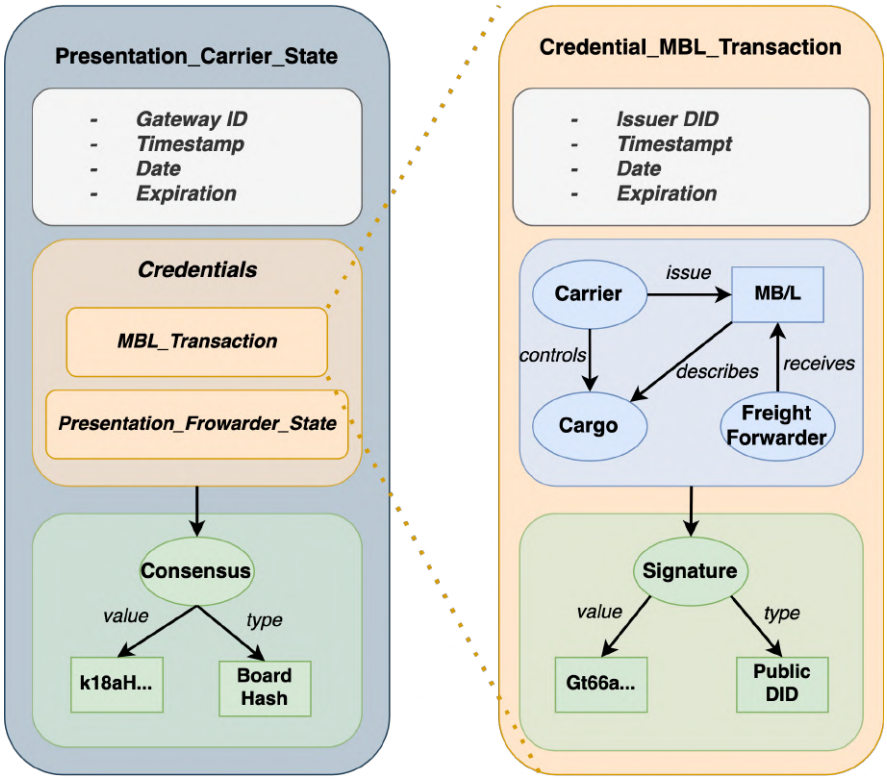


FIGURE 5.9: Verifiable presentation of carrier.

Therefore, the carrier can proof the relationship between the peer on his local ledger and the peer operating the gateway against the metadata of his own transaction. That is how the carrier is able to validate the presentation published why the freight forwarder from platform *A*. The cross-chain *DID* validation process is shown in Figure 5.10.

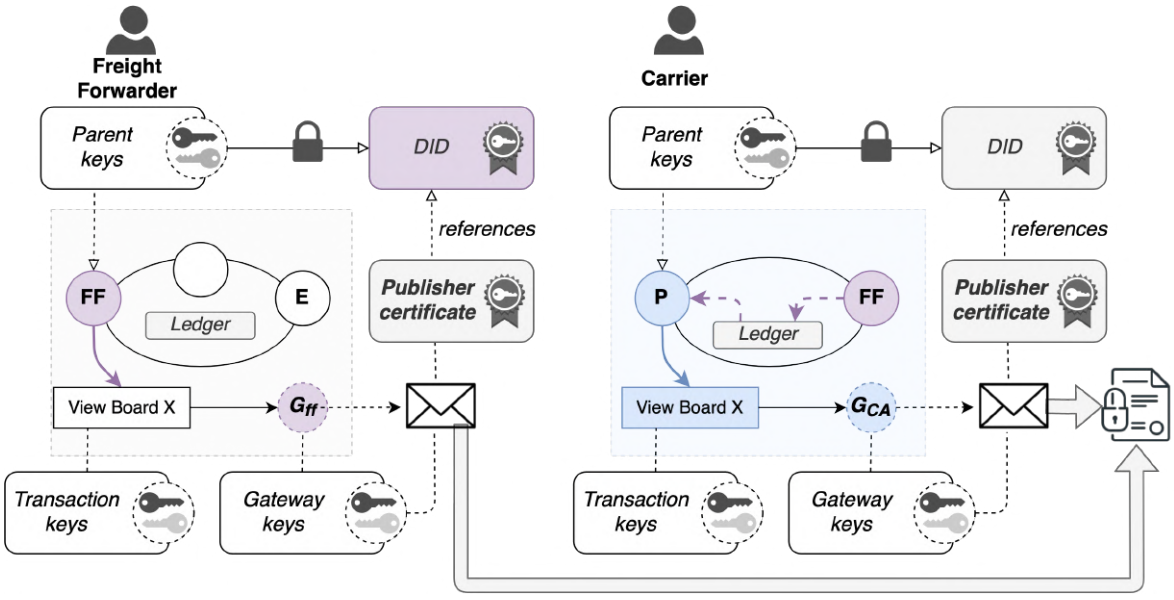


FIGURE 5.10: Cross-chain DID validation.

5.5 DAG Application

In the event visibility layer a *DAG* is able to keep track of the validation of cross-chain state views. As shown Figure 5.11, the resulting *DAG* for the use case is relatively simple.

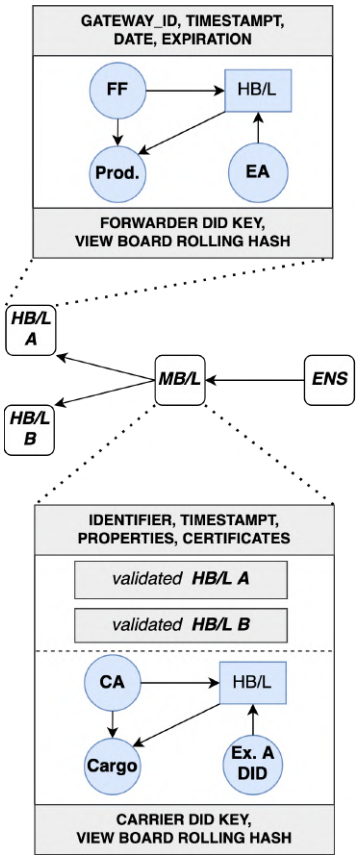


FIGURE 5.11: Resulting *DAG* ledger.

Since the use case only has a forwarder and a carrier, the benefits of using different views of the ledger for different users is less evident. Also, the utility of the asynchronous accumulators for the addition of elements is more pronounced when using a larger *DAG*. Therefore, the *DAG* formed when the use case is applied multiple times with the same carrier and different freight forwarders is shown in Figure 5.12.

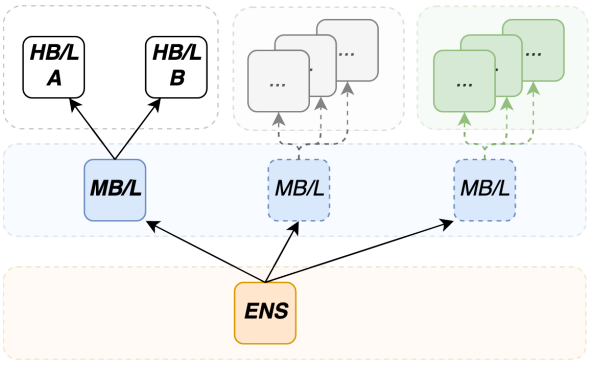


FIGURE 5.12: Perspectives of *DAG* ledger.

A participant in the *DAG* stores a view of the levels he takes parts in. For example, the freight forwarder will only see the *HB/L* nodes he added, so their *MMR* representation will be used as proof for that level (Figure 5.13). Similarly, a carrier has access to the same records as the forwarder, in addition to the other *MB/L*s on the second level. Therefore, the carrier can concatenate the *MMR* proofs of the different levels and reconstruct the original three. What is added to the *DAG* ledgers are not the verifiable presentations transmitted by the gateways. Rather, each participant stores these proofs per level and update their local accumulator. These proofs will be layer used to validate their additions to the ledger.

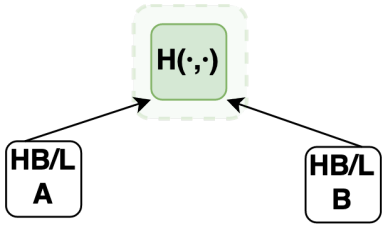


FIGURE 5.13: Freight forwarder cluster proof.

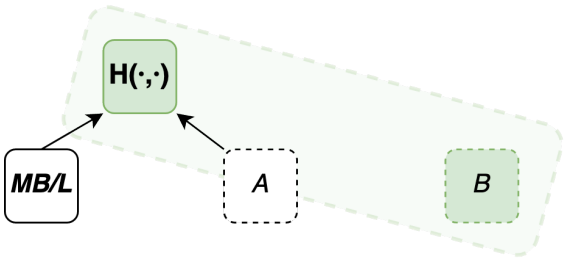


FIGURE 5.14: Carrier level proof.

The aggregation of the *MB/L* level is used to generate the presentation that will be eventually shared with customs. Static accumulators would not allow to create such a ledger, whose topology can be used to model the information flows themselves from the transactions being validated. By storing each cluster within each level using the *MMR* technique, it is possible to grow a *DAG* by adding more elements in a specific level. This is shown in Figure 5.15. Although the proof for the accumulated *MB/L* level has changed by introducing a new element, the authentication path towards *HB/L A* and *HB/L B* is still valid, because the previous accumulated value is part of the new accumulated set. This way the future *ENS* data set can be built gradually by adding more elements on each level.

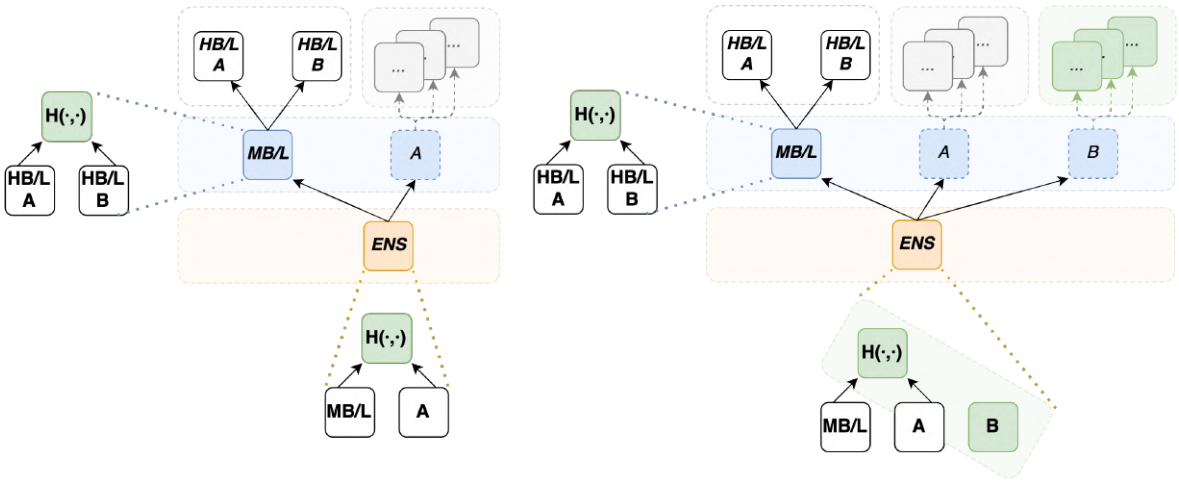


FIGURE 5.15: Addition of leaf records after node validation.

5.6 Conclusion

The goal of this chapter was to answer *RQ5*: *How would the current import declaration procedure be implemented using the specified peer-to-peer data sharing architecture?* This has been done by illustrating how the different architecture layers would be used to create links between *B/L* information stored in multiple blockchain platforms. These links aggregated by declarants to provide European customs with an *ENS* formed by bundles of links to the information necessary to conduct import risk assessments. It has been shown how the proposed architecture can be used to reduce the number of intermediaries that process the logistic data eventually included in import declarations, which can reduce incomplete and contradictory entries in the final *ENS* processed by customs.

Chapter 6

Evaluation

This chapter covers the last phase of the design science research methodology, the evaluation of the architecture. Since this phase is being performed prior to the construction and deployment of the artefact and the application environment will be explored, this chapter presents a naturalistic *ex ante* evaluation [177]. The goals of a design science evaluation can be diverse. Requirement compliance, expert knowledge on the application context, comparison to similar artefacts and assessments of relevant side effects after implementation are common evaluation topics [90].

For the current research, the objective is twofold. The first one is to determine whether the data sharing architecture complies with the requirements specified during the *Requirement Definition* phase. The second one is to evaluate if using the data sharing architecture fits the design principles and satisfies the expectations of stakeholders [179]. Therefore, the following research question is answered by means of requirement analysis and consultation with a commercial blockchain platform expert:

RQ6: *Does the data sharing architecture comply with the requirements to a sufficient extent to be considered a feasible solution that contributes to the application domain?*

The structure of the chapter is as follows. First, the relationship between architecture components and the design requirements is covered in section 6.1. A review of the architecture design by an industry expert is presented in section 6.2. The practical implications and potential improvements derived from the two previous sections is covered in section 6.3. Finally, a conclusion on the design suitability is covered in section 6.4.

6.1 Requirement Analysis

An overview of the design requirements and the architecture components that support them and the that motivated their selection are shown in shown in Table 6.1 and Table 6.2. The following paragraphs elaborate on how the functionalities of these components and their interaction with the rest of the architecture are aligned with the design requirements.

TABLE 6.1: Functional requirements and supporting components.

requirement	*supporting component*
FR1	Overlay Network and Trusted Gateway Protocol
FR2	Directed Acyclic Graph Ledger
FR3	Decentralised Identifiers
FR4	Decentralised Identifiers
FR5	Trusted Gateway Protocol and Hash Time-lock Contracts
FR6	Overlay Network, Trusted Gateway Protocol, Hash Time-lock Contracts and Decentralised Identifiers

TABLE 6.2: Nnon-functional requirements and supporting components.

requirement	*supporting component*
NFR1	Overlay Network and Trusted Gateway Protocol
NFR2	Directed Acyclic Graph Ledger
NFR3	Directed Acyclic Graph Ledger
NFR4	Decentralised Identifiers
NFR5	Decentralised Identifiers and Asynchronous Accumulators
NFR6	Decentralised Identifiers
NFR7	Asynchronous Accumulators
NFR8	Hash Time-lock Contracts and Asynchronous Accumulators
NFR9	State View Board, Decentralised Identifiers and Asynchronous accumulator
NFR10	State View Board and Directed Acyclic Graph Ledger
NFR11	Hash Time-lock Contracts and State View Board
NFR12	Overlay Network, Trusted Gateway Protocol and Decentralised Identifiers

Overlay Network & Trusted Gateway Protocol

An overlay gateway network has been chosen as the backbone of the cross-chain communication of the architecture. This decision allows customs administrations to interact directly with the locations among which original declaration data is distributed (*FR1*). The use of relay gateways to support a cross-platform transfer protocol allows to use a pull mechanism that characterises the data pipeline concept, which has been expressed a strategically desired approach by customs administrations(NFR1). The gateway mechanism is also motivated by the need to respect the current configuration of blockchain platforms, as it can be implemented as an additional layer on top of the existing local networks (*NFR12*). The resource transfer protocol included in the design takes advantage of this network, but does not require discrete point-to-point transfers executed in real-time. This enables customs administrations, and other entities, to passively piggyback on data via links to resources generated and/or stored in a number of blockchain platforms (*FR5*, *FR6* and *NFR1*).

The use of *HTLC*s in the protocol makes it possible for data subjects to encode revocation mechanisms in the links shared with customs and other trade partners. This can be understood as the ability to decide when to break the links, and therefore, exercise the right to erasure and other features that improve the data sovereignty of the ecosystem *NFR8*. Additionally, by linking data transfer rules to the verifiable proofs published directly from platform ledger data, *HTLC*s prevent users from unilaterally providing executing data exchanges using proofs outside their local platform consensus. This avoids the propagation of invalid proofs that could collide with platform consensus states in the future (*NFR11*).

State View Board

Similarly to the overlay bridges, state view boards offer a modular and plug compatible solution to inform a gateway about the ledger states observed by a node. It can be used in a number of platforms without modifying their internal configurations (*NFR12*). Moreover, they serve as a common canvas where multiple platforms users can publish updates regarding the same digital asset (*NFR9*) and arbitrary events registered in a paltform (*NFR10*).

Decentralised Identifiers

*DID*s are the cornerstone of data sovereignty in the design. Their main contribution for the proposed application is enhanced trust between supply chain stakeholders when it comes to information sharing. A *DID* document describes how to securely interact with the owner of a *DID* without revealing any information about the entity behind the *DID* (*FR3* and *FR4*). In essence, it is used to produce self-certifying certificates, which can be understood as decentralised zero-knowledge proof of identity for data exchange.

The limitations of recognisablility are provided by the *DID* core data model, as it allows data subjects to take control over the governance of their own data. They force potential verifiers to present proofs of their relationship towards the credential owner or data subject (*NFR4*, *NFR5* and *NFR6*), as well as to appoint trusted representatives (see *DID* delegates in subsection 4.4.4) that can support the *DID* owner during transactions (*NFR5*, *NFR6* and *NFR9*). This holds for personal and non-personal data, as *DID*s can be assigned to both legal persons and abstract digital objects. Therefore, the *DID* model enables entities operating in different permissioned environments to trust and process assets published by the users of other platforms (*FR6*). The use of the *DID* model is also motivated by its interoperability advantages. It can adopted by legacy systems running on centralised certificate systems with limited additions to their current infrastructure (*NFR12*).

Asynchronous Accumulators

Cryptographic features are always expected to be part of any information system built around a blockchain environment. They are required to maintain tamper-evident registries and achieve consensus between network participants. However, two additional features where required by the researched application: statements should be verifiable by a dynamic group of entities operating in different blockchain platforms, and the topology of these statements should be able to go beyond the linear sequences found in blockchains.

Asynchronous accumulators have been the proposed approach to tackle this two issues. They are first used by the *pub-sub* system to handle the addition and revocation of certificates and access permissions. Such modular access control is very useful in the proposed platform interoperability context, as it allows *ENS* declarants to change their *DID* delegates over time (*NFR5*, *NFR7* and *NFR9*). Additionally, the structure of *RDF* information graphs maintained by these accumulator are dynamic, meaning that the right to erasure (and associated data sovereignty features) can be easily implemented (*NFR8*).

Directed Acyclic Graph Ledger

Choosing *DAG* as the distributed ledger technology to record cross-chain transactions brings immediate benefits in terms of event visibility and interoperability, not only technical but also organisational. When combined with the right cryptographic components (the aforementioned asynchronous accumulators in this case) they can be used by groups of entities to build up information graphs modelled after cross-chain services.

The chosen *DAG* protocol is designed to aggregate these services and allow stakeholders to collaborate. Since each participant maintains an independent view of the application logic, more than one framework agreement can be modelled without compromising the confidentiality of a particular view (*FR2*). This way, trade patterns, such as cargo custody chains, sequences of contracts of carriage, can be combined into an augmented, auditable logic (*NFR2* and *NFR3*). Using the *DAG* protocol to power distributed applications allows to

decentralise the publication of events registered in multiple blockchain platforms (*NFR10*). Moreover, this can be achieved with limited changes to the current architecture of the event sources, either modern blockchain platforms or centralised legacy systems (*NFR12*).

6.2 Expert Validation

TradeLens is the largest blockchain-based platform that provides support to logistics service providers [188]. With more than one billion events and ten million documents processed every year [171], it plays a leading role in the digitisation of trade finance. The large number of ocean carriers and shippers involved in the *TradeLens* ecosystem implies that a considerable portion of the declaration data processed by European customs originates in this platform. *TradeLens* is therefore a great example of a blockchain environment European customs desires to interact with. The architecture has been presented to a representative of this platform to judge its practical value. The details of this interaction is shown in Table 6.3.

TABLE 6.3: Details of the expert validation.

	purpose	*date*	*format*	*institution*
Expert 1	External validation of design and requirements	07/09	E-mail	IBM, representative TradeLens Platform

The purpose of the external validation is not a detailed assessment of the technical feasibility of the design. Rather, it focuses on the suitability of the design components that provide key functionalities not present in current solutions. It also intends to show how the design succeeds at aligning the interests of commercial blockchain platforms and European customs. These components include the *DID*-based verifiable credentials used in the credential management layer and the *DAG* ledgers used in the event visibility to create distributed applications where multiple platform users can collaborate.

First, the functional and non-functional requirements presented in chapter 3 have been validated by *expert 1*. These requirements are considered realistic and applicable to *TradeLens*. Also, the requirements resemble requirements applied internally in *TradeLens*, which shows the link between the research and current development trends in the private sector.

Second, the novelty of *DID*-based verifiable credentials has been confirmed by *expert 1*. It is acknowledged that they are an essential feature to ensure the scalability of secure and interoperable solutions in the shipping industry. The interaction between the cross-chain transfer protocol and the identity management of the architecture is identified by *expert 1* as a valid use case of *DID*s in the supply chain domain. Furthermore, it emphasises the relevance of the technical challenges addressed during the *Design & Development* phase, such as public key rotation to reduce collaboration friction in asynchronous service choreographies.

Lastly, *expert 1* confirms the contribution of the proposed architecture to ledger interoperability in a fragmented platform environment. From a functional viewpoint, the overlay network and *DAG* application model are considered useful for the purpose of the architecture. However, *expert 1* addressed the need to study in more detail the operational requirements for its implementations. These include the limitations imposed by implementation costs and considerations on the governance of the architecture covered in section 6.3.

Overall, the external validation has proved the suitability of the architecture. The integration of the proposed architecture components in *TradeLens* are considered by *expert 1* a significant engineering investment, but certainly possible. A definitive evaluation of the compatibility between the proposed architecture and the platform would require a more detailed technical analysis by the platform´s architects and developers.

6.3 Practical Limitations & Improvements

Additional aspects related to the implementation of the proposed architecture have been gathered during the expert consultation. This section focuses on the limitations of the design regarding two main topics. First, the role of governance, which is addressed by *expert 1*. Also, the foreseeable distribution effects of costs and benefits are discussed as a potential implementation barrier.

Governance

The research has focused on the design of a data sharing architecture to allow public entities to collect information produced in private domains. As part of a larger social coordination system, governance is a design factor closely linked to the alignment of interests between stakeholders [49]. The requirements included in the research do not account for governance requirements, which are defined by *van Engelenburg et. al* [49] as "*decision rights that parties should be able to exercise based on stakeholder dynamics and the design choices*" (pp. 199). Therefore, a link between these rights and the design decisions should be established. Shown in Figure 6.1 is a framework to analyse the relationship between governance requirements in *B2G* communication settings and the design of blockchain-based information systems [49].

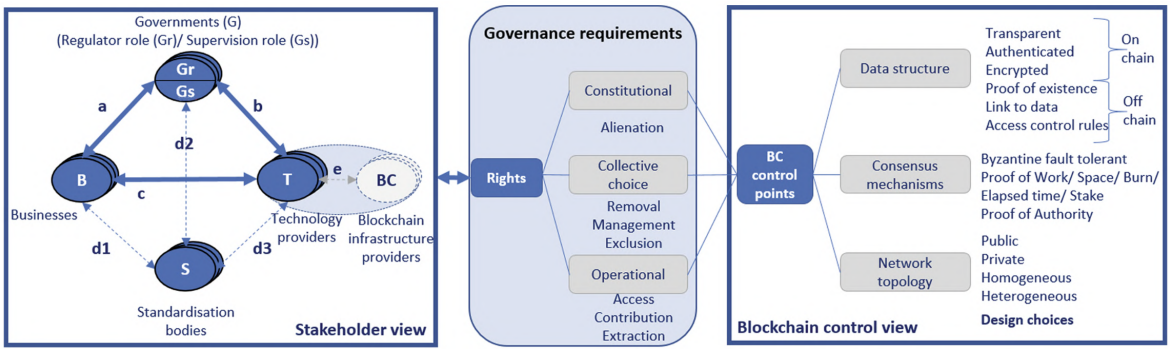


FIGURE 6.1: Blockchain governance framework for *B2G* communication [49].

The framework uses governance rights to link stakeholder dynamics with governance, and BCT control points. Certain governance aspects have been covered implicitly through the data sovereignty design principle. For instance, the assessment of European legislation studied how to translate current operational rights into a data governance model that promotes trust. However, the relevance of the research could benefit from a systematic review of governance implications using the aforementioned framework so that governance rights are captured throughout the entire design.

Two examples of operational scenarios vaguely addressed in the research and closely related to governance are the following. First, the collaboration of supply chain actors whose economic activities fall outside European jurisdiction with European customs. The latter are required to wait until a carrier has processed declaration information. If European customs

could directly access the data produced by cargo custodians before it is loaded in the vessel calling a European port, a better monitoring of incoming cargo could be developed. This way more frequent and earlier cross-validations could be performed [148]. A better understanding of the governance requirements is needed to assess the feasibility of persuading supply chain actors operating outside the *EU* to identify themselves to European customs.

The second example is an architecture feature, whose implementation is still surrounded by uncertainty. The feature in question is the definition of *DID*-delegates. This feature allows declarants to designate trusted agents to act on their behalf in the provision of declaration data. Given the international nature of the multi-actor system in which the architecture is to be deployed, a better study on the collective and constitutional rights to protect supply chain actors against the legal ramifications of this practice is required.

Implementation Costs

The research aimed to explore the required technical components to improve an existing *B2G* communication process. The architecture implementation costs have thus been left aside during the *Requirement Generation* and *Design & Development* phases. Operational costs are however a complex and fundamental aspect of information system design that can help identifying the effect of hidden or uninformed design assumptions [14]. Let us assume an interface based on the proposed architecture to allow supply chain actors to prepare their import declaration links for European customs. Let us also ignore the costs related to software development. Table 6.4 shows the costs associated with the implementation. These costs are divided in three main categories and include some of the common hidden costs in the deployment of information system identified by *Barreau* [14].

TABLE 6.4: Overview of implementation costs.

type	*category*	*description*
Preliminary	Testing and planning	Project management costs to plan and monitor the implementation progress (unit testing, interface validation, etc.).
	Conversion of legacy systems	Preparing the migration of data and redesign of processes registered in previous systems. Insufficient conversion can be translated to additional operational and maintenance costs [14] (*e.g.*, maintenance of multiple access tools).
	Documentation and training	Preparing documentation, manuals and protocols, as well as training activities for the users of the system.
Operational	Hardware	Vendor selection, purchase and installation of new equipment.
	Supporting infrastructure	Secondary infrastructure providing support to the architecture (*e.g.*, servers, equipment cooling or energy supply [14]).
	Overhead	Decision-making hierarchy to ensure the system remains operational (*e.g.*, overhead of management structure and user staff [14]).
Maintenance	Performance monitoring	Effort to extract and analyse information from the system to measure and predict performance.
	Updates and improvements	Development of software updates and recommendations for potential improvements.
	Support staff	Administrative and training costs of technical staff in support centers [14].

The distribution of these costs depends on the role and needs of each stakeholder, so understanding the effect of design decisions on these costs can improve the feasibility of the design. For example, the scalability of the system in terms of transaction throughout will not affect the training costs of new users, while it may affect the maintenance costs of the maintainer of the system. Therefore, an explicit link between components and costs throughout the development phase would lead to better design trade-offs.

In the same way hidden costs can inflate implementation budgets, insufficient analysis of the distribution of intangible benefits in a multi-actor system can prevent an accurate assessment of its practical desirability. This is particularly relevant when cost-benefit analysis is used to allocate public funding to initiatives of general interest.

In order to increase the accuracy of the expected distribution of benefits, it is useful to establish a link between the design components and the stakeholder dynamics of the application domain. The distribution effects of *DLT*-based information systems for supply chain applications is studied by *Roeack et. al* [138]. These relationships are shown in Table 6.5. A link between effects and the architecture design can be built by mapping effects into design principles, requirements and eventually design decisions.

TABLE 6.5: *DLT* characteristics and effects in supply chain transactions [138].

Name of effect		Type of effect	Relation to extant TCE	DLT-enabler		
				Transparency	Trust	Disintermediation
TCE assumptions	DLT-enabled assistance effect	Cost avoidance effect due to better decision-making by embanked bounded rationality	Confirming	D ++	–	–
	DLT-enhanced substitution effect	Cost reduction effect due to DLT-enabled trust as substituting assumption for opportunism	Expanding	I ++	D ++	–
Transaction dimension	DLT-enhanced disclosure effect	Cost reduction effect due to better performance evaluation of partners based on DLT data	Confirming	I ++	–	–
	DLT-caused scale-pan effect	Cost reduction (increase) effect due to equalised (reinforced) information asymmetry	Expanding	D ++	–	–
	DLT-enabled demonopolisation effect	Power shift due to diminished role of third party	Expanding	I +	–	D ++
	DLT-caused network effect	Dependency increasing due to network effect for gapless transparency	Refining	I +	–	–
Transaction costs and governance mode	DLT-enabled segregation effect	Cost reduction due to facilitated searching for transaction partners	Expanding	D ++	–	–
	DLT-enhanced automation effect	Cost reduction due to automated monitoring and enforcement based on verified data	Refining	D +	–	–
	DLT-caused torpedo effect	Power shift due to the potential to lose bargaining power	Refining	D 0	–	–

Type of relation to DLT-enabled cause
D Direct
I Indirect
– Not related

Strength of relation
++ Very strong influence (appearance in all cases, multiple codes)
+ Strong influence (appearance in all cases, one code)
0 Weak influence (appearance in some cases).

6.4 Evaluation Conclusion

This chapter was aimed at answering *RQ6*: *Does the data sharing architecture comply with the requirements to a sufficient extent to be considered a feasible solution that contributes to the application domain?* A naturalistic *ex ante* evaluation has been performed to assess the compliance with design requirements and the applicability of the data sharing architecture in the proposed practical setting.

The successful link between functional and non-functional requirements and the components of the architecture has been discussed. The relevance and suitability of the designed architecture has been tested against the industry knowledge of an expert, who has validated its implementation potential in a leading commercial blockchain platform. Also, governance and implementation costs have been identified as two main weaknesses of the design, and an approach to tackle both has been proposed. First, a framework to link governance and technical requirements has been discussed. Then, an overview of the expected implementation costs has been covered, as well as the importance of identifying potential transaction benefits of a *DLT*-based design approach for supply chain applications.

Chapter 7

Conclusion

This is the last chapter of the research. It covers a reflection on the work done and its contribution to the research objectives. The answer to the research questions is covered in section 7.1. The main research question is answered in section 7.2. The scientific contribution of the research is discussed in section 7.3 and the societal relevance in section 7.4. Finally, the recommendations for future research are covered in section 7.5.

7.1 Answering Research Questions

Problem Explication

This phase was intended to describe the nature of contractual agreement data sharing between supply chain actors, the European customs declaration process and the role of blockchain technology in increasing supply chain visibility. The following paragraphs provide answer to *RQ1*: *What is the relationship between supply chain visibility, import declarations and the risk assessments performed by customs administrations?*.

In the recent past, European customs have been willing to trade-off the consistency of physical cargo inspections for the collection of high quality declaration data. Shipping companies have been incentivised to share more information with a relaxation of customs requirements and privileges in customs facilities. The rise of commercial blockchain platforms and the digitisation of *B/Ls* are seen by European customs as an opportunity not to solely rely on economic incentives to increase their visibility over supply chains entering the *EU*. They envision an ecosystem able to provide a constant stream of high quality declaration data and thus increase supply chain visibility. However, European customs should not be satisfied with interacting individually with these data sharing service providers.

The participation of customs administrations in all blockchain platforms could decrease the administrative hurdle of customs declarations. However, siloed information flows would still be destined to form. As a result, a different approach to communicate between private and public entities is needed. The main driver is thus equipping blockchain platforms with capabilities to create verifiable links with other platforms, and allow customs administrations to extract declaration data from these links.

TABLE 7.1: Overview of design principles (*DP's*).

code	**name**	**description**
DP1	Logistic Event Visibility	Provide features that allow European customs to maintain the visibility of logistic events across the logistics domain
DP2	Stakeholder Data Sovereignty	Provide features that ensure the preservation of the data sovereignty rights of supply chain stakeholders
DP3	Architecture Interoperability	Provide features that support the compatibility of data exchanges between European customs and different blockchain platforms

Requirement Generation

This phase answers *RQ2 - What are the design requirements to preserve the data sovereignty of supply chain actors when creating links to data stored in multiple ledgers?* - and *RQ3 - What are the design requirements to allow multiple blockchain platforms to share interoperable links to their ledger states?* Table 7.1 shows the three design principles used to generate the requirements. The first principle, *logistic event visibility*, was added on top of the research questions to represent the essential practical value of the architecture for European customs administrations. The two remaining principles are the practical challenges addressed in the research questions: *stakeholder data sovereignty* and *architecture interoperability*. An overview of the functional and non-functional requirements is shown in Table 7.2 and Table 7.3.

TABLE 7.2: Overview of functional requirements.

code	*description*
	DP1: Logistic Event Visibility
FR1	The artefact should allow European customs to access information produced up-stream supply chains entering the European Union
FR2	The artefact should allow to monitor trade framework agreements and detect anomalies in the business transactions between supply chain actors
	DP2: Stakeholder Data Sovereignty
FR3	The artefact should allow the enforcement of the recognisability principles implemented in all procedures carried out by European customs administrations
FR4	The artefact should allow the certification of the user identities
	DP3: Architecture Interoperability
FR5	The artefact should allow European customs administrations to reuse information produced and/or stored by different blockchain protocols
FR6	The artefact should support the exchange of digital assets between platforms

TABLE 7.3: Overview of non-functional requirements.

code	*description*
	DP1: Logistic Event Visibility
NFR1	Data access should be provided by a pull mechanism with links between data sources
NFR2	Cargo flows should be monitored by processing data gathered by the artefact
NFR3	Transport flows should be monitored by processing data gathered by the artefact
	DP2: Stakeholder Data Sovereignty
NFR4	The enforcement of recognisability must be structurally restricted to European customs administrations by the architecture of the artefact
NFR5	The access to data must be exclusive to the data subjects, persons acting on their behalf and customs administrations
NFR6	The interpretation of data must be exclusive to the data subjects, persons acting on their behalf and customs administrations
NFR7	Amendments to data performed through the artefact, their motivation and authors must become traceable by customs administrations
NFR8	Exercising the right to erasure must be enabled to the data subjects of the data collected, processed and stored by the artefact
	DP3: Architecture Interoperability
NFR9	The digital assets processed by the artefact should be accessible and editable by a number of users related to the data subject
NFR10	The artefact must include features to exchange arbitrary events between platforms
NFR11	The cross-platform communication performed through the artefact must preserve consistency between the consensus protocols of different platforms
NFR12	Modifications to the internal specifications and functionalities of a commercial platform must not be required in order to interact with the artefact

Design & Development

This research phase addressed *RQ4*: *What architecture components can used by customs administrations to gather decla- ration data stored in multiple commercial blockchains?*. Based on the requirements generated in the previous research phase, a number of components have been included in the archtiecture. An overview of the chosen components and the functionality that motivated their selection can be found in Table 7.4.

TABLE 7.4: Summary of design choices and driving functions.

architecture component	*driving function*
Overlay Network	Cross-platform peer-to-peer communication
Trusted Gateway Protocol	Exchange of ledger states proofs to feed the application logic of cross-platform clients
Hash Time-lock Contracts	Propagation of self-sovereign smart contract logic
State View Board	Plug-compatible permissioned ledger state observation
Decentralised Identifiers	Self-sovereign identity management
Asynchronous Accumulators	Dynamic data access rules and asynchronous ledger state proof updates
Directed Acyclic Graph Ledger	Modular distributed application logic while maintaining parallel ledger views

Demonstration

The demonstration phase was intended to answer *RQ5*: *How would the current import declaration procedure be implemented using the peer- to-peer data sharing architecture?* By means of a use case, it is shown how declaration data can be created as a network of links towards document stored in blockchain platforms. The verification properties of these links are maintained with the supply chain actors that own the data. The configuration and presentation of these links allows customs to pull the required declaration data. Moreover, it is shown that the declaration procedure used by customs administrations can become more modular and less prone to data inconsistencies using the architecture to achieve *ENS*-friendly *B/L* issuing compared to the current declaration procedure.

Evaluation

This chapter was aimed at answering *RQ6*: *Does the data sharing architecture comply with the requirements to a sufficient extent to be considered a feasible solution that contributes to the application domain?* The compliance with design requirements and the contribution of the data sharing architecture the proposed practical setting have been tested.

The successful link between functional and non-functional requirements and the components of the architecture has been discussed. The relevance and suitability of the designed architecture has been tested against the industry knowledge of an expert, who has validated its implementation potential in a leading commercial blockchain platform. Also, governance and implementation costs have been identified as two main weaknesses of the design, and an approach to tackle both has been proposed. First, a framework to link governance and technical requirements has been discussed. Then, an overview of the expected implementation costs has been covered, as well as the importance of identifying potential transaction benefits of a *DLT*-based design approach for supply chain applications.

7.2 Answer to Main Research Question

The main research question - *What interoperable peer-to-peer data sharing architecture can be used by European customs administrations to gather declaration data from commercial blockchain platforms while preserving the data sovereignty of supply chain actors?* - can be answered by combining the findings of each research phase.

At a network level, overlay bridges are used to create peer-to-peer connections between platforms. These bridges are maintained through an authenticated gateway, with which platform participants interact to share information with external peers. Groups of peers take part in a cross-chain protocol with three phases: an preliminary phase two register and validate your identity, an initial phase in which a verifiable state required by an external peer is generated and published, and a resource exposure phase used when the actual migration of an asset between storage location is needed. By means of the verifiable credential model [154], peers generate verifiable presentations of the internal state of their respective ledgers. These states contain the transactions that validate a claim that needs to be verified by an external peer.

These claims are verified in order to complement the logic implemented in another platform. Due to the asynchrony of the system, arbitrary conditions on the validity of cross-chain transactions are implemented by means of hash time-lock contracts. These are used to encode necessary proofs of consensus or data availability that need to be delivered throughout the protocol phases. These verifiable presentations of the state of a permissioned ledger as observed by a trusted node is used to feed cross-chain applications built using *DAG* ledgers. Multiple participants can contribute to the a ledger, while each participant holds an individual view of the latter, which restricts the amount of transactions that can be validated and observed.

These restrictions are based on an access control based on self-certifying identifiers provided during the first phase of the cross-chain protocol. The access control to these applications is ruled by a credential database, in which verifiable proofs of the relationships between the publishers and subscribers of application updates is maintained. These relationships are resolution proofs of *DID* documents, with which entities can proof their link towards an entity without revealing details of his own entity or of the other party of the relationship.

7.3 Scientific Contribution

The main scientific contribution of the research has been proposing the combined used of *DAG* ledgers, asynchronous accumulators and self-certifying identifiers to create links between data stored in different blockchain platforms. In general, it can be interpreted as an approach to model data sharing patterns in multi-actor systems, taking into account commercial secrecy and auditability requirements. In this case, it has been applied to enhance trust between supply chain stakeholders while improving the efficiency of the generation of customs declaration data.

From a technical perspective, the research has shown that cryptographic accumulators can be used for the distributed coordination of data sharing processes, besides their more traditional security applications. It has also shown that directed acyclic graph ledgers can be used as a tool to build interoperable bridges between permissioned blockchains. This allows an entity to retrieve a verifiable view of a process state by aggregating the individual

views of a number of trusted participants and activate external process logic. This was been applied to customs administrations following traces of contractual supply chain data, although it could be applied to other trustless multi-actor systems.

In terms of identity management, the research has piggybacked on the work of [135] and [199] to explore the use of asynchronous accumulators for *DPKI*s. It is shown that this approach is particularly useful in environments where cryptographic proofs of data authorship are already generated. Also, it has been shown feasible to reduce the discoverability limitations of cross-chain communication protocols by using *DID*s to model the identifies of participants and to overcome key rotation and certificate revocation issues.

From a design science perspective, the research has delivered a new type of information object with the potential to solve the data gathering and distribution problems faced by customs administrations. However, the resulting architecture is also a different approach towards *DLT* design, which is a very relevant aspect given the scarcity of design and implementation frameworks of this nature.

7.4 Societal Relevance

Besides contributing to closing a gap in a growing research field, the use of the architecture provides benefits to society in different ways. The most direct one is increasing the amount of cargo that European customs can process. The result is less fraud and better port planning, which means that the *EU* is involved in safer and more efficient trade. This brings economic incentives to industries and populations of the member states. Within the logistics sector, the architecture represents the reduction of collaboration friction, the increase of operational efficiency and the reduction of costs. The resulting success of European enterprises can lead to creation of new employment and the improvement of the quality of the current employment within the sector.

From a strategic viewpoint, the implementation of the proposed architecture contributes to the widespread adoption of *DLT* to solve challenges in *B2G* information sharing. It is a step forward in the acceptation of distributed information sharing as a tool to improve the effectiveness of public institutions. Increasing the maturity of *DLT* solutions for public use in transport and commerce serves a use case transferable to other areas of public interest where *DLT*, *BCT* in particular, show great implementation potential, such as healthcare [4], energy [48], agriculture [145].

7.5 Future Research

The research has focused on enabling document-based information sharing. This is very relevant because contractual information is encoded within document flows to automate the acceptance and release of cargo at transfer points throughout supply chains. However, this approach leaves aside a growing number of additional data sources, such as sensor-based data collected at transport terminals. Some of the components included in the proposed design have been inspired by their use in these other applications, logistics *IoT* in particular. An example is the use of *DAG*s to coordinate the operational decisions taken by independent sensor networks [104, 185, 197]. Since the level of maturity is higher in such applications, future research should explore how to complement the architecture with these data sources in order to integrate document-based data sharing with logistic data gathered on site.

The two document flows most represented in the research are *ENS* declarations and the issue process of different types of *B/L*. Although *ENS* declarations are based on the content of *B/Ls*, they might not share format. The research has assumed that the format and content differences between the two are negligible to focus on the high-level design of architecture components. An example of these differences are the detailed product definitions of a *HB/L* and the *HS* codes included in *ENS* data and used by authorities to assign tariff rules to product categories [54]. This leaves room for further research on the requirements for the cross-reference of documents between platforms to ensure end-to-end semantic compatibility during cross-validations performed by customs.

From a technical perspective, the long-term performance consequences of the selected components have not been contemplated by the research. For instance, the conceptual suitability of a consensus service based on asynchronous byzantine fault tolerance has been stated. Future research should evaluate the technical feasibility of an additional consensus layer an its requirements in terms of speed and scalability. This is important in order to confirm that the transaction throughput required by customs is in line with the architecture´s ability to maintain consensus on a growing number of cross-platform references.

In terms of the proposed application context, the declaration process has been simplified during the demonstration. The architecture has been presented as a one-way system, in which only supply chain actors provide verifiable presentations of data links to European customs administrations. In practice, a two-way system is required, as customs administrations also provide messages to supply chain actors. These can include *"do no load"* warnings or decisions about applications submitted by economic operators, such as temporary storage or updates on the progress of a customs investigation [54, 55]. This opens an opportunity for supply chain actors to integrate cargo acceptance messages issued by customs administrations as collaterals in cross-chain smart contracts, which could help entities externalise certain risks of import and export agreements.

Bibliography

[1] NaQi . et al. "Analysis and Research of the RSA Algorithm". In: *Information Technology Journal* 12.9 (Aug. 2013). ISSN: 18125638. DOI: 10.3923/itj.2013.1818.1824.

[2] Ermyas Abebe et al. "Enabling Enterprise Blockchain Interoperability with Trusted Data Transfer (Industry Track)". In: *Proceedings of the 20th International Middleware Conference Industrial Track*. ACM, Aug. 2019. ISBN: 9781450370417. DOI: 10.1145/3366626.3368129.

[3] Ermyas Abebe et al. "Verifiable Observation of Permissioned Ledgers". In: *arXiv* (Aug. 2021). URL: https://arxiv.org/pdf/2012.07339.pdf.

[4] Israa Abu-elezz et al. "The benefits and threats of blockchain technology in healthcare: A scoping review". In: *International Journal of Medical Informatics* 142 (2020), p. 104246. ISSN: 1386-5056. DOI: https://doi.org/10.1016/j.ijmedinf.2020.104246. URL: https://www.sciencedirect.com/science/article/pii/S1386505620301544.

[5] Kevin MacG. Adams. *Nonfunctional Requirements in Systems Analysis and Design*. 1st ed. Vol. 28. Springer International Publishing, 2015. ISBN: 978-3-319-18343-5. DOI: 10.1007/978-3-319-18344-2.

[6] Ohoud Albogami et al. "Public Key Infrastructure Traditional andModern Implementation". In: *International Journal of Network Security* 23.2 (Aug. 2021), pp. 343–350.

[7] Steven Alter. "Defining information systems as work systems: implications for the IS field". In: *European Journal of Information Systems* 17.5 (Oct. 2008), pp. 448–469. ISSN: 0960-085X. DOI: 10.1057/ejis.2008.37.

[8] Mohammad Hassan Ameri et al. "A Key-Policy Attribute-Based Temporary Keyword Search scheme for Secure Cloud Storage". In: *IEEE Transactions on Cloud Computing* 8.3 (July 2020). ISSN: 2168-7161. DOI: 10.1109/TCC.2018.2825983.

[9] Mohammad Javad Amiri, Divyakant Agrawal, and Amr El Abbadi. "CAPER: a cross-application permissioned blockchain". In: *Proceedings of the VLDB Endowment* 12.11 (Aug. 2019). ISSN: 2150-8097. DOI: 10.14778/3342263.3342275.

[10] Anisha Mirchandani. "The GDPR-Blockchain Paradox: Exempting Permissioned Blockchains from the GDPR". In: *Fordham Intellectual Property, Media and Entertainment Law Journal* 29.4 (2019). URL: https://ir.lawnet.fordham.edu/iplj/vol29/iss4/5/.

[11] Leemon Baird. *THE SWIRLDS HASHGRAPH CONSENSUS ALGORITHM: FAIR, FAST, BYZANTINE FAULT TOLERANCE*. Aug. 2016. URL: https://www.swirlds.com/downloads/SWIRLDS-TR-2016-01.pdf.

[12] Mark C Ballandies, Marcus M Dapp, and Evangelos Pournaras. "Decrypting distributed ledger design—taxonomy, classification and blockchain community evaluation". In: *Cluster Computing* (Aug. 2021). ISSN: 1386-7857. DOI: 10.1007/s10586-021-03256-w.

[13] James G. Barnes and James E. Byrne. "E-Commerce and Letter of Credit Law and Practice". In: *The International Lawyer* 35.1 (2001), pp. 23–29. URL: TheInternationalLawyer.

[14] Deborah Barreau. "The hidden costs of implementing and maintaining information systems". In: *The Bottom Line* 14.4 (Dec. 2001). ISSN: 0888-045X. DOI: 10.1108/08880450110408481.

[15] Rafael Belchior et al. "HERMES: Fault-Tolerant Middleware for Blockchain Interoperability". In: *TechRxiv. Preprint* (2021). URL: https://www.techrxiv.org/ndownloader/files/26628587/1.

[16] Josh Benaloh and Michael de Mare. "One-Way Accumulators: A Decentralized Alternative to Digital Signatures". In: *Advances in Cryptology — EUROCRYPT '93. Lecture Notes in Computer Science*. Ed. by Helleseth T. Vol. 765. Berlin: Springer, July 2001. Chap. 24. ISBN: 978-3-540-48285-7. DOI: 10.1007/3-540-48285-7.

[17] Roberto Bergami. "The Link Between Incoterms 2000 and Letter of Credit Documentation Requirement and Payment Risk". In: *Journal of Business Systems, Governance and Ethics* 1.4 (Aug. 2006). ISSN: 1833-4318. DOI: 10.15209/jbsge.v1i4.91.

[18] Liton Chandra Biswas. "A Discussion and Analysis of the Bill of Lading as a Document of Title". In: *SSRN Electronic Journal* (Aug. 2011). ISSN: 1556-5068. DOI: 10.2139/ssrn.2089523.

[19] Sebastian Boell and Dubravka Cecez-Kecmanovic. "What is an Information System?" In: vol. 2015. Aug. 2015. DOI: 10.1109/HICSS.2015.587.

[20] Dan Boneh, Benedikt Bünz, and Ben Fisch. "Batching Techniques for Accumulators with Applications to IOPs and Stateless Blockchains". In: *Lecture Notes in Computer Science*. Ed. by Boldyreva A. and Micciancio D. Vol. 11692. Cham: Springer, Aug. 2019. Chap. 20, pp. 561–586. DOI: 10.1007/978-3-030-26948-7.

[21] Richard Braun et al. *Proposal for Requirements Driven Design Science Research*. 2015. DOI: 10.1007/978-3-319-18714-3.

[22] Jan vom Brocke et al. "Tool-Support for Design Science Research: Design Principles and Instantiation". In: *SSRN Electronic Journal* (2017). ISSN: 1556-5068. DOI: 10.2139/ssrn.2972803.

[23] Benedikt Bunz et al. "Bulletproofs: Short Proofs for Confidential Transactions and More". In: *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, May 2018, pp. 315–334. ISBN: 978-1-5386-4353-2. DOI: 10.1109/SP.2018.00020.

[24] David Burdett and Nickolas Kavantzas. *Web Services Choreography Model Overview - W3C Working Draft*. Aug. 2004. URL: https://www.w3.org/TR/ws-chor-model/.

[25] C Cachin et al. "The Transaction Graph for Modeling Blockchain Semantics". In: *IACR Cryptol. ePrint Arch.* 2017 (2017), p. 1070.

[26] Jan Camenisch and Anna Lysyanskaya. "Dynamic Accumulators and Application to Efficient Revocation of Anonymous Credentials". In: *Advances in Cryptology — CRYPTO 2002. Lecture Notes in Computer Science*. Vol. 2442. Berling: Springer, Sept. 2002. DOI: 10.1007/3-540-45708-9.

[27] Bin Cao et al. "When Internet of Things Meets Blockchain:Challenges in Distributed Consensus". In: *arXiv* (Aug. 2019). URL: https://arxiv.org/pdf/1905.06022.pdf.

[28] Maria Caridi et al. "The benefits of supply chain visibility: A value assessment model". In: *International Journal of Production Economics* 151.C (2014), pp. 1–19. URL: https://EconPapers.repec.org/RePEc:eee:proeco:v:151:y:2014:i:c:p:1-19.

[29] J.Lawrence Carter and Mark N. Wegman. "Universal classes of hash functions". In: *Journal of Computer and System Sciences* 18.2 (Apr. 1979). ISSN: 00220000. DOI: 10.1016/0022-0000(79)90044-8.

[30] Central Commission for the Navigation of the Rhine Danube Commission. *Budapest Convention on the Contract for the Carriage of Goods by Inland Waterways (CMNI)*. Aug. 2000. URL: https://unece.org/fileadmin/DAM/trans/main/sc3/cmniconf/cmnidoc/finalconf02e.pdf.

[31] David W Chadwick et al. "Improved Identity Management with Verifiable Credentials and FIDO". In: *IEEE Communications Standards Magazine* 3.4 (Aug. 2019). ISSN: 2471-2825. DOI: 10.1109/MCOMSTD.001.1900020.

[32] Y Chang, E Iakovou, and W Shi. "Blockchain in global supply chains and cross border trade: a critical synthesis of the state-of-the-art, challenges and opportunities". In: *International Journal of Production Research* 58.7 (2020), pp. 2082–2099.

[33] Wei Chen et al. "Developing a Concurrent Service Orchestration Engine Based on Event-Driven Architecture". In: 2008. DOI: https://doi.org/10.1007/978-3-540-88871-0{_}47.

[34] Yi-Cheng Chen, Yueh-Peng Chou, and Yung-Chen Chou. "An Image Authentication Scheme Using Merkle Tree Mechanisms". In: *Future Internet* 11.7 (July 2019). ISSN: 1999-5903. DOI: 10.3390/fi11070149.

[35] Sang-Min Choi et al. *Fantom: A scalable framework for asynchronous distributed systems*. 2018. URL: https://arxiv.org/abs/1810.10360v1.

[36] T . Choi et al. "The mean-variance approach for global supply chain risk analysis with air logistics in the blockchain technology era". In: *Transportation Research Part E: Logistics and Transportation Review* 127 (2019), pp. 178–191.

[37] YeonSung Chu et al. "SS-DPKI: Self-Signed Certificate Based Decentralized Public Key Infrastructure for Secure Communication". In: *2020 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, Aug. 2020. ISBN: 978-1-7281-5186-1. DOI: 10.1109/ICCE46568.2020.9043086.

[38] Clyde & Co LPP. *The Legal Status of Electronic Bills of Lading: A Report for the ICC Banking Commission*. 2018. URL: https://iccwbo.org/content/uploads/sites/3/2018/10/the-legal-status-of-e-bills-of-lading-oct2018.pdf.

[39] Council of the European Union. "Council Decision 2009/917/JHA of 30 November 2009 on the use of information technology for customs purposes". In: *Official Journal of the European Union* 323 (Aug. 2009), pp. 20–30. URL: http://data.europa.eu/eli/dec/2009/917/oj.

[40] Stefan Cronholm and Hannes Göbel. "Evaluation of the Information Systems Research Framework: Empirical Evidence from a Design Science Research Project". In: *Electronic Journal Information Systems Evaluation* 19 (Aug. 2016), pp. 157–167.

[41] Mohammad Dabbagh, Sai Peck Lee, and Reza Meimandi Parizi. "Functional and non-functional requirements prioritization: empirical evaluation of IPA, AHP-based, and HAM-based approaches". In: *Soft Computing* 20.11 (Aug. 2016). ISSN: 1432-7643. DOI: 10.1007/s00500-015-1760-z.

[42] Bingrong Dai et al. "Research and Implementation of Cross-Chain Transaction Model Based on Improved Hash-Locking". In: *Blockchain and Trustworthy Systems*. Ed. by Zheng Z. et al. Singapore: Springer, Nov. 2020. Chap. 17. ISBN: 978-981-15-9213-3. DOI: 10.1007/978-981-15-9213-3{_}17.

[43] Simon Dalmolen et al. "Trust in a multi-tenant, logistics, data sharing infrastructure: Opportunities for blockchain technology." In: *5th International Physical Internet Conference*. 2018.

[44] Farhad Daneshgar, Omid Sianaki, and Prabhat Guruwacharya. *Blockchain: A Research Framework for Data Security and Privacy*. Aug. 2019. DOI: 10.1007/978-3-030-15035-8{_}95.

[45] Liping Deng et al. "Research on Cross-Chain Technology Based on Sidechain and Hash-Locking". In: *Edge Computing - EDGE 2018*. Ed. by Liu S. et al. Cham: Springer, June 2018. Chap. 12. ISBN: 978-3-319-94340-4. DOI: 10.1007/978-3-319-94340-4{_}12.

[46] David Derler, Christian Hanser, and Daniel Slamanig. "Revisiting Cryptographic Accumulators, Additional Properties and Relations to Other Primitives". In: *Topics in Cryptology – CT-RSA 2015. Lecture Notes in Computer Science*. Ed. by Nyberg K. Vol. 9048. Cham: Springer, Mar. 2015. Chap. 7. ISBN: 978-3-319-16715-2. DOI: 10.1007/978-3-319-16715-2{_}7.

[47] Daniel Drescher. *Blockchain Basics*. Apress, 2017. ISBN: 978-1-4842-2603-2. DOI: 10.1007/978-1-4842-2604-9.

[48] Pankaj Dutta et al. "Blockchain technology in supply chain operations: Applications, challenges and research opportunities". In: *Transportation Research Part E: Logistics and Transportation Review* 142 (2020), p. 102067. ISSN: 1366-5545. DOI: https://doi.org/10.1016/j.tre.2020.102067. URL: https://www.sciencedirect.com/science/article/pii/S1366554520307183.

[49] Sélinde van Engelenburg et al. "Aligning Stakeholder Interests, Governance Requirements and Blockchain Design in Business and Government Information Sharing". In: *Electronic Government - 19th IFIP WG 8.5 International Conference, EGOV 2020, Proceedings*. Springer, 2020, pp. 197–209. DOI: 10.1007/978-3-030-57599-1{_}15.

[50] European Commission. "Proposal for a Regulation of the Parliament and of the Council on European data governance (Data Governance Act)". In: (Aug. 2020).

[51] European Cyber Security Organisation (ECSO). *European Cyber Security Certification A Meta-Scheme Approach v1.0: WG1– Standardisation, certification, labelling and supply chain management*. Aug. 2017. URL: http://dev.ecsorg.eu/documents/uploads/european-cyber-security-certification-a-meta-scheme-approach.pdf.

[52] European Parliament. "Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?" In: *European Parliamentary Research Service* PE 634.445 (Aug. 2019).

[53] European Parliament and Council of the European Union. "Regulation (EU) No 952/2013 of the European Parliament and of the Council of 9 October 2013 laying down the Union Customs Code". In: *Official Journal of the European Union* 269 (Aug. 2013), pp. 1–101. URL: http://data.europa.eu/eli/reg/2013/952/oj.

[54] European Parliament and Council of the European Union. "Commission Delegated Regulation (EU) 2015/2446 of 28 July 2015 supplementing Regulation (EU) No 952/2013 of the European Parliament and of the Council as regards detailed rules concerning certain provisions of the Union Customs Code". In: *Official Journal of the European Union* 343 (Aug. 2015), pp. 1–557. URL: http://data.europa.eu/eli/reg_del/2015/2446/oj.

[55] European Parliament and Council of the European Union. "Commission Implementing Regulation (EU) 2015/2447 of 24 November 2015 laying down detailed rules for implementing certain provisions of Regulation (EU) No 952/2013 of the European Parliament and of the Council laying down the Union Customs Code". In: *Official Journal of the European Union* 343 (Aug. 2015), pp. 558–893. URL: http://data.europa.eu/eli/reg_impl/2015/2447/oj.

[56] European Parliament and Council of the European Union. "Commission Delegated Regulation (EU) 2016/341 of 17 December 2015 supplementing Regulation (EU) No 952/2013 of the European Parliament and of the Council as regards transitional rules for certain provisions of the Union Customs Code where the relevant electronic systems are not yet operational and amending Delegated Regulation (EU) 2015/2446". In: *Official Journal of the European Union* 69 (Aug. 2016), pp. 1–313. URL: http://data.europa.eu/eli/reg_del/2016/341/oj.

[57] European Parliament and Council of the European Union. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)". In: *Official Journal of the European Union* 119 (Aug. 2016), pp. 1–88. URL: http://data.europa.eu/eli/reg/2016/679/oj.

[58] European Parliament and Council of the European Union. "Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) (Text with EEA relevance)". In: *Official Journal of the European Union* 151.PE/86/2018/REV/1 (Aug. 2019), pp. 15–69. URL: http://data.europa.eu/eli/reg/2019/881/oj.

[59] European Parliamentary Research Service: Panel for the Future of Science and Technology. *Blockchain for Supply Chains and International Trade: Report on Key Features, Impacts and Policy Options*. Aug. 2020. URL: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641544/EPRS_STU(2020)641544_EN.pdf.

[60] FEDeRated: Network of Platforms. "SEMANTICS - Discussion paper: Detailing the FEDeRATED Reference Framework". In: *European Commission Connecting Europe Facility* (2020). URL: http://www.federatedplatforms.eu/index.php/library/item/detailing-the-reference-model-semantics-discussion-paper.

[61] FEDeRated: Network of Platforms. "Supply Chain Visibility (SCVL) A Ledger-based Demonstrator". In: *European Commission Connecting Europe Facility* (Aug. 2020). URL: http://www.federatedplatforms.eu/index.php/library/item/semantics-supply-chain-visibility-ledger-discussion-paper.

[62] Philipp Frauenthaler et al. *Leveraging Blockchain Relays for Cross-Chain Token Transfers*. Aug. 2020.

[63] O Gass and A Maedche. "Enabling End-user-driven Data Interoperability - A Design Science Research Project". In: *AMCIS*. 2011.

[64] Jorge Gracia and Eduardo Mena. "Dealing with Semantic Heterogeneity Issues on the Web". In: *IEEE Internet Computing* (2019). ISSN: 1089-7801. DOI: 10.1109/MIC.2011.129.

[65] A Gurtu and J Johny. "Potential of blockchain technology in supply chain management: a literature review". In: *International Journal of Physical Distribution and Logistics Management* 49.9 (2019), pp. 881–900.

[66] D Gurzick and W G Lutters. "Towards a design theory for online communities". In: *4th International Conference on Design Science Research in Information Systems and Technology (DESRIST'09)*. Association for Computing Machinery, New York, Aug. 2009, pp. –.

[67] Arni Halldorsson and Jan Stentoft. *Research Methodologies in Supply Chain Management — What Do We Know?* Aug. 2005. DOI: 10.1007/3-7908-1636-1.

[68] Harry Halpin. "Deconstructing the Decentralization Trilemma". In: *Proceedings of the 17th International Joint Conference on e-Business and Telecommunications* (2020). DOI: 10.5220/0009892405050512. URL: http://dx.doi.org/10.5220/0009892405050512.

[69] Ki-Moon Han, Sae-Woon Park, and Sunhae Lee. "Anti-Fraud in International Supply Chain Finance: Focusing on Moneual Case". In: *Journal of Korea Trade* 24.1 (Aug. 2020). DOI: 10.35611/jkt.2020.24.1.59.

[70] Thomas Hardjono. "Blockchain Gateways, Bridges and Delegated Hash-Locks". In: *arXiv* (Aug. 2021). URL: https://arxiv.org/pdf/2102.03933.pdf.

[71] Karina Hauser, Helgi S. Sigurdsson, and Katherine M. Chudoba. "EDSOA: An Event-Driven Service-Oriented Architecture Model For Enterprise Applications". In: *International Journal of Management & Information Systems (IJMIS)* 14.3 (Jan. 2011). ISSN: 2157-9628. DOI: 10.19030/ijmis.v14i3.839.

[72] P Helo and Y Hao. "Blockchains in operations and supply chains: A model and reference implementation". In: *Computers and Industrial Engineering* 136 (2019), pp. 242–251.

[73] A Henderson and J Burnie. "Putting names to things: reconciling cryptocurrency heterogeneity and regulatory continuity". In: *Journal of International Banking and Financial Law* 33.4 (2018).

[74] Maurice Herlihy. "Atomic Cross-Chain Swaps". In: *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing*. New York, NY, USA: ACM, July 2018. ISBN: 9781450357951. DOI: 10.1145/3212734.3212736.

[75] David Hesketh. "Weaknesses in the supply chain: who packed the box?" In: *World Customs Journal* 4.2 (Aug. 2010), pp. 3–20.

[76] Bernard Hoekman and Laura Puccio. "EU Trade Policy: Challenges and Opportunities". In: *European University Institute: Robert Schuman Centre for Advanced Studies* (2019). URL: http://respect.eui.eu/wp-content/uploads/sites/6/2019/02/RSCAS_PP_2019_06-1.pdf.

[77] Wout Hofman. "Supply Chain Risk Analysis with Linked Open Data". In: *Formal Ontologies Meet Industry (FOMI) Proceedings of the Fifth International Workshop*. Aug. 2011.

[78] Wout J Hofman. "A Methodological Approach for Development and Deployment of Data Sharing in Complex Organizational Supply and Logistics Networks with Blockchain Technology". In: *IFAC-PapersOnLine* 52.3 (2019), pp. 55–60. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2019.06.010. URL: https://www.sciencedirect.com/science/article/pii/S2405896319300941.

[79] Alen Hrga, Tomislav Capuder, and Ivana Podnar Zarko. "Demystifying Distributed Ledger Technologies: Limits, Challenges, and Potentials in the Energy Sector". In: *IEEE Access* 8 (2020). ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3007935.

[80] Huawei Huang et al. "A Survey of State-of-the-Art on Blockchains: Theories, Modelings, and Tools". In: *arXiv* (Aug. 2020). URL: https://arxiv.org/pdf/2007.03520.pdf.

[81] J Hulstijn et al. "Towards Trusted Tradelanes". In: *Proceedings of the 15th IFIP E-Government conference (EGOV 2016): Electronic Government*. Ed. by H J Scholl et al. Lecture Notes in Computer Science, Springer International Publishing, 2016, pp. 299–311.

[82] Hans-Henrik Hvolby et al. "Information Exchange and Block Chains in Short Sea Maritime Supply Chains". In: *Procedia Computer Science* 181 (2021), pp. 722–729. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2021.01.224. URL: https://www.sciencedirect.com/science/article/pii/S1877050921002672.

[83] Hyperledger. "Peer Channel-based Event Services". In: *Hyperledger Fabric: A Blockchain Platform for the Enterprise*. 2020. URL: https://hyperledger-fabric.readthedocs.io/en/release-2.3/peer_event_services.html.

[84] Juhani Iivari and John Venable. "Action research and design science research - Seemingly similar but decisively dissimilar". In: *17th European Conference on Information Systems, ECIS 2009*. Aug. 2009, pp. 1642–1653.

[85] ISO/IEC Joint Technical Committee 1 - Subcommittee 7 - Software and Systems Engineering. *ISO/IEC/IEEE 42010:2011 - Systems and software engineering — Architecture description*. 1st ed. International Organization for Standardization, 2011. URL: https://www.iso.org/standard/50508.html.

[86] Yassine Issaoui et al. "Smart logistics: Study of the application of blockchain technology". In: *Procedia Computer Science* 160 (2019). ISSN: 18770509. DOI: 10.1016/j.procs.2019.09.467.

[87] J. Hintsa et al. "Does better visibility help mitigate security risks in cross-border supply chains? - Case FP7-CASSANDRA". In: *e-Freight Conference 2012*. Aug. 2012.

[88] Sohail Jabbar et al. "Blockchain-enabled supply chain: analysis, challenges, and future directions". In: *Multimedia Systems* (Aug. 2020). ISSN: 0942-4962. DOI: 10.1007/s00530-020-00687-0.

[89] Hai Jin, Xiaohai Dai, and Jiang Xiao. "Towards a Novel Architecture for Enabling Interoperability amongst Multiple Blockchains". In: *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, Aug. 2018. ISBN: 978-1-5386-6871-9. DOI: 10.1109/ICDCS.2018.00120.

[90] Paul Johannesson and Erik Perjons. *An Introduction to Design Science*. 2014. DOI: https://doi.org/10.1007/978-3-319-10632-8.

[91] Audun Jøsang et al. "Local user-centric identity management". In: *Journal of Trust Management* 2.1 (Aug. 2015). ISSN: 2196-064X. DOI: 10.1186/s40493-014-0009-6.

[92] Ivan Jovović et al. "5G, Blockchain and IPFS: A General Survey wić with Possible Innovative Applications in Industry 4.0". In: *Proceedings of the 3rd EAI International Conference on Management of Manufacturing Systems*. EAI, 2018. ISBN: 978-1-63190-167-6. DOI: 10.4108/eai.6-11-2018.2279695.

[93] Niclas Kannengießer et al. "Trade-offs between Distributed Ledger Technology Characteristics". In: *ACM Computing Surveys* 53.2 (Aug. 2020). ISSN: 0360-0300. DOI: 10.1145/3379463.

[94] Chris Khan et al. "A Distributed-Ledger Consortium Model for Collaborative Innovation". In: *Computer* 50.9 (2017). ISSN: 0018-9162. DOI: 10.1109/MC.2017.3571057.

[95] Shafaq Naheed Khan et al. "Blockchain smart contracts: Applications, challenges, and future trends". In: *Peer-to-Peer Networking and Applications* (Aug. 2021). ISSN: 1936-6442. DOI: 10.1007/s12083-021-01127-0.

[96] Bram Klievink et al. "Enhancing Visibility in International Supply Chains". In: *International Journal of Electronic Government Research* 8.4 (Aug. 2012). ISSN: 1548-3886. DOI: 10.4018/jegr.2012100102.

[97] Ann Hibner Koblitz, Neal Koblitz, and Alfred Menezes. "Elliptic curve cryptography: The serpentine course of a paradigm shift". In: *Journal of Number Theory* 131.5 (Aug. 2011). ISSN: 0022314X. DOI: 10.1016/j.jnt.2009.01.006.

[98] T Koens and E Poll. "Assessing interoperability solutions for distributed ledgers". In: *Pervasive and Mobile Computing* 59 (2019), p. 101079. ISSN: 1574-1192. DOI: https://doi.org/10.1016/j.pmcj.2019.101079. URL: https://www.sciencedirect.com/science/article/pii/S1574119218306266.

[99] Norbert Koppenhagen, Oliver Gaß, and Benjamin Müller. *Design Science Research in Action - Anatomy of Success Critical Activities for Rigor and Relevance*. 2012.

[100] Czesław Kościelny, Mirosław Kurkowski, and Marian Srebrny. *Foundations of Asymmetric Cryptography*. 2013. DOI: 10.1007/978-3-642-41386-5{_}4.

[101] Nir Kshetri. "1 Blockchain's roles in meeting key supply chain management objectives". In: *International Journal of Information Management* 39 (2018), pp. 80–89.

[102] Pascal Lafourcade and Marius Lombard-Platet. "About blockchain interoperability". In: *Information Processing Letters* 161 (Aug. 2020). ISSN: 00200190. DOI: 10.1016/j.ipl.2020.105976.

[103] Youngsu Lee and Suk-Chul Rim. "Quantitative Model for Supply Chain Visibility: Process Capability Perspective". In: *Mathematical Problems in Engineering* 2016 (Aug. 2016), pp. 1–11. DOI: 10.1155/2016/4049174.

[104] Yixin Li et al. "Direct Acyclic Graph based Ledger for Internet of Things: Performance and Security Analysis". In: *arXiv* (Aug. 2020). URL: https://arxiv.org/pdf/1905.10925.pdf.

[105] Xidong Liu. "Research and Application of Electronic Invoice Based on Blockchain". In: *MATEC Web of Conferences* 232 (Aug. 2018). ISSN: 2261-236X. DOI: 10.1051/matecconf/201823204012.

[106] Yue Liu et al. "Design Pattern as a Service for Blockchain-Based Self-Sovereign Identity". In: *IEEE Software* 37.5 (Sept. 2020). ISSN: 0740-7459. DOI: 10.1109/MS.2020.2992783.

[107] Jacob Lohmer and Rainer Lasch. "Blockchain in operations management and manufacturing: Potential and barriers". In: *Computers & Industrial Engineering* 149 (2020), p. 106789. ISSN: 0360-8352. DOI: https://doi.org/10.1016/j.cie.2020.106789. URL: https://www.sciencedirect.com/science/article/pii/S0360835220304988.

[108] Mahabir Prasad Jhanwar and Pratyush Ranjan Tiwari. " Trading Accumulation Size for Witness Size: A Merkle Tree Based Universal Accumulator Via Subset Differences". In: *Cryptology ePrint Archive* (Oct. 2019). URL: https://eprint.iacr.org/2019/1186.

[109] T Männistö and J Hintsa. "PROFILE: Enhancing Customs Risk Management". In: *World Customs Organization News* Issue 89 (2019).

[110] Zvonko Merkaš, Davor Perkov, and Vesna Bonin. "The Significance of Blockchain Technology in Digital Transformation of Logistics and Transportation". In: *International Journal of E-Services and Mobile Applications* 12.1 (Aug. 2020). ISSN: 1941-627X. DOI: 10.4018/IJESMA.2020010101.

[111] Mike Hearn and Richard Gendal Brown. *Corda: A distributed ledger*. Aug. 2019. URL: https://www.r3.com/wp-content/uploads/2019/08/corda-technical-whitepaper-August-29-2019.pdf.

[112] Arshad Muhammad et al. "A Secure Gateway Service for Accessing Networked Appliances". In: *2010 Fifth International Conference on Systems and Networks Communications*. IEEE, Aug. 2010. ISBN: 978-1-4244-7789-0. DOI: 10.1109/ICSNC.2010.35.

[113] R S NEERAJ. "Trade Rules for the Digital Economy: Charting New Waters at the WTO". In: *World Trade Review* 18.S1 (Aug. 2019). ISSN: 1474-7456. DOI: 10.1017/S1474745618000423.

[114] Q K Nguyen. "Blockchain - A Financial Technology for Future Sustainable Development". In: *2016 3rd International Conference on Green Technology and Sustainable Development (GTSD)*. 2016, pp. 51–54. DOI: 10.1109/GTSD.2016.22.

[115] Nick Szabo. "Smart contracts: Building blocks for digital markets". In: *Extropy: Journal of Transhumanist Thought* 16 (1996), pp. 53–60.

[116] Andriy Nikolov et al. *Overcoming Schema Heterogeneity between Linked Semantic Repositories to Improve Coreference Resolution*. Ed. by Gómez-Pérez A., Yu Y., and Ding Y. 2009. DOI: 10.1007/978-3-642-10871-6{_}23.

[117] Theo Notteboom, Athanasios Pallis, and Jean-Paul Rodrigue. *Port Economics, Management and Policy*. URL: https://porteconomicsmanagement.org/.

[118] Svein Ølnes, Jolien Ubacht, and Marijn Janssen. "Blockchain in government: Benefits and implications of distributed ledger technology for information sharing". In: *Government Information Quarterly* 34.3 (Sept. 2017). ISSN: 0740624X. DOI: 10.1016/j.giq.2017.09.007.

[119] P. Zappalà et al. "Game theoretical framework for analyzing Blockchains Robustness". In: *Cryptology ePrint Archive* (May 2020). URL: https://eprint.iacr.org/2020/626.pdf.

[120] Yan Pang. "A New Consensus Protocol for Blockchain Interoperability Architecture". In: *IEEE Access* 8 (2020). ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3017549.

[121] A Panos, G Kapnissis, and H C Leligou. "The Blockchain and DLTs in the Maritime Industry: Potential and Barriers". In: *European Journal of Electrical Engineering and Computer Science* 4.5 (Aug. 2020). ISSN: 2506-9853. DOI: 10.24018/ejece.2020.4.5.243.

[122] G Perboli, S Musso, and M Rosano. "Blockchain in Logistics and Supply Chain: A Lean Approach for Designing Real-World Use Cases". In: *IEEE Access* 6 (2018), pp. 62018–62028.

[123] Babu Pillai, Kamanashis Biswas, and Vallipuram Muthukkumarasamy. "Cross-chain interoperability among blockchain-based systems using transactions". In: *The Knowledge Engineering Review* 35 (Aug. 2020). ISSN: 0269-8889. DOI: 10.1017/S0269888920000314.

[124] Evi Plomaritou and Yiannis Voudouris. "The Relationships of Bill of Lading, Charterparty and Other Transport Documents". In: *Journal of Economics, Management and Trade* (Aug. 2019). ISSN: 2456-9216. DOI: 10.9734/jemt/2019/v24i630182.

[125] Serguei Popov. *The Tangle*. Aug. 2018. URL: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf.

[126] Sudeep Kumar Pradhan and Srikanta Routroy. "Improving supply chain performance by Supplier Development program through enhanced visibility". In: *Materials Today: Proceedings* 5.2, Part 1 (2018), pp. 3629–3638. ISSN: 2214-7853. DOI: https://doi.org/10.1016/j.matpr.2017.11.613. URL: https://www.sciencedirect.com/science/article/pii/S2214785317328869.

[127] Alex Preukschat and Drummond Reed. *Self-Sovereign Identity: Decentralized digital identity and verifiable credentials*. 1st ed. Manning, Aug. 2021. ISBN: 9781617296598.

[128] PROFILE. *Deliverable 2.4 — Possibilities of Blockchain Technologies for Trusted Data Sharing (Internal Report)*. Tech. rep. 2021. URL: https://www.profile-project.eu/.

[129] Potchara Pruksasri et al. "Data concealing of supply chain transactions using the Distributed Trust Backbone". In: *The 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014)*. IEEE, Aug. 2014. ISBN: 978-1-908320-39-1. DOI: 10.1109/ICITST.2014.7038796.

[130] R3 Ltd. "Writing Oracle Services". In: *Corda OS 4.8: Documentation and Training for Corda Developers and Operators*. 2021. URL: https://docs.corda.net/docs/corda-os/4.8/oracles.html.

[131] RAFAEL BELCHIOR et al. "A Survey on Blockchain Interoperability: Past, Present, and Future Trends". In: *arXiv* (Mar. 2021).

[132] Gowri Sankar Ramachandran et al. "Trinity: A Byzantine Fault-Tolerant Distributed Publish-Subscribe System with Immutable Blockchain-based Persistence". In: *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, May 2019. ISBN: 978-1-7281-1328-9. DOI: 10.1109/BLOC.2019.8751388.

[133] Drummond Reed et al. *Decentralized Identifiers (DIDs): Core architecture, data model, and representations - W3C Proposed Recommendation Draft*. Ed. by Drummond Reed, Manu Sporny, and Markus Sabadello. W3C, Aug. 2021, –undefined. URL: https://www.w3.org/TR/2021/PR-did-core-20210803/.

[134] Wei Ren, Xutao Wan, and Pengcheng Gan. "A double-blockchain solution for agricultural sampled data security in Internet of Things network". In: *Future Generation Computer Systems* 117 (2021), pp. 453–461. ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2020.12.007. URL: https://www.sciencedirect.com/science/article/pii/S0167739X20330697.

[135] Leonid Reyzin and Sophia Yakoubov. *Efficient Asynchronous Accumulators for Distributed PKI*. Aug. 2016.

[136] Marten Risius and Kai Spohrer. "A Blockchain Research Framework: What We (don't) Know, Where We Go from Here, and How We Will Get There". In: *Business & Information Systems Engineering* 59 (Aug. 2017), pp. 385–409. DOI: 10.1007/s12599-017-0506-0.

[137] Gianluigi Maria Riva. "What Happens in Blockchain Stays in Blockchain. A Legal Solution to Conflicts Between Digital Ledgers and Privacy Rights". In: *Frontiers in Blockchain* 3 (Aug. 2020). ISSN: 2624-7852. DOI: 10.3389/fbloc.2020.00036.

[138] Dominik Roeck, Henrik Sternberg, and Erik Hofmann. "Distributed ledger technology in supply chains: a transaction cost perspective". In: *International Journal of Production Research* 58.7 (Apr. 2020). ISSN: 0020-7543. DOI: 10.1080/00207543.2019.1657247.

[139] Dorit Ron and Adi Shamir. *Quantitative Analysis of the Full Bitcoin Transaction Graph*. 2013. DOI: 10.1007/978-3-642-39884-1.

[140] Matti Rossi et al. "Blockchain Research in Information Systems: Current Trends and an Inclusive Future Research Agenda". In: *Journal of the Association for Information Systems* 20 (Aug. 2019), pp. 1388–1403. DOI: 10.17705/1jais.00571.

[141] Robby Rouben and Alexander Snyers. "Cryptocurrencies and Blockchain: Legal context and implications for financial crime, money laundering and tax evasion". In: *European Parliament's Special Committee on Financial Crimes, Tax Evasion and Tax Avoidance* PE 619.024 (Aug. 2018).

[142] Boriana Rukanova, Roel Huiden, and Yao-Hua Tan. *Coordinated Border Management Through Digital Trade Infrastructures and Trans-National Government Cooperation: The FloraHolland Case*. 2017. DOI: 10.1007/978-3-319-64677-0{_}20.

[143] Sara Ghaemi et al. "A Pub-Sub Architecture to Promote Blockchain Interoperability". In: *arXiv* (Jan. 2021). URL: https://arxiv.org/pdf/2101.12331v1.pdf.

[144] Tatsuya Sato and Yosuke Himura. "Smart-Contract Based System Operations for Permissioned Blockchain". In: *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. IEEE, Aug. 2018. ISBN: 978-1-5386-3662-6. DOI: 10.1109/NTMS.2018.8328745.

[145] Samant Saurabh and Kushankur Dey. "Blockchain technology adoption, architecture, and sustainable agri-food supply chains". In: *Journal of Cleaner Production* 284 (2021), p. 124731. ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2020.124731. URL: https://www.sciencedirect.com/science/article/pii/S0959652620347752.

[146] Fabian Schär. "Blockchain Forks: A FormalClassification Framework and PersistencyAnalysis". In: *Munich Personal RePEc Archive* (Aug. 2020). URL: https://mpra.ub.uni-muenchen.de/101712/.

[147] Eder J Scheid et al. "Bifröst: a Modular Blockchain Interoperability API". In: *2019 IEEE 44th Conference on Local Computer Networks (LCN)*. IEEE, Aug. 2019. ISBN: 978-1-7281-1028-8. DOI: 10.1109/LCN44214.2019.8990860.

[148] Lennard Segers et al. "The use of a blockchain-based smart import declaration to reduce the need for manual cross-validation by customs authorities". In: *Proceedings of the 20th Annual International Conference on Digital Government Research*. ACM, Aug. 2019. ISBN: 9781450372046. DOI: 10.1145/3325112.3325264.

[149] Nathalie Silva et al. "Improving Supply Chain Visibility With Artificial Neural Networks". In: *Procedia Manufacturing* 11 (2017), pp. 2083–2090. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2017.07.329. URL: https://www.sciencedirect.com/science/article/pii/S2351978917305371.

[150] Frank G M Smeele. *The bill of lading contracts under European national laws (civil law approaches to explaining the legal position of the consignee under bills of lading)*. Aug. 2020. DOI: 10.4324/9781003122869-12.

[151] Yonatan Sompolinsky, Shai Wyborski, and Aviv Zohar. *PHANTOM and GHOSTDAG: A Scalable Generalization of Nakamoto Consensus*. Aug. 2020. URL: https://eprint.iacr.org/2018/104.pdf.

[152] Jie SONG et al. "Research advances on blockchain-as-a-service: architectures, applications and challenges". In: *Digital Communications and Networks* (Aug. 2021). ISSN: 23528648. DOI: 10.1016/j.dcan.2021.02.001.

[153] Robert Spekman and Niklas Myhr. "An Empirical Investigation into Supply Chain Management: A Perspective on Partnerships". In: *Supply Chain Management: An International Journal* 3 (Aug. 1998), pp. 53–67. DOI: 10.1108/13598549810215379.

[154] Manu Sporny et al. *Verifiable Credentials Data Model: Expressing verifiable information on the Web - W3C Recommendation*. Ed. by Manu Sporny, Dave Longley, and David Chadwick. W3C, Aug. 2019. URL: https://www.w3.org/TR/2019/REC-vc-data-model-20191119/.

[155] Eveline van Stijn et al. "Enhancing business and government interactions in global trade". In: *Third International Engineering Systems Symposium CESUN 2012*. Aug. 2012.

[156] John Strassner et al. "The Use of Context-Aware Policies and Ontologies to Facilitate Business-Aware Network Management". In: *Journal of Network and Systems Management* 17.3 (Aug. 2009). ISSN: 1064-7570. DOI: 10.1007/s10922-009-9126-4.

[157] John Strassner et al. "An architecture for using metadata to manage ubiquitous communications and services: A position paper". In: *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE, Aug. 2010. ISBN: 978-1-4244-6605-4. DOI: 10.1109/PERCOMW.2010.5470674. URL: https://ieeexplore.ieee.org/document/5470674.

[158] Jianwen Su et al. *Towards a Theory of Web Service Choreographies*. DOI: 10.1007/978-3-540-79230-7{_}1.

[159] Ali Sunyaev. "Information Systems Architecture". In: *Internet Computing: Principles of Distributed Systems and Emerging Internet-Based Technologies*. Cham: Springer International Publishing, Feb. 2020. Chap. 2. DOI: 10.1007/978-3-030-34957-8{_}2.

[160] Ali Sunyaev et al. "Token Economy". In: *Business & Information Systems Engineering* (Aug. 2021). ISSN: 2363-7005. DOI: 10.1007/s12599-021-00684-1.

[161] Sut Sakchutchawan. "An Inquiry into the Strict Compliance of the International Chamber of Commerce Trade Rules in Financing Process". In: *Global Journal of International Business Research* 2.2 (Aug. 2009).

[162] Yao-Hua Tan et al., eds. *Accelerating Global Supply Chains with IT-Innovation*. Springer Berlin Heidelberg, 2011. ISBN: 978-3-642-15668-7. DOI: 10.1007/978-3-642-15669-4.

[163] Shuyang Tang. *Bracing A Transaction DAG with A Backbone Chain*. 2020. URL: https://eprint.iacr.org/2020/472.pdf.

[164] Paolo Tasca and Claudio Tessone. "A Taxonomy of Blockchain Technologies: Principles of Identification and Classification". In: *Ledger* 4 (Aug. 2019). DOI: 10.5195/ledger.2019.140.

[165] Paolo Tasca and Thayabaran Thanabalasingham. "Ontology of Blockchain Technologies. Principles of Identification and Classification". In: *SSRN Electronic Journal* (2017). ISSN: 1556-5068. DOI: 10.2139/ssrn.2977811.

[166] The Digital Transport Logistics Forum. "Enabling organisations to reap the benefits of data sharing in logistics and supply chain". Aug. 2018. URL: https://www.dtlf.eu/sites/default/files/public/uploads/fields/page/field_file/executive_summary2_reading__0.pdf.

[167] Thomas Hardjono, Alexander Lipton, and Alex Pentland. "Wallet Attestations for Virtual Asset Service Providers and Crypto-Assets Insurance". In: *arXiv* (May 2020). URL: https://arxiv.org/pdf/2005.14689.pdf.

[168] Edvard Tijan et al. "Blockchain Technology Implementation in Logistics". In: *Sustainability* 11.4 (Aug. 2019). ISSN: 2071-1050. DOI: 10.3390/su11041185.

[169] S Tönnissen and F Teuteberg. "Analysing the impact of blockchain-technology for operations and supply chain management: An explanatory model drawn from multiple case studies". In: *International Journal of Information Management* 52 (2020).

[170] Trade Commission. "Trade for All: Towards a More Responsible Trade and Investment Policy". In: *European Commission: Trade Commissioner Cabinet* (2015). URL: https://trade.ec.europa.eu/doclib/docs/2015/october/tradoc_153846.pdf.

[171] TradeLens. *Solution Brief: Edition 3*. Tech. rep. URL: https://www.ibm.com/downloads/cas/B4K3R1MP.

[172] Transport Commission. "Ports 2030: Gateways for the Trans European Transport Network". In: *European Commission: Transport Commissioner Cabinet* (2013). URL: https://ec.europa.eu/transport/infrastructure/tentec/tentec-portal/site/brochures_images/ports2013_brochure_lowres.pdf.

[173] Sofia Umaroh and Afriyanti Kartika. "A framework for an IT use policy development". In: *EEA - Electrotehnica, Electronica, Automatica* 66 (Aug. 2018), pp. 159–165.

[174] United Nations. "United Nations Convention on the Carriage of Goods by Sea (Hamburg Rules)". In: *Treaty Series* 1695 (Aug. 1978). URL: https://treaties.un.org/doc/Treaties/1992/10/19921001%2004-36%20AM/Ch_XI_D_3.pdf.

[175] United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT). "United Nations Layout Key for Trade Documents Recommended Practice and Guidelines: Recommendation No. 1". In: *United Nations Economic Commission for Europe* ECE/TRADE/432 (2017).

[176] United Nations Commission of International Trade Law. "United Nations Convention on Contracts for the International Carriage of Goods Wholly or Partly by Sea". In: *United Nations Publishing and Library Section* (2009). URL: https://uncitral.un.org/sites/uncitral.un.org/files/media-documents/uncitral/en/rotterdam-rules-e.pdf.

[177] John Venable, Jan Pries-Heje, and Richard Baskerville. "A Comprehensive Framework for Evaluation in Design Science Research". In: *Design Science Research in Information Systems. Advances in Theory and Practice*. Ed. by K. Peffers, M. Rothenberger, and B. Kuechler. Vol. 7286. Belin/Heidelberg: Springer, 2012. DOI: 10.1007/978-3-642-29863-9{_}31.

[178] Peter Verhoeven, Florian Sinn, and Tino Herden. "Examples from Blockchain Implementations in Logistics and Supply Chain Management: Exploring the Mindful Use of a New Technology". In: *Logistics* 2.3 (Aug. 2018). ISSN: 2305-6290. DOI: 10.3390/logistics2030020.

[179] Piet Verschuren and Rob Hartog. "Evaluation in Design-Oriented Research". In: *Quality & Quantity* 39.6 (Dec. 2005). ISSN: 0033-5177. DOI: 10.1007/s11135-005-3150-6.

[180] Jyri Vilko, Paavo Ritala, and Jukka Hallikas. "Risk management abilities in multimodal maritime supply chains: Visibility and control perspectives". In: *Accident Analysis & Prevention* 123 (2019), pp. 469–481. ISSN: 0001-4575. DOI: https://doi.org/10.1016/j.aap.2016.11.010. URL: https://www.sciencedirect.com/science/article/pii/S0001457516304031.

[181] Wattana Viriyasitavat et al. "Blockchain-based business process management (BPM) framework for service composition in industry 4.0". In: *Journal of Intelligent Manufacturing* 31.7 (Aug. 2020). ISSN: 0956-5515. DOI: 10.1007/s10845-018-1422-y.

[182] Joseph G Walls, George R Widmeyer, and Omar A El Sawy. "Building an Information System Design Theory for Vigilant EIS". In: *Information Systems Research* 3.1 (Aug. 1992), pp. 36–59. ISSN: 1047-7047. DOI: 10.1287/isre.3.1.36.

[183] Gang Wang. "SoK: Exploring Blockchains Interoperability". In: *IACR Cryptology ePrint Archive* (Aug. 2021). URL: https://eprint.iacr.org/2021/537.pdf.

[184] Licheng Wang et al. "Cryptographic primitives in blockchains". In: *Journal of Network and Computer Applications* 127 (Feb. 2019). ISSN: 10848045. DOI: 10.1016/j.jnca.2018.11.003.

[185] Qin Wang et al. *SoK: Diving into DAG-based Blockchain Systems*. Aug. 2020. URL: https://arxiv.org/pdf/2012.06128.pdf.

[186] Ping Wah Wong and N Memon. "Secret and public key image watermarking schemes for image authentication and ownership verification". In: *IEEE Transactions on Image Processing* 10.10 (2001). ISSN: 10577149. DOI: 10.1109/83.951543.

[187] World Economic Forum and Deloitte. "Inclusive Deployment of Blockchain for Supply Chains: Part 6 – A Framework for Blockchain Interoperability". In: (Aug. 2020). URL: http://www3.weforum.org/docs/WEF_A_Framework_for_Blockchain_Interoperability_2020.pdf.

[188] Wout Hofman. *Personal Communication*. Sept. 2021.

[189] Kimchai Yeow et al. "Decentralized Consensus for Edge-Centric Internet of Things: A Review, Taxonomy, and Research Issues". In: *IEEE Access* 6 (2018). ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2779263.

[190] Liyang Yu. *The Building Block for the Semantic Web: RDF*. 2014. DOI: 10.1007/978-3-662-43796-4{_}2.

[191] H Yusuf and I Surjandari. "Comparison of Performance Between Kafka and Raft as Ordering Service Nodes Implementation in Hyperledger Fabric". In: *International Jour nal of Advanced Science and Technology* 29.7 (2020), pp. 3549–3554. URL: http://sersc.org/journals/index.php/IJAST/article/view/17652.

[192] Zac Mitton. "Adding Merkle-Mountain-Ranges and Fly-Proofs to ETC". In: *Ethereum Classim Summit*. Vancouver, 2019.

[193] Alexei Zamyatin et al. *SoK: Communication Across Distributed Ledgers*. Aug. 2019. URL: https://eprint.iacr.org/2019/1128.pdf.

[194] Jian Zhang. "Deploying Blockchain Technology in the Supply Chain". In: *Computer Security Threats*. IntechOpen, Sept. 2020. DOI: 10.5772/intechopen.86530.

[195] Shijie Zhang and Jong-Hyouk Lee. "Analysis of the main consensus protocols of blockchain". In: *ICT Express* 6.2 (Aug. 2020). ISSN: 24059595. DOI: 10.1016/j.icte.2019.08.001.

[196] Xuesong Zhang, Lorne Olfman, and Daniel Firpo. "An Information Systems Design Theory for Collaborative ePortfolio Systems". In: *2011 44th Hawaii International Conference on System Sciences*. IEEE, Aug. 2011. ISBN: 978-1-4244-9618-1. DOI: 10.1109/HICSS.2011.56.

[197] Zhiyi Zhang et al. "DLedger: An IoT-Friendly Private Distributed Ledger System Based on DAG". In: *arXiv* (Aug. 2019). URL: https://arxiv.org/pdf/1902.09031.pdf.

[198] Tong Zhou, Xiaofeng Li, and He Zhao. "DLattice: A Permission-Less Blockchain Based on DPoS-BA-DAG Consensus for Data Tokenization". In: *IEEE Access* 7 (2019). ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2906637.

[199] Ziyuan Li et al. "Implementing a sidechain-based asynchronous DPKI". 2021. URL: https://www.researchgate.net/publication/349097289_Implementing_a_sidechain-based_asynchronous_DPKI.

[200] G R Zomer. "Smart Trade Logistics - Compliance as an Opportunity". In: *10th International Electronic Government Conference 2011*. Ed. by Sietse Overbeek, Yao-Hua Tan, and Gerwin Zomer. CEUR Workshop Proceedings, Aug. 2011, pp. 9–19.

Appendix A

Scientific Paper

An interoperable and self-sovereign data sharing architecture to aggregate import declaration data from multiple blockchain platforms

L.A. CABRERA MOSCA*, Delft University of Technology, The Netherlands

Abstract: Increasingly specialised logistic services are triggering the disaggregation of supply chain functions and fostering the generation of information silos. This is perceived by European customs as a threat, since it affects the reliability of the import risk assessments they conduct. Recently, commercial data sharing platforms based on blockchain technology (*BCT*) are allowing to expedite the verification of trade finance documents. The cross-organisational trust achieved in these platforms has driven the digitisation of the bill of lading, from which entry summary declarations (*ENS*) used by European customs are formed. The latter see combining data from multiple platforms an opportunity to improve supply chain visibility, turn declarations more agile and risk assessments more effective. However, it remains unclear how to integrate declaration procedures in these data ecosystems. There are two major barriers to overcome. Firstly, the lack of interoperability solutions to make platform architectures compatible for the aggregation of declaration data. Secondly, the need to adapt available identity management solutions to the distributed nature of these platforms to promote trust between declarants and their clients. To tackle this, a peer-to-peer architecture is presented as a novel solution to migrate from declaration based on data duplication towards information sharing based on links to the original and trusted data stored in *BCT* platforms.

CCS Concepts: • **Computer systems organization** → **Peer-to-peer architectures**; • **Security and privacy** → **Information accountability and usage control**; • **Applied computing** → **Transportation**; • **Social and professional topics** → **Transborder data flow**.

Keywords: supply chain visibility, international shipping, data sovereignty, import declarations, blockchain interoperability

1 INTRODUCTION

As the nature of global trade and commerce shift, so does the structure of the logistic processes driving border risks. For instance, the tendency to disaggregate supply chain functions leads to an increasing number of actors, which in turn produces decentralized knowledge within supply chains [65]. The currently expected global cargo mobility implies complex custody chains involving providers of increasingly specialised logsitic services, such as freight forwarding, warehousing and other commonly outsourced activities.

On the eyes of customs administrations, the result is information fragmented between these service providers and hidden by an opaque and complex network of commercial agreements. Therefore, the ability to assess the impact of imports on national interests commences to rely on an unprecedented level of collaboration between customs and a growing number of enterprises. This is perceived by European institutions as a threat, which raises the flag on their obligation to maintain sufficient end-to-end visibility on the economic activities carried out across supply chains entering the European Union (*EU*).

Author's address: L.A. Cabrera Mosca, Delft University of Technology, Technology Policy and Management Faculty, Jaffalaan 5, 2628BX Delft, The Netherlands, l.a.cabreramosca@student.tudelft.nl.

In parallel, commercial platforms based on blockchain technology (*BCT*) aimed at improving overall supply chain visibility and solving inefficiencies in international shipping have gained considerable popularity among shipping lines, freight forwarders and other supply chain actors. *BCT* enables trustless collaboration systems where entities can share information without the need to assess their degree of trust towards other participants [37]. It's main advantage is the set of modern cryptographic mechanisms through which transaction records become immutable, meaning that they can not be altered once stored in a blockchain [50].

These platforms leverage *BCT* to provide enterprises with a secure environment where to deploy distributed applications to process trade documents, such as letters of credit (*L/C*) or bills of lading (*B/L*). A *B/L* is the foundation of a cargo custody chain, and is used to legally bind the conditions of contracts of carriage between logistic service providers [63]. Data is retrieved from documents generated across supply chains, such as the *B/L*, and then aggregated in the import declarations used by European customs to perform risks assessments, known as entry summary declarations (*ENS*) [18]. Therefore, the digitisation of the *B/L* is a milestone in the automation of import declarations and the modernisation of customs procedures in general. However, *B/L* data is still processed multiple times by different actors before their use to generate an *ENS*, which makes it prone to incompleteness and inconsistencies [15, 45].

In this context, European customs identify the rise of commercial blockchain platforms as an opportunity to leverage increased data availability to feed the *ENS* generation process with original, accurate, trusted logistic data collected directly from its original source [19, 44]. Instead of relying on the last cargo custodian to forward the information accumulated downstream a supply chain, customs risk assessments could benefit from retrieving logistic data earlier at each stage of a supply chain [60] or maintain trusted links to this data [26, 45]. However, despite the positive impact that these platforms are bringing to the private sector, there is a lack of research on the integration of customs declaration procedures in blockchain ecosystems such as the one shown in Figure 1.

To bridge this gap, this paper explores the design of a peer-to-peer data sharing architecture able to aggregate distributed applications for the generation of *ENS* from verifiable links to the internal ledger state proofs of multiple blockchain platforms. This is expected to benefit European customs by reducing fraud and enhancing the reliability of their import risk assessments, as well as to exporters, importers and shipping lines by reducing the bureaucratic friction of customs declarations.

The structure of the paper is the following. A review of the relevant literature on the use of *BCT* in supply chains, blockchain interoperability solutions and trends in decentralised credential management is discussed in section 2. This is followed by section 3

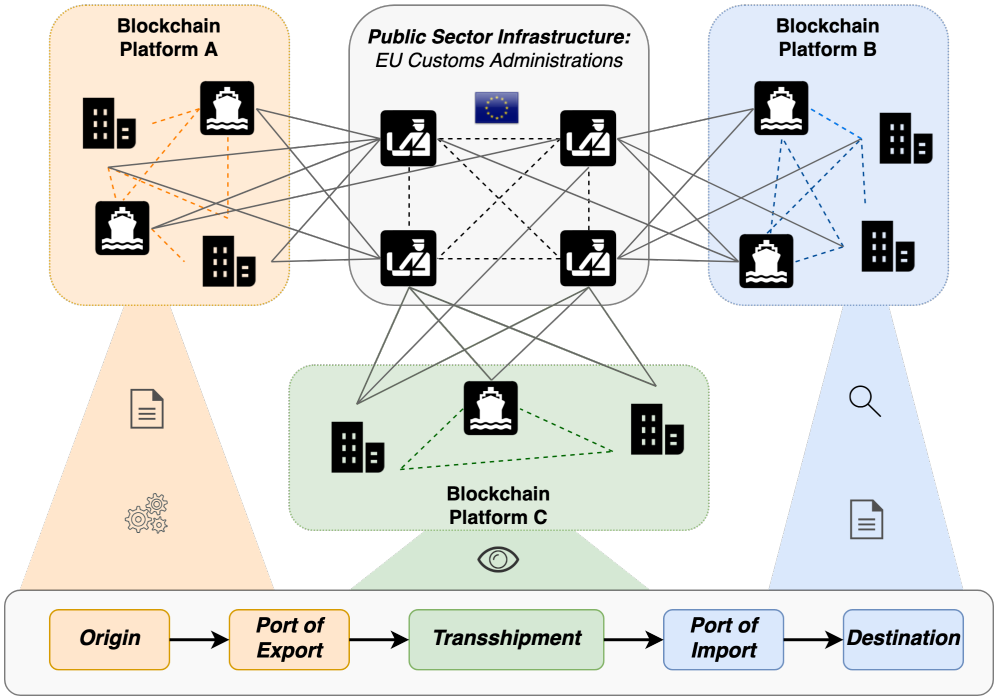


Fig. 1. Blockchain platform ecosystem.

where the function of each layer is covered. Finally, section 4 discusses the research conclusions and recommendations for further research.

2 LITERATURE REVIEW

2.1 Blockchain Technology in Supply Chain

A blockchain is a record of transactions bundled and assigned to a blocks connected sequentially building a chain that is updated simultaneously by all the participants in a network [17]. When applied to a distributed computer network, the result is a decentralised peer-to-peer database, in which every node keeps an independent copy of the history of executed transactions [54]. One of the most attractive features is their immutability: it impossible for a single party in the network to alter the history of transactions once stored in the blockchain [66].

Overall, *BCT* is considered to have the ability to improve global supply chain visibility standards [29]. In the recent past, *BCT* has emerged as a disruptive technology with innumerable applications in the transport and shipping industry [19]. Improved cost effectiveness by simplifying the tracking of items and transactions [22], increasing shipping flexibility [34] and a considerable reduction of supply chain risks [11] have all been shown achievable benefits of integrating *BTC* in the logistics sector.

The source of the interest in using *BCT* to tackle the latest challenges in large-scale logistic data dissemination is threefold. First, a dire need to cope with a fast increase in supply chain complexity, both at operational and organizational level [9]. Second, the leading businesses' fear to fail at staying at the vanguard of their industries [62], fueled by *BCT*'s recognised potential to reshape the foundations of existing business models [22]. And lastly, the success examples of early adopters. The leverage of *BCT* for non-financial applications is a reality, being widely accepted to have reached sufficient level of maturity to produce tangible results in real-world problems [62]. From supply chain transparency for ethical sourcing in the fashion industry to counterfeit prevention and product authentication in pharmaceutical distribution [11, 25], *BCT* is being slowly adopted in varied logistics applications.

Recently, the logistics sector has reduced collaboration friction and enhanced trust with blockchain-based data sharing platforms. These platforms offer logistic service providers with secure information exchange services that help optimising contractual information flows and decreasing risks by improving the reliability of forecasts. An example is *TradeLens*, a joint venture between *Maersk* and *IBM* aimed at improving supply chain visibility, and in doing so, increasing the efficiency in containerised shipping [28].

Further work on the standardisation and privacy of blockchain-based data objects is required before automated trade document processing is embraced as a reliable solution for inter-organisational data sharing in the shipping industry [72]. Initiatives such as the aforementioned *Tradelens* intend to close this gap by providing enterprises and institutions with a single source of shipping data [28, 42]. However, these initiatives tend to focus on specific supply chain segments. They do not capture all the transactions executed throughout the complete logistics domain. For example, *Tradelens*

supports an ample ecosystem of actors, but focuses on ocean shipping and excludes activities performed outside port terminals (*e.g.*, pre-carriage and on-carriage multimodal transshipment [44]. This reduces the visibility between enterprises and institutions and is one of the reasons it is difficult to integrate information flows under the same blockchain architecture. In addition, there is a lack of research on the interaction between commercial *BCT* ecosystems and customs administrations, such as the lodging of *ENS* [54].

2.2 Blockchain Interoperability

In the research context, blockchain interoperability refers to the communication standards between blockchain platforms that allow them to agree on the interpretation of information [32]. Interoperability is crucial for long-term industry transformations, because the widespread implementation of *BCT* solutions with insufficient interoperability can prevent blockchain applications from solving multi-actor collaboration issues at a socio-technical scale [61]. There is insufficient academic effort on the standardization of communication protocols that allow two different ledgers to share their internal states and coordinate application logic [32, 67]. There is also no consensus on the classification of interoperability solutions. An approach is to differentiate between chain-based, bridge-based and *dApp*-based solutions [67], which can be alternatively referred to as public connectors, hybrid connectors and *blockchains of blockchains* respectively [47].

Chain-based solutions focus on the chain-to-chain interactions behind atomic swaps. Bridge-based solutions build connections between blockchains to reduce or remove large technical incompatibilities between layer components. The purpose of *dApp*-based solutions is to ease the implementation of and interaction between decentralised peer-to-peer applications, and represent a more holistic approach to interoperability linked to the emerging *Blockchain-as-a-Service* design paradigm (*BaaS*) [36, 55]. Among these categories, literature distinguishes four subcategories relevant for the proposed application: sidechains, notary schemes, hash-locks and trusted relays [2, 16, 20, 23, 32, 41, 47, 59, 67, 71].

Sidechains act as complementary chains build around a mainchain [67]. They are used as buffers to delegate certain phases of resource transfer protocols, and can be designed as one-way or two-way systems [47]. However, the number of sidechains required in the blockchain environment described in Figure 1 would grow at an unsustainable rate, as well as adding unpredictable complexity to the maintenance of the design.

Notary schemes rely on a trusted third party to monitor multiple blockchains, witness the terms of cross-chain commitments and trigger the execution of contracts [16, 32]. A popular application of notary schemes are centralised cryptocurrency exchanges, where the security of token transfers is guaranteed by the platform provider [47, 67].

The next subcategory are hash-locks. They can be described as decentralised escrow services that can alter the ownership of assets without relying on a trusted third party, unlike notary schemes [23, 32]. They can be chained after each other [41], which makes them particularly useful when transaction sequences want to be programmed between entities with no direct connections [47]. Hash-locks can be implemented as smart contracts triggered by arbitrary conditions, such as the time limits for the provision of cryptographic proofs used in hash time-lock contracts (*HTLC*) [67]. Also, when combined with the appropriate network configuration, hash-locks allow a group of entities operating in independent blockchains to exchange proofs of the internal state of their ledgers [5, 13].

The last solution type are trusted relays, with a focus on trusted gateway bridges [23]. Also referred to as *relay services* [2] or *chain relays* [71], they handle requests to fetch ledger state proofs between remote networks and verify application logic [2, 48]. Relay services make it possible for an entity on a chain to verify events registered in other chains by building a bridge that provides smart contract services between platforms [67]. This means that a smart contract implemented in one chain can become a *client* of another chain [71]. For the research context, the main advantage is that they allow *clients* to define arbitrary business logic that can be fed with evidences of external data without a centralised entity [20, 47]. Research on more advanced blockchain-agnostic protocols, consensus engines and security infrastructures to aggregate complete blockchain architectures are being developed [47]. However, they fail to offer backward compatibility, implying that legacy systems would need to be heavily modified. In view of this, trusted relays are seen as the most realistic solution to link network layers between permisisoned blockchains while maintaining a design philosophy inspired in the *d-App* paradigm [23, 47].

2.3 Decentralised Credential Management

Digital identities have experienced a transformation in the recent past, starting at centralised identities, moving towards federated identities and arriving to modern decentralised identities [**?**]. Centralised identities link a data subject to a digital environment through dedicated credentials. These credentials are managed under internal environment rules and are not recognised by other digital environments.

The figure of *identity provider* (*IDP*) was introduced later in order to reduce the limitations of centralised identities [**?**]. An *IDP* acts as intermediary between subjects and digital environments enabling reusable credentials to be trusted by more than one digital environment, which is the core concept behind federated identities. Although the level of centralisation is reduced, data subjects still rely on an external party to control the legitimacy of their digital identities and not all digital environments might use the same *IDP*.

The next step in the evolution of digital identities are decentralised identities, which remove the need to rely on intermediaries to engage in digital transactions by fully operating one's digital identity. The characteristic attribute of decentralised identities is the swift from *account-based* access control towards trusted links between digital peers, whether data subjects represent a natural person, an organisation or a digital resource [**?**]. In this context, the term *self-sovereign identity* represents not depending on any organisation to make use of your digital identity to legitimise digital transactions.

Decentralised identity certification is becoming an integral part of information sharing in the *BCT* era [14, 19]. However, architectures

including self-sovereign identity management are miss-represented in literature: providing enterprises incentives to share data by equipping them with data governance privileges to control the exposure of their information.

The use of the decentralised identities to combat the weaknesses of conventional identity management has been strongly influenced by the adoption of *BCT* [4, 12]. Innovative public-key authentication and verifiable data registries can provide the certainty that an entity is linked to the public key being used in transactions [4, 56]. This is known in general as *public key infrastructure* (*PKI*), which has traditionally depended on the aforementioned *IDP*s [69]. To overcome this, *BCT* can be the backbone technology of trusted verifiable registries to achieve functional decentralised identities without the need for *IDP*s [74]. In that case, a more appropriate term for the use of *DLT* for *PKI* applications is *distributed public key infrastructure* (*DPKI*) [51].

In the absence of *IDP*s, self-certifying identifiers fulfil their two main functions: binding a public key to an identifier and binding an identifier to a credential holder [49]. This model is shown in Figure 2. Trust in identifiers is achieved by creating cryptographic bonds with public keys. This is effective as the cryptosystems used to authenticate these links are the same trusted methods used to bind private and public keys, such as Rivest–Shamir–Adleman (*RSA*) algorithms, elliptic curves or the Diffie-Hellman algorithm [1, 31, 33]. *IDP*s are not needed as long as the authentication is trusted, because a legitimate holder is the only entity able to generate pairs of identifiers and public keys compliant with the aforementioned cryptosystems.

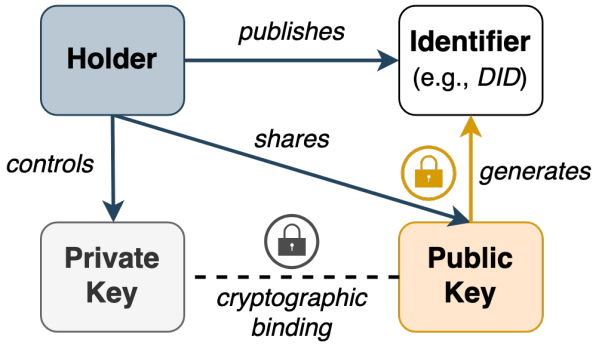


Fig. 2. Self-certifying identifier model, adapted [43].

3 ARCHITECTURE DESIGN

3.1 Design Approach

The problem at hand is driven by three principles: visibility of logistic events throughout supply chains, data sovereignty of supply chain actors and the interoperability of their information systems. An overview of the relationship between principles is described in Figure 3. Each pair of design principles has been found to bring different benefits in the context of supply chain data sharing. Event visibility entails that a larger number of actors across all supply chain segments keep track of the logistic events associated with their economic activities. Interoperability makes it technically possible to exchange data between the information systems used in each supply chain segment. Data sovereignty is an actor´s ability to control the level of exposure of his own identities and business information. Event visibility and interoperability produce potential connections between information systems. Event visibility and data sovereignty incentivise supply chain actors to incorporate externally validated data in their business processes, which in turn elevates trust. Finally, interoperability and data sovereignty foster the integration of these data exchanges in their business activities. Enhanced data quality is achieved in practice next to the three relationships.

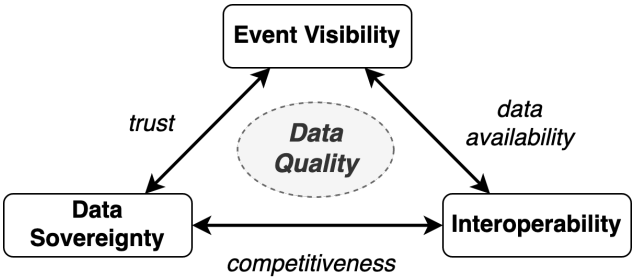


Fig. 3. Design value framework.

Enhanced data quality, and therefore increased reliability during risk assessments, is the main benefit for European customs. However, it would be naive to take competitiveness incentives and consolidated trust between trade stakeholders for granted. This means that the quality of the declaration data does not only depend on the data infrastructure used by customs to interact with carriers to lodge an *ENS*, but also on the previous data exchanges between carriers and other logistic service providers. Even if an interface able to automatically collect declaration data based on the operations of a carrier was used, data quality still relies on the information provided to the carrier by freight forwarders and other logistic service providers. Therefore, the challenge faced by customs originates in the data exchanges previous to the generation of the *ENS*.

From a process-oriented perspective, the challenge can interpreted as the need to model behavior in a multi-actor system based on data sharing patterns [58]. This is closely related to *service orchestration*: using standardised protocols and high-level languages to integrate architectures, and in doing so, achieve stronger relationships between businesses and their information systems [10, 24]. This approach consists on finding links between the application and business logic supporting these services, which can be very effective against semantic heterogeneity and making collaboration between supply chain actors easier.

However, the context in which services and transactions are executed must be also taken into consideration, which can be done with an additional domain-based framework able to represent the underlying entity interactions in more detail. Synchronising information flows between private domains requires a data sharing architecture in which transaction queries and index data are somehow captured [57]. The example shown in Figure 4 provides domain interoperability by using a common data repository. It stores links between

pieces of private data so that information produced in one domain can propagate application logic in other domains. This way, supply chain actors can operate following internal domain policies while allowing their data to feed otherwise isolated services.

Different versions of this concept have been adapted to facilitate ontology integration [40], the implementation of context-aware data dissemination policies [57], *Linked Open Data* (*LOD*) for *B2G* communication [26, 45, 64] and blockchain-based supply chain monitoring [48].

3.2 Architecture Overview

The research presents a novel approach to migrate from import declarations based on duplication towards information sharing based on links to original and trusted data stored in blockchain platforms. Three layers are used: *Cross-chain Communication*, *Credential Management* and *Event Visibility*. An overview is shown in Figure 5. The *Cross-chain Communication* layer uses an overlay network to enable cross-platform peer-to-peer interactions, observe and share ledger states via trusted gateways and propagate self-sovereign smart contract logic. The *Credential Management* layer uses decentralised identifiers (*DID*s) as an alternative to the currently available identity certification solutions, allowing data owners to be in full control of the exposure of their digital identities and business information and avoid challenges related to key rotation and certificate revocation. Lastly, *Event Visibility* proposes a directed acyclic graph (*DAG*) ledger environment where to deploy decentralised applications. These applications are used to combine the ledger states of independent logistic blockchains, model the issue of trade documents throughout the whole cargo custody chain, and eventually detect anomalies in the framework agreements between logistic partners.

3.3 Cross-chain Communication Layer

Gateways are used to relay (connect) a *client* and a *source*. They are normally dedicated nodes, but can be implemented as an additional service layer within a permissioned network [23]. This depends on the level of centralisation in the architecture and consensus mechanism used. In any case, gateways that represent a group of nodes maintaining independent ledgers can be grouped. Assuming every node can interact with its gateway, peers can leverage the functionalities of their ledger to offer *ad hoc* services to external clients [39]. Creating a logical layer above the networks represented by each gateway to provide these services results in an overlay network, which is shown graphically in Figure 6.

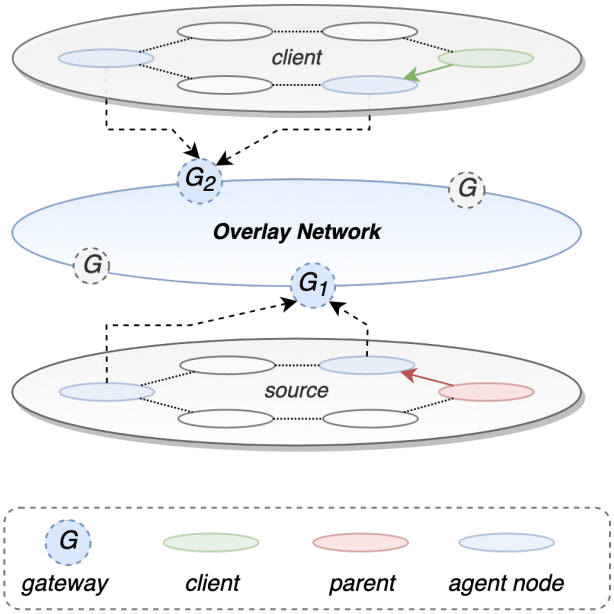


Fig. 6. Overlay network.

The model includes a *client* and a *source*. This terminology should not be confused with the terms *sender* and *receiver* used in atomic swaps, *e.g.*, transfers of fungible tokens linked to account balances for financial applications. The goal is rather to describe an environment where information (or traces towards information) about ledger states becomes accessible under certain rules to legitimate entities. Depending on the use case, this may mean migrating the copy of an asset between ledgers or only forwarding transaction proofs to propagate application logic between independent ledgers.

Agent nodes are also part of the model. Generally referred to as committee members [3], they have increased visibility over a ledger´s activity due to governance privileges. They might be irrelevant for less sophisticated platform architectures, but they play a crucial role in the publication of internal ledger states for platforms using certain *BCT*s, such as peer agents in *Hyperledger Fabric* [3, 23] or oracle nodes in *R3 Corda* [38, 46]. The design takes into account, that a blockchain node might access its gateway in different ways depending on the agent node configuration of its platform.

One of the drawbacks of relayed networks is their static nature, meaning that their participants must know each other´s identities and configurations *a priori* [2, 47]. While this might not be an issue for permissioned environments with fixed participants, network discovery is important when participants are added and removed dynamically, which is the case for the research context. Modular designs are able to improve dynamic discoverability with a credential registry next to a publisher-subscriber (*pub-sup*) system [52, 53]. The logic of the latter is shown in Figure 7. It allows sources to share application logic and *clients* to receive verifiable updates via their gateways. The credential management layer controls access to the overlay infrastructure by processing self-certifying identities based on a decentralised identifier (*DID*) method (see subsection 3.4).

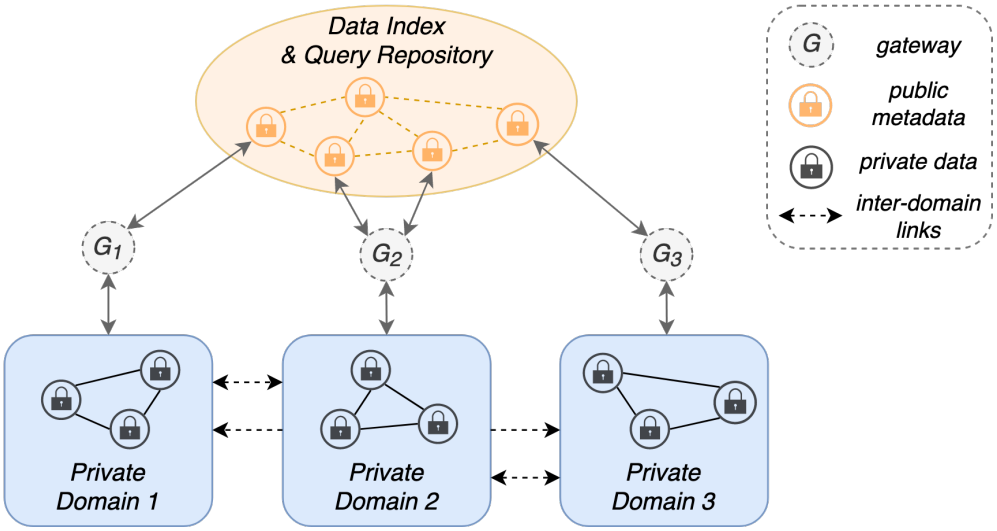


Fig. 4. Conceptual data sharing model, adapted [57].

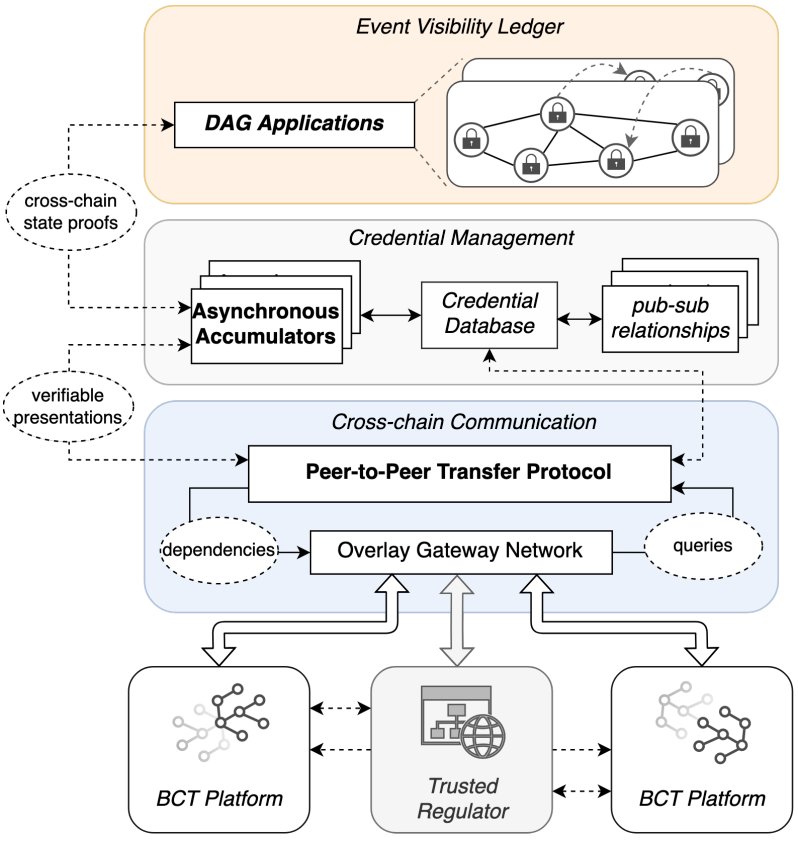


Fig. 5. Detailed architecture design.

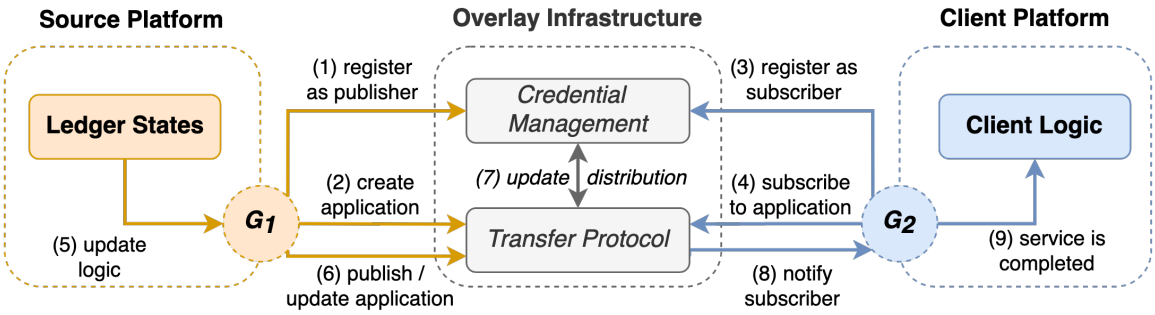


Fig. 7. Publisher-subscriber system, adapted [52].

3.4 Credential Management Layer

The digital identifies of organisations and resources are modelled in the design using verifiable credentials based on decentralised identifiers (*DIDs*). The term identifier indicates a uniform resource identifier (*URI*), which is a string of characters used to represent physical resources that are not network-accessible (*i.e.*, persons, locations, etc.), as well as logical representations of objects retrieved from an information system, such as electronic documents or any other digital asset [70]. A *DID* is a *URI* management schema meant to be a component of larger information systems built around the verifiable credential [56].

DIDs represent a solution to enable key rotation and recovery when using self-certifying identifiers [43, 49]. From a design perspective this is specially relevant for data piggybacking. Traditionally, copies of digital resources have been used for data sharing, which implies coordinating numerous entities on how to generate credentials, authenticate identities and regulate access control to digital environments where resource versions are stored.

The result are networks of networks of co-referenced resources with complex resolution requirements [40]. This means it is becoming difficult to ensure the visibility of information stored in varying formats and authenticated through different methods. Also, public institutions face the risk of failing at monitoring the behavior of entities operating with data in these networks. An example for the research context is the visibility of records stored in blockchains representing the state of logistic processes and agreements digitally. If both trusted identifiers and discoverable resources could be implemented through a universal schema, the data sovereignty and interoperability barriers set by semantic heterogeneity could be reduced [21, 32]. Thus, credential authentication and access control based on *DIDs* can help public institutions piggyback on digital resources operated by logistic service providers in different digital platforms.

DIDs support the delegates, which are entities to whom a controller has granted permission to execute a verification method associated with its *DID*. This is useful for supply chain data piggybacking, because supply chain actors can delegate the verification of each other´s credentials once they have been shared to other parties in other platforms downstream a cargo custody chain.

Another advantage of implementing *DID*-based trust for information sharing, is that the model does not depend on a particular cryptography for the interpretation of *DIDs*. Therefore, the trust model can be implemented as an additional layer on top of legacy systems that rely on identifiers based on the centralised or federated paradigm.

3.5 Event Visibility Layer

The architecture leverages the properties of *directed acyclic graph* ledgers, or *DAG*. Mathematically, a *directed acyclic graph* is a finite set of nodes connected by unidirectional edges where no directed cycles exist [27], meaning that feedback loops cannot be generated. In this approach verifiable presentations are not bundled and stored in blocks, thus its nickname *block-less* ledger. Instead, they are linked directly between each other in a network of ledger states binned together by similar cryptographic techniques used in traditional blockchains. This technology is considered the next iteration in *BCT* [27], being sometimes referred to as *Blockchain 3.0* [8], and is particularly promising for use in permissioned ledger interoperability and *Internet of Things* (*IoT*) [35, 68, 73]. Also, *DAG* technology is an interesting option to process and organise cross-chain transactions, acting as an independent reference to validate states between ledgers while allowing third parties to act as auditors (such as customs or any other regulator) [7].

The architecture is based on the CAPER protocol: an asynchronous ledger where different applications run on a number of nodes known as *agents* [6]. An application refers to a private smart contract in which a specific logic is encoded as the rules to process internal transactions. These contracts only run in the nodes of the application. Additionally, rules to process cross-application transactions can be included in public contracts. Languages widely used to encode smart contracts, such as Solidity [30], can be used for both private and public contracts to ensure the deterministic execution of transactions [6].

Sensitive business logic can be kept confidential within an application while standardised procedures can be encoded as public contracts to facilitate the exchange of information to trigger smart contracts in other applications. There are no tested and reliable solutions that allow internal and cross-application transactions between untrusted applications within a ledger. The CAPER concept can overcome this barrier that is hindering the development of efficient, scalable and secure cross-application communication.

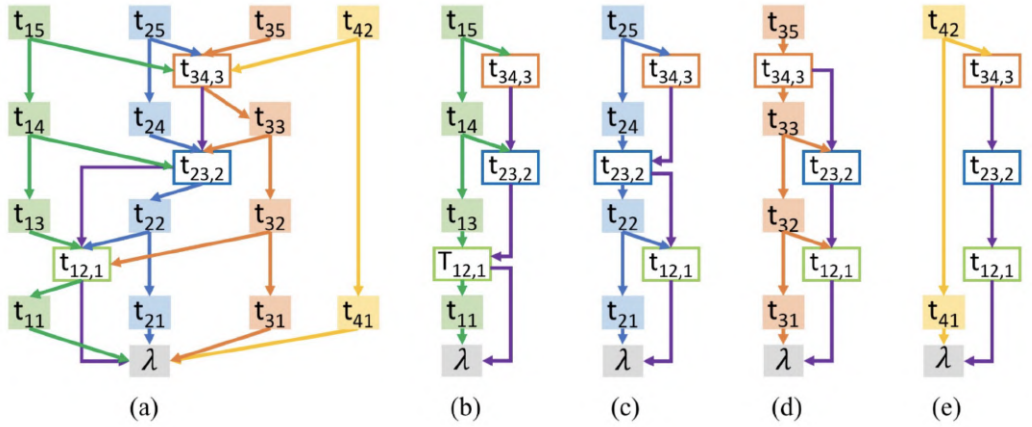


Fig. 8. Example of distributed applications: main *DAG* ledger (a), consisting of four parallel applications (b, c, d, e) [6].

The proposed separation of applications can be used against some data distribution problems faced by carriers and customs administrations. A single record could act as the genesis record of multiple applications. This can be used to trigger additional private contracts in internal application logic to update a private perspective of a supply chain. Moreover, less frequent data duplication can help customs administrations detect dependencies between risk assessment data effectively.

4 CONCLUSIONS & RECOMMENDATIONS

Before European customs can interact with commercial blockchain platforms, a data gathering strategy that takes into account the commercial relationship between actors during data dissemination, while making data access less complex and positively contribute to the institutional duties of European customs, is required. The proposed architecture is a solution help European customs interact with commercial blockchain platforms to retrieve declaration data.

An overlay network is an interesting alternative to power a trusted gateway protocol for cross-chain interactions between permissioned environments. The combination of self-certifying identifiers and the decentralised identity model using *BCT* as verifiable registry can elevate the coordination of services between private and public entities. *DAG* protocols are identified as a powerful tool to enable distributed applications to model and detect anomalies in framework agreements.

The architecture design has focused on trade documents, leaving aside a growing number of data sources, such as sensor-based data collected at transport terminals. Future research should explore how to complement the architecture with these data sources in order to integrate document-based data sharing with logistic data gathered on site. The research has considered the format and content differences between *B/Ls* and import declarations negligible to focus on the high-level design of architecture components. This leaves room for further research on the requirements for the cross-reference of documents between platforms to ensure end-to-end semantic compatibility during cross-validations performed by customs. Lastly, additional research should evaluate the long-term performance consequences of the selected components, such as the feasibility of an additional consensus layer an its requirements in terms of speed and scalability. This is important in order to confirm that the transaction throughput required by European customs is in line with the architecture´s ability to maintain consensus on a growing number of cross-platform references.

ACKNOWLEDGMENTS

This research has been performed in collaboration with the Dutch Organisation for Applied Scientific Research as part of the PROFILE project, an initiative funded by the European Union for the Horizon 2020 research and innovation programme under the Grant Agreement No 786748.

REFERENCES

[1] NaQi ., Wei Wei, Jing Zhang, Wei Wang, Jinwei Zhao, Junhuai Li, Peiyi Shen, Xiaoyan Yin, Xiangrong Xiao, and Jie Hu. 2013. Analysis and Research of the RSA Algorithm. *Information Technology Journal* 12, 9 (8 2013). https://doi.org/10.3923/itj.2013.1818.1824

[2] Ermyas Abebe, Dushyant Behl, Chander Govindarajan, Yining Hu, Dileban Karunamoorthy, Petr Novotny, Vinayaka Pandit, Venkatraman Ramakrishna, and Christian Vecchiola. 2019. Enabling Enterprise Blockchain Interoperability with Trusted Data Transfer (Industry Track). In *Proceedings of the 20th International Middleware Conference Industrial Track*. ACM. https://doi.org/10.1145/3366626.3368129

[3] Ermyas Abebe, Yining Hu, Allison Irvin, Dileban Karunamoorthy, Vinayaka Pandit, Venkatraman Ramakrishna, and Jiangshan Yu. 2021. Verifiable Observation of Permissioned Ledgers. *arXiv* (8 2021). https://arxiv.org/pdf/2012.07339.pdf

[4] Ohoud Albogami, Manal Alruqi, Kholood Almalki, and Asia Aljahdali. 2021. Public Key Infrastructure Traditional andModern Implementation. *International Journal of Network Security* 23, 2 (8 2021), 343–350.

[5] Mohammad Hassan Ameri, Mahshid Delavar, Javad Mohajeri, and Mahmoud Salmasizadeh. 2020. A Key-Policy Attribute-Based Temporary Keyword Search scheme for Secure Cloud Storage. *IEEE Transactions on Cloud Computing* 8, 3 (7 2020). https://doi.org/10.1109/TCC.2018.2825983

[6] Mohammad Javad Amiri, Divyakant Agrawal, and Amr El Abbadi. 2019. CAPER: a cross-application permissioned blockchain. *Proceedings of the VLDB Endowment* 12, 11 (8 2019). https://doi.org/10.14778/3342263.3342275

[7] C Cachin, A D Caro, Pedro Moreno-Sanchez, Björn Tackmann, and M Vukolic. 2017. The Transaction Graph for Modeling Blockchain Semantics. *IACR Cryptol. ePrint Arch.* 2017 (2017), 1070.

[8] Bin Cao, Yixin Li, Lei Zhang, Long Zhang, Shahid Mumtaz, Zhenyu Zhou, and Mugen Peng. 2019. When Internet of Things Meets Blockchain:Challenges in

Distributed Consensus. *arXiv* (8 2019). https://arxiv.org/pdf/1905.06022.pdf

[9] Y Chang, E Iakovou, and W Shi. 2020. Blockchain in global supply chains and cross border trade: a critical synthesis of the state-of-the-art, challenges and opportunities. *International Journal of Production Research* 58, 7 (2020), 2082–2099.

[10] Wei Chen, Jun Wei, Guoquan Wu, and Xiaoqiang Qiao. 2008. Developing a Concurrent Service Orchestration Engine Based on Event-Driven Architecture. https://doi.org/10.1007/978-3-540-88871-0

[11] T . Choi, X Wen, X Sun, and S . Chung. 2019. The mean-variance approach for global supply chain risk analysis with air logistics in the blockchain technology era. *Transportation Research Part E: Logistics and Transportation Review* 127 (2019), 178–191.

[12] YeonSung Chu, Jae Min Kim, YoonJick Lee, SungHoon Shim, and Junho Huh. 2020. SS-DPKI: Self-Signed Certificate Based Decentralized Public Key Infrastructure for Secure Communication. In *2020 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE. https://doi.org/10.1109/ICCE46568.2020.9043086

[13] Bingrong Dai, Shengming Jiang, Menglu Zhu, Ming Lu, Dunwei Li, and Chao Li. 2020. Research and Implementation of Cross-Chain Transaction Model Based on Improved Hash-Locking. In *Blockchain and Trustworthy Systems*, Zheng Z., Dai HN., Fu X., and Chen B. (Eds.). Springer, Singapore, Chapter 17. https://doi.org/10.1007/978-981-15-9213-3{_}17

[14] Simon Dalmolen, Harrie Bastiaansen, Hans Moonen, Wout Hofman, Matthijs Punter, and Erik Cornelisse. 2018. Trust in a multi-tenant, logistics, data sharing infrastructure: Opportunities for blockchain technology.. In *5th International Physical Internet Conference*.

[15] David Hesketh. 2010. Weaknesses in the supply chain: who packed the box? *World Customs Journal, 4 (2), 3–20* 4, 2 (9 2010), 3–20. https://worldcustomsjournal.org/Archives/Volume%204%2C%20Number%202%20(Sep%202010)/02%20Hesketh.pdf

[16] Liping Deng, Huan Chen, Jing Zeng, and Liang-Jie Zhang. 2018. Research on Cross-Chain Technology Based on Sidechain and Hash-Locking. In *Edge Computing - EDGE 2018*, Liu S., Tekinerdogan B., Aoyama M., and Zhang LJ. (Eds.). Springer, Cham, Chapter 12. https://doi.org/10.1007/978-3-319-94340-4{_}12

[17] Daniel Drescher. 2017. *Blockchain Basics*. Apress. https://doi.org/10.1007/978-1-4842-2604-9

[18] European Parliament and Council of the European Union. 2013. Regulation (EU) No 952/2013 of the European Parliament and of the Council of 9 October 2013 laying down the Union Customs Code. *Official Journal of the European Union* 269 (8 2013), 1–101. http://data.europa.eu/eli/reg/2013/952/oj

[19] European Parliamentary Research Service: Panel for the Future of Science and Technology. 2020. Blockchain for Supply Chains and International Trade: Report on Key Features, Impacts and Policy Options. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641544/EPRS_STU(2020)641544_EN.pdf

[20] Philipp Frauenthaler, Marten Sigwart, Christof Spanring, and Stefan Schulte. 2020. Leveraging Blockchain Relays for Cross-Chain Token Transfers.

[21] Jorge Gracia and Eduardo Mena. 2019. Dealing with Semantic Heterogeneity Issues on the Web. *IEEE Internet Computing* (2019). https://doi.org/10.1109/MIC.2011.129

[22] A Gurtu and J Johny. 2019. Potential of blockchain technology in supply chain management: a literature review. *International Journal of Physical Distribution and Logistics Management* 49, 9 (2019), 881–900.

[23] Thomas Hardjono. 2021. Blockchain Gateways, Bridges and Delegated Hash-Locks. *arXiv* (8 2021). https://arxiv.org/pdf/2102.03933.pdf

[24] Karina Hauser, Helgi S. Sigurdsson, and Katherine M. Chudoba. 2011. EDSOA: An Event-Driven Service-Oriented Architecture Model For Enterprise Applications. *International Journal of Management & Information Systems (IJMIS)* 14, 3 (1 2011). https://doi.org/10.19030/ijmis.v14i3.839

[25] P Helo and Y Hao. 2019. Blockchains in operations and supply chains: A model and reference implementation. *Computers and Industrial Engineering* 136 (2019), 242–251.

[26] Wout Hofman. 2011. Supply Chain Risk Analysis with Linked Open Data. In *Formal Ontologies Meet Industry (FOMI) Proceedings of the Fifth International Workshop*.

[27] Huawei Huang, Wei Kong, Sicong Zhou, Zibin Zheng, and Song Guo. 2020. A Survey of State-of-the-Art on Blockchains: Theories, Modelings, and Tools. *arXiv* (8 2020). https://arxiv.org/pdf/2007.03520.pdf

[28] Hans-Henrik Hvolby, Kenn Steger-Jensen, Anders Bech, Sven Vestergaard, Carsten Svensson, and Mihai Neagoe. 2021. Information Exchange and Block Chains in Short Sea Maritime Supply Chains. *Procedia Computer Science* 181 (2021), 722–729. https://doi.org/10.1016/j.procs.2021.01.224

[29] Yassine Issaoui, Azeddine Khiat, Ayoub Bahnasse, and Hassan Ouajji. 2019. Smart logistics: Study of the application of blockchain technology. *Procedia Computer Science* 160 (2019). https://doi.org/10.1016/j.procs.2019.09.467

[30] Shafaq Naheed Khan, Faiza Loukil, Chirine Ghedira-Guegan, Elhadj Benkhelifa, and Anoud Bani-Hani. 2021. Blockchain smart contracts: Applications, challenges, and future trends. *Peer-to-Peer Networking and Applications* (8 2021). https://doi.org/10.1007/s12083-021-01127-0

[31] Ann Hibner Koblitz, Neal Koblitz, and Alfred Menezes. 2011. Elliptic curve cryptography: The serpentine course of a paradigm shift. *Journal of Number Theory* 131, 5 (8 2011). https://doi.org/10.1016/j.jnt.2009.01.006

[32] T Koens and E Poll. 2019. Assessing interoperability solutions for distributed ledgers. *Pervasive and Mobile Computing* 59 (2019), 101079. https://doi.org/10.1016/j.pmcj.2019.101079

[33] Czesław Kościelny, Mirosław Kurkowski, and Marian Srebrny. 2013. Foundations of Asymmetric Cryptography. https://doi.org/10.1007/978-3-642-41386-5{_}4

[34] Nir Kshetri. 2018. 1 Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management* 39 (2018), 80–89.

[35] Yixin Li, Bin Cao, Mugen Peng, Long Zhang, Lei Zhang, Daquan Feng, and Jihong Yu. 2020. Direct Acyclic Graph based Ledger for Internet of Things: Performance and Security Analysis. *arXiv* (8 2020). https://arxiv.org/pdf/1905.10925.pdf

[36] Yue Liu, Qinghua Lu, Hye-Young Paik, Xiwei Xu, Shiping Chen, and Liming Zhu. 2020. Design Pattern as a Service for Blockchain-Based Self-Sovereign Identity. *IEEE Software* 37, 5 (9 2020). https://doi.org/10.1109/MS.2020.2992783

[37] Jacob Lohmer and Rainer Lasch. 2020. Blockchain in operations management and manufacturing: Potential and barriers. *Computers & Industrial Engineering* 149 (2020), 106789. https://doi.org/10.1016/j.cie.2020.106789

[38] Mike Hearn and Richard Gendal Brown. 2019. *Corda: A distributed ledger*. https://www.r3.com/wp-content/uploads/2019/08/corda-technical-whitepaper-August-29-2019.pdf

[39] Arshad Muhammad, Abdullahi Arabo, Madjid Merabti, Qi Shi, and Bob Askwith. 2010. A Secure Gateway Service for Accessing Networked Appliances. In *2010 Fifth International Conference on Systems and Networks Communications*. IEEE. https://doi.org/10.1109/ICSNC.2010.35

[40] Andriy Nikolov, Victoria Uren, Enrico Motta, and Anne de Roeck. 2009. Overcoming Schema Heterogeneity between Linked Semantic Repositories to Improve Coreference Resolution. https://doi.org/10.1007/978-3-642-10871-6{_}23

[41] P. Zappalà, M. Belotti, M. Potop-Butucaru, and S. Secci. 2020. Game theoretical framework for analyzing Blockchains Robustness. *Cryptology ePrint Archive* (5 2020). https://eprint.iacr.org/2020/626.pdf

[42] A Panos, G Kapnissis, and H C Leligou. 2020. The Blockchain and DLTs in the Maritime Industry: Potential and Barriers. *European Journal of Electrical Engineering and Computer Science* 4, 5 (8 2020). https://doi.org/10.24018/ejece.2020.4.5.243

[43] Alex Preukschat and Drummond Reed. 2021. *Self-Sovereign Identity: Decentralized digital identity and verifiable credentials* (1 ed.). Manning.

[44] PROFILE. 2021. *Deliverable 2.4 — Possibilities of Blockchain Technologies for Trusted Data Sharing (Internal Report)*. Technical Report. https://www.profile-project.eu/

[45] Potchara Pruksasri, Jan van den Berg, Wout Hofman, and Yao-Hua Tan. 2014. Data concealing of supply chain transactions using the Distributed Trust Backbone. In *The 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014)*. IEEE. https://doi.org/10.1109/ICITST.2014.7038796

[46] R3 Ltd. 2021. Writing Oracle Services. In *Corda OS 4.8: Documentation and Training for Corda Developers and Operators*. https://docs.corda.net/docs/corda-os/4.8/oracles.html

[47] RAFAEL BELCHIOR, ANDRÉ VASCONCELOS, SÉRGIO GUERREIRO, and MIGUEL CORREI. 2021. A Survey on Blockchain Interoperability: Past, Present, and Future Trends. *arXiv* (3 2021).

[48] Gowri Sankar Ramachandran, Kwame-Lante Wright, Licheng Zheng, Pavas Navaney, Muhammad Naveed, Bhaskar Krishnamachari, and Jagjit Dhaliwal. 2019. Trinity: A Byzantine Fault-Tolerant Distributed Publish-Subscribe System with Immutable Blockchain-based Persistence. In *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE. https://doi.org/10.1109/BLOC.2019.8751388

[49] Drummond Reed, Manu Sporny, Dave Longley, Christopher Allen, Ryan Grant, and Markus Sabadello. 2021. *Decentralized Identifiers (DIDs): Core architecture, data model, and representations - W3C Proposed Recommendation Draft*. W3C. –undefined pages. https://www.w3.org/TR/2021/PR-did-core-20210803/

[50] Wei Ren, Xutao Wan, and Pengcheng Gan. 2021. A double-blockchain solution for agricultural sampled data security in Internet of Things network. *Future Generation Computer Systems* 117 (2021), 453–461. https://doi.org/10.1016/j.future.2020.12.007

[51] Leonid Reyzin and Sophia Yakoubov. 2016. Efficient Asynchronous Accumulators for Distributed PKI.

[52] Sara Ghaemi, Sara Rouhani, Rafael Belchior, Rui S. Cruz, Hamzeh Khazaei, and Petr Musilek. 2021. A Pub-Sub Architecture to Promote Blockchain Interoperability. *arXiv* (1 2021). https://arxiv.org/pdf/2101.12331v1.pdf

[53] Eder J Scheid, Timo Hegnauer, Bruno Rodrigues, and Burkhard Stiller. 2019. Bifröst: a Modular Blockchain Interoperability API. In *2019 IEEE 44th Conference on Local Computer Networks (LCN)*. IEEE. https://doi.org/10.1109/LCN44214.2019.8990860

[54] Lennard Segers, Jolien Ubacht, Yao-Hua Tan, and Boriana D Rukanova. 2019. The use of a blockchain-based smart import declaration to reduce the need for manual cross-validation by customs authorities. In *Proceedings of the 20th Annual International Conference on Digital Government Research*. ACM. https://doi.org/

10.1145/3325112.3325264

[55] Jie SONG, Pengyi ZHANG, Mohammed ALKUBATI, Yubin BAO, and Ge YU. 2021. Research advances on blockchain-as-a-service: architectures, applications and challenges. *Digital Communications and Networks* (8 2021). https://doi.org/10.1016/j.dcan.2021.02.001

[56] Manu Sporny, Grant Noble, Dave Longley, Daniel C Burnett, and Brent Zundel. 2019. *Verifiable Credentials Data Model: Expressing verifiable information on the Web - W3C Recommendation* . W3C. https://www.w3.org/TR/2019/REC-vc-data-model-20191119/

[57] John Strassner, Sung-Su Kim, Tom Pfeifer, and Sven van der Meer. 2010. An architecture for using metadata to manage ubiquitous communications and services: A position paper. In *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE. https://doi.org/10.1109/PERCOMW.2010.5470674

[58] Jianwen Su, Tevfik Bultan, Xiang Fu, and Xiangpeng Zhao. [n. d.]. Towards a Theory of Web Service Choreographies. https://doi.org/10.1007/978-3-540-79230-7{_}1

[59] Ali Sunyaev, Niclas Kannengießer, Roman Beck, Horst Treiblmaier, Mary Lacity, Johann Kranz, Gilbert Fridgen, Ulli Spankowski, and André Luckow. 2021. Token Economy. *Business & Information Systems Engineering* (8 2021). https://doi.org/10.1007/s12599-021-00684-1

[60] Yao-Hua Tan, Niels Björn-Andersen, Stefan Klein, and Boriana Rukanova (Eds.). 2011. *Accelerating Global Supply Chains with IT-Innovation*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-15669-4

[61] Paolo Tasca and Claudio Tessone. 2019. A Taxonomy of Blockchain Technologies: Principles of Identification and Classification. *Ledger* 4 (8 2019). https://doi.org/10.5195/ledger.2019.140

[62] S Tönnissen and F Teuteberg. 2020. Analysing the impact of blockchain-technology for operations and supply chain management: An explanatory model drawn from multiple case studies. *International Journal of Information Management* 52 (2020).

[63] United Nations. 1978. United Nations Convention on the Carriage of Goods by Sea (Hamburg Rules). *Treaty Series* 1695 (8 1978). https://treaties.un.org/doc/Treaties/1992/10/19921001%2004-36%20AM/Ch_XI_D_3.pdf

[64] Sélinde van Engelenburg, Boriana Rukanova, Wout Hofman, Jolien Ubacht, Yao-Hua Tan, and Marijn Janssen. 2020. Aligning Stakeholder Interests, Governance Requirements and Blockchain Design in Business and Government Information Sharing. In *Electronic Government - 19th IFIP WG 8.5 International Conference, EGOV 2020, Proceedings*. Springer, 197–209. https://doi.org/10.1007/978-3-030-57599-1{_}15

[65] Jyri Vilko, Paavo Ritala, and Jukka Hallikas. 2019. Risk management abilities in multimodal maritime supply chains: Visibility and control perspectives. *Accident Analysis & Prevention* 123 (2019), 469–481. https://doi.org/10.1016/j.aap.2016.11.010

[66] Wattana Viriyasitavat, Li Da Xu, Zhuming Bi, and Assadaporn Sapsomboon. 2020. Blockchain-based business process management (BPM) framework for service composition in industry 4.0. *Journal of Intelligent Manufacturing* 31, 7 (8 2020). https://doi.org/10.1007/s10845-018-1422-y

[67] Gang Wang. 2021. SoK: Exploring Blockchains Interoperability. *IACR Cryptology ePrint Archive* (8 2021). https://eprint.iacr.org/2021/537.pdf

[68] Qin Wang, Jiangshan Yu, Shiping Chen, and Yang Xiang. 2020. SoK: Diving into DAG-based Blockchain Systems. https://arxiv.org/pdf/2012.06128.pdf

[69] Ping Wah Wong and N Memon. 2001. Secret and public key image watermarking schemes for image authentication and ownership verification. *IEEE Transactions on Image Processing* 10, 10 (2001). https://doi.org/10.1109/83.951543

[70] Liyang Yu. 2014. The Building Block for the Semantic Web: RDF. https://doi.org/10.1007/978-3-662-43796-4{_}2

[71] Alexei Zamyatin, Mustafa Al-Bassam, Dionysis Zindros, Eleftherios Kokoris-Kogias, Pedro Moreno-Sanchez, Aggelos Kiayias, and William J Knottenbelt. 2019. SoK: Communication Across Distributed Ledgers. https://eprint.iacr.org/2019/1128.pdf

[72] Jian Zhang. 2020. Deploying Blockchain Technology in the Supply Chain. In *Computer Security Threats*. IntechOpen. https://doi.org/10.5772/intechopen.86530

[73] Zhiyi Zhang, Vishrant Vasavada, Xinyu Ma, and Lixia Zhang. 2019. DLedger: An IoT-Friendly Private Distributed Ledger System Based on DAG. *arXiv* (8 2019). https://arxiv.org/pdf/1902.09031.pdf

[74] Ziyuan Li, Huimei Wang, Jian Liu, and Ming Xian. 2021. Implementing a sidechain-based asynchronous DPKI. (2021). https://www.researchgate.net/publication/349097289_Implementing_a_sidechain-based_asynchronous_DPKI

Appendix B

Example Information Graphs

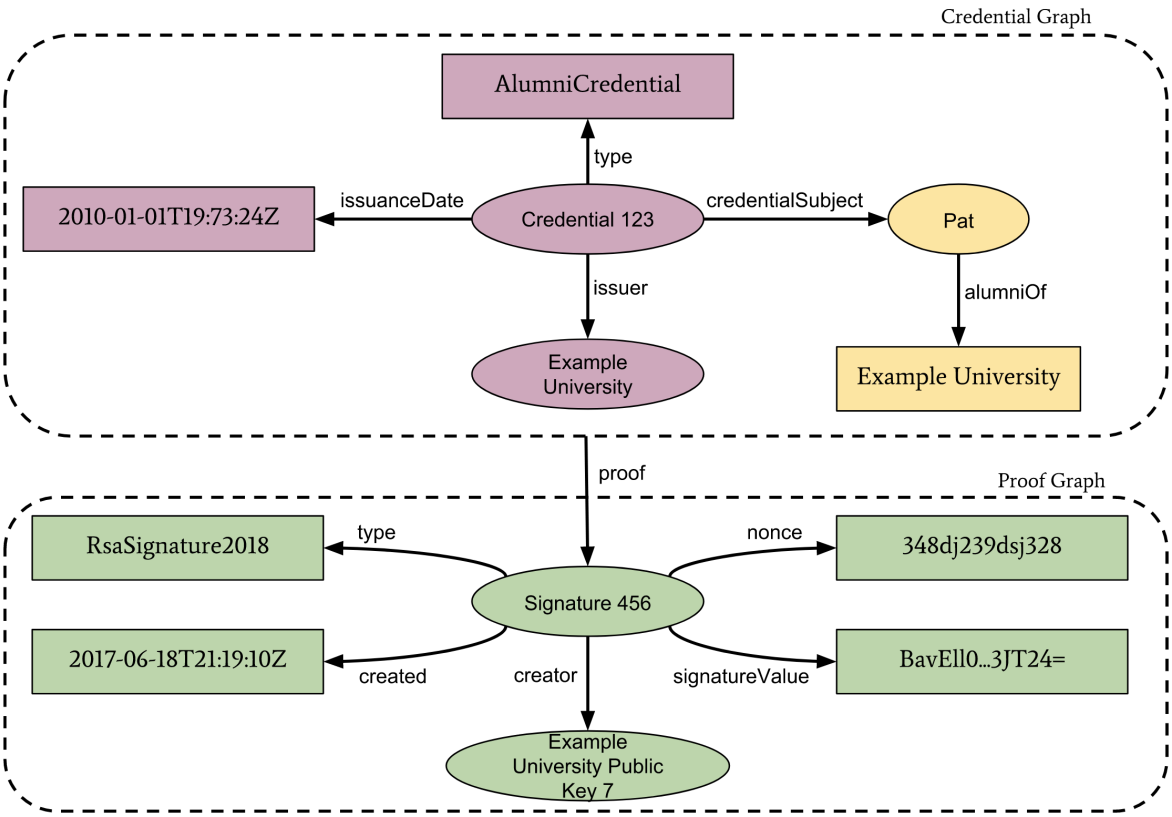


FIGURE B.1: Example of a verifiable credential information graph [154].

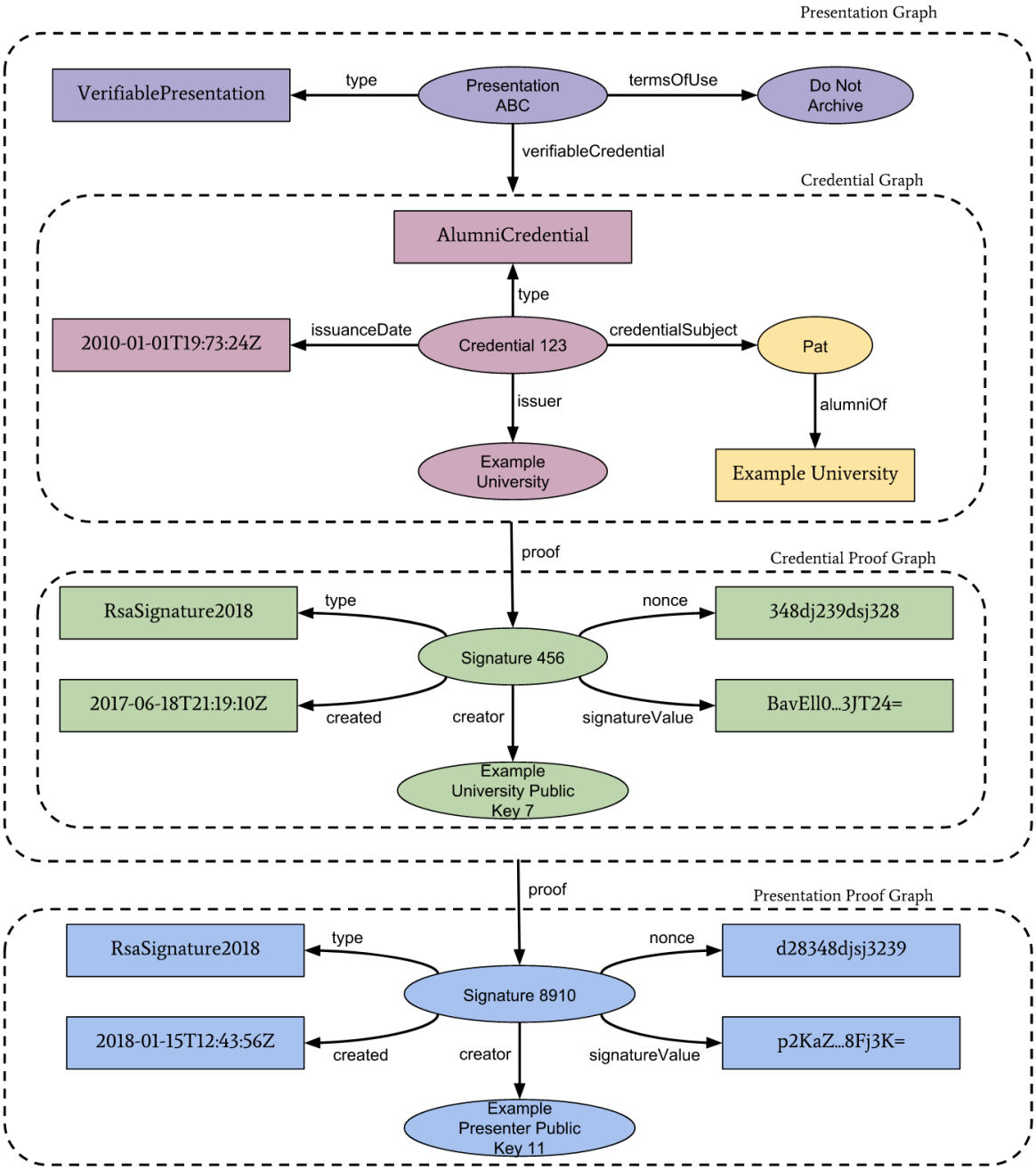


FIGURE B.2: Example of a verifiable presentation information graph [154].

Appendix C

Additional *DID* Specifications

DID Development Goals

TABLE C.1: Design goals of *DID*s, adapted [133].

goal	*description*
Decentralisation	Eliminate single point failures in identifier management and the registration of globally unique identifiers and public verification keys.
Control	Give entities, both human and non-human, the power to directly control their digital identifiers without the need to rely on *IDP*.
Privacy	Enable entities to control the privacy of their information, including minimal, selective, and progressive disclosure of attributes.
Security	Enable sufficient security for requesting parties to depend on *DID* documents for their required level of assurance.
Proof-based	Enable credential holders to provide cryptographic proof when interacting with other entities.
Discoverability	Make it possible for entities to discover *DID*s for other entities, to learn more about or interact with those entities.
Interoperability	Use interoperable standards so *DID* infrastructure can make use of existing tools and software libraries designed for interoperability.
Portability	Be system- and network-independent and enable entities to use digital identifiers with any system that supports *DID*'s.
Simplicity	Favor a reduced set of simple features to make the technology easier to understand, implement, and deploy.
Extensibility	Where possible, enable extensibility provided it does not greatly hinder interoperability, portability, or simplicity.

DID Method Requirements

TABLE C.2: *DID* method specification requirements [133].

#	*description*
1	A *DID* method specification MUST define how authorization is performed to execute all operations, including any necessary cryptographic processes.
2	A *DID* method specification MUST specify how a *DID* controller creates a *DID* and its associated *DID* document.
3	A *DID* method MUST specify how a *DID* resolver uses a *DID* to resolve a *DID* document, including how the *DID* resolver can verify the authenticity of the response.
4	A *DID* method MUST specify what constitutes an update to a *DID* document and how a *DID* controller can update a *DID* document or state that updates are not possible.
5	The *DID* method MUST specify how a *DID* controller can deactivate a *DID* or state that deactivation is not possible.

TABLE C.3: *DID* method security requirements, adapted [133].

#	*description*
1	The Security Considerations section MUST document the following forms of attack for the *DID* operations defined in the *DID* method specification: eavesdropping, replay, message insertion, deletion, modification, denial of service, amplification, and man-in-the-middle. Other known forms of attack SHOULD also be documented.
2	The Security Considerations section MUST discuss residual risks, such as the risks from compromise in a related protocol, incorrect implementation, or cipher after threat mitigation was deployed.
3	The Security Considerations section MUST provide integrity protection and update authentication for all operations required by Table C.2.
4	If authentication is involved, particularly user-host authentication, the security characteristics of the authentication method MUST be clearly documented.
5	The Security Considerations section MUST discuss the policy mechanism by which DIDs are proven to be uniquely assigned.
6	Method-specific endpoint authentication MUST be discussed. Where *DID* methods make use of DLTs with varying network topology, sometimes offered as light node or thin client implementations to reduce required computing resources, the security assumptions of the topology available to implementations of the *DID* method MUST be discussed.
7	If a protocol incorporates cryptographic protection mechanisms, the *DID* method specification MUST clearly indicate which portions of the data are protected and by what protections, and it SHOULD give an indication of the sorts of attacks to which the cryptographic protection is susceptible. Some examples are integrity only, confidentiality, and endpoint authentication.
8	Data which is to be held secret (keying material, random seeds, and so on) SHOULD be clearly labeled.
9	*DID* method specifications SHOULD explain and specify the implementation of signatures on *DID* documents, if applicable.
10	Where *DID* methods use peer-to-peer computing resources, such as with all known *DLT*s, the expected burdens of those resources SHOULD be discussed in relation to denial of service.
11	*DID* methods that introduce new authentication service types SHOULD consider the security requirements of the supported authentication protocol.

Detailed *DID* Architecture

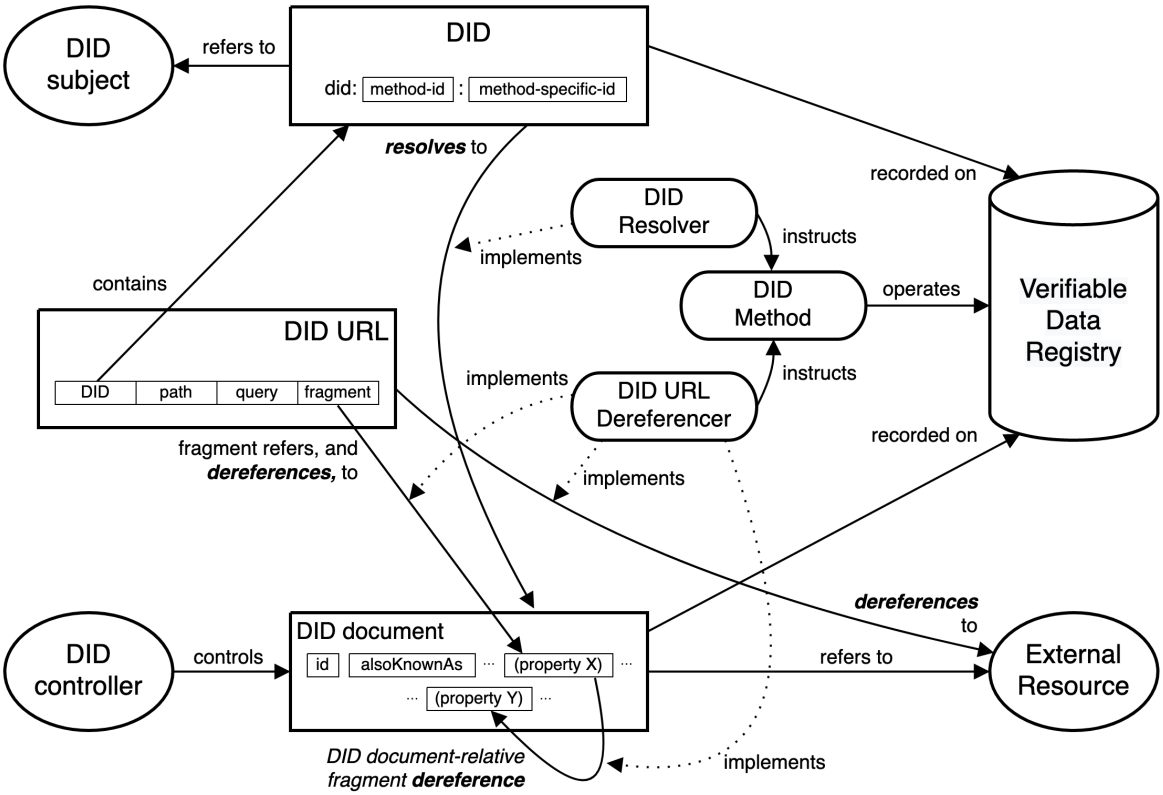
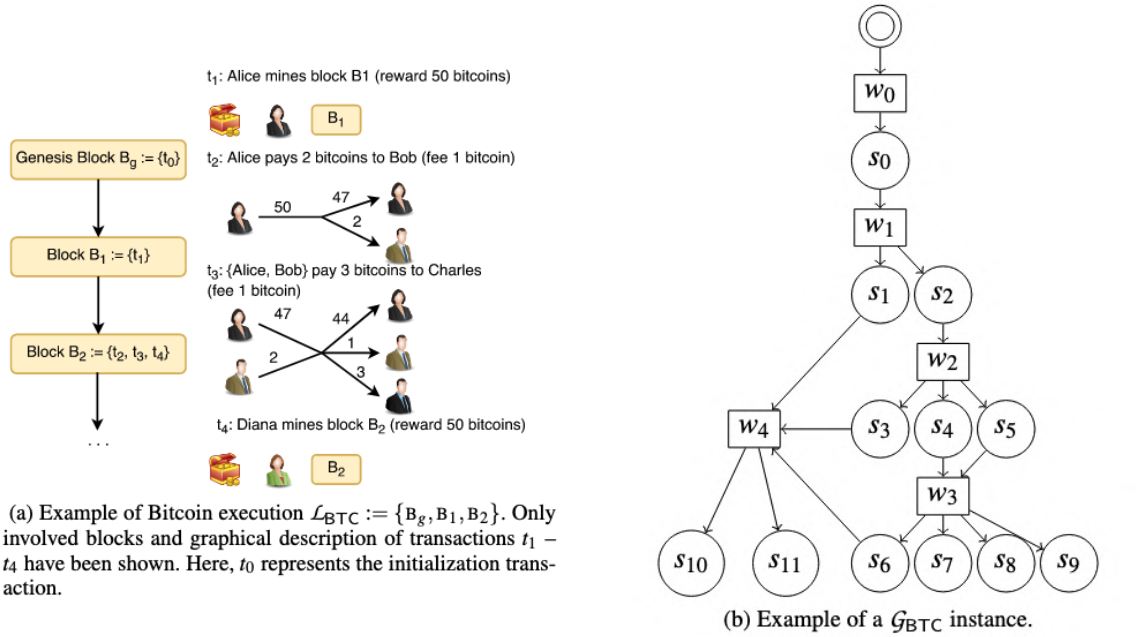


FIGURE C.1: Detailed overview of *DID* architecture [133].

Appendix D

Applied TDAG Examples



(a) Example of Bitcoin execution $\mathcal{L}_{\mathsf{BTC}} := \{B_g, B_1, B_2\}$. Only involved blocks and graphical description of transactions $t_1 - t_4$ have been shown. Here, t_0 represents the initialization transaction.

(b) Example of a $\mathcal{G}_{\mathsf{BTC}}$ instance.

FIGURE D.1: Bitcoin transactions represented using TDAG [25].

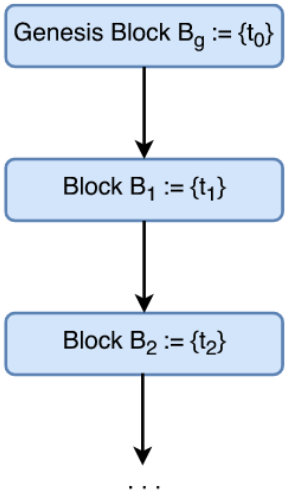
ALGORITHM 1: Function f_1

PUTSTATE ("a", 0);
PUTSTATE ("b", 5);
PUTSTATE ("c", 3);

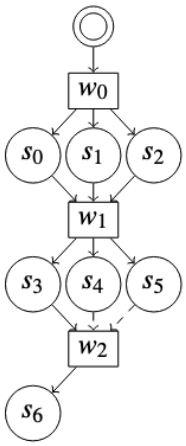
ALGORITHM 2: Function f_2

$a \leftarrow$ GETSTATE ("a");
$b \leftarrow$ GETSTATE ("b");
$c \leftarrow$ GETSTATE ("c");
PUTSTATE ("a", $a + b + c$);

(a) Set of GETSTATE and PUTSTATE operations for each function defined in this example.



(b) Example HLF execution $\mathcal{L}_{\mathsf{HLF}} := \{B_g, B_1, B_2\}$. Here, t_1 invokes f_1 and t_2 invokes f_2.



(c) Example of $\mathcal{G}_{\mathsf{HLF}}$ instance.

FIGURE D.2: Hyperledger Fabric transactions represented using TDAG [25].

Appendix E

Consensus in the CAPER Protocol

Local Application Consensus

The performance and interoperability limitations of conventional blockchains are driven by its batched block structure, as well as consensus algorithms designed to operate on a single chain where forking is not permitted (such as proof of work or proof of stake) [27]. An example is shown in Figure E.1, which consists on producing parallel chains of nodes that will be eventually rejected by the consensus protocol for not being part of the longest chain [79]. Uncontrolled forking can produce irreparable disagreements on current blockchain states, thus creating incompatible database instances that compete with each other [146].

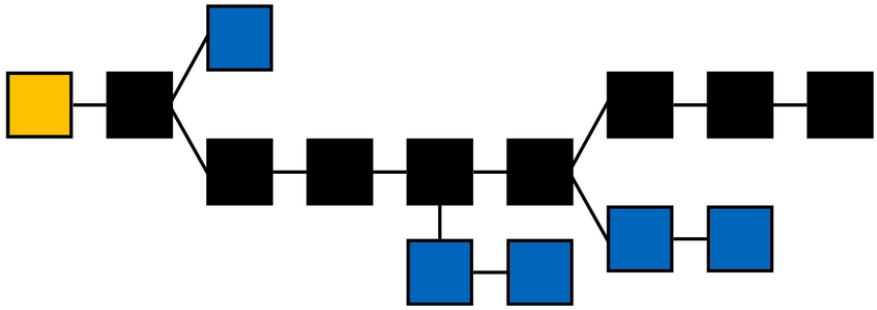


FIGURE E.1: Example of blockchain ledger fork (fork nodes in blue) [79].

Along the forking challenge comes the possibility of double-spending attacks [79], which is the risk of a digital asset being spent (*consumed*) more than once simultaneously. Interestingly, a pure *DAG* network is able to counter this challenge by adopting forking as the essence of the network topology. As a result, the main advantage is the ability to append more than one transaction (node) in parallel, which increases the potential confirmation rate and transaction throughout of the network [27].

The proposed system uses pluggable local consensus [9], meaning that the set of nodes of each application can choose a crash fault-tolerant protocol (*CFT*), practical (*PBFT*) byzantine fault-tolerant protocol, or even a trusted node in charge of ordering transactions. This is particularly useful when the goal is to coordinate applications that represent interactions between different blockchain consensus protocols or different third party transaction ordering services, such as *Raft* in *Hyperledger Fabric* [191] or *DAG*-based asynchronous byzantine fault-tolerant (*aBFT*) consensus services, including *Hashgraph* [11] and *Tangle* [125].

Global Consensus

The consensus for cross-application transactions can be achieved in three different ways. The first one is relaying on independent nodes not taking part in any application, called *orderers*. The *orderers* will witness these transactions from an impartial perspective and provide a global consensus. In this case, the consensus protocol is once again pluggable, and

any type of fault-tolerant protocol can be implemented. However, the implementation and control of these nodes might be difficult from a governance perspective.

The second option is a hierarchical global consensus, in which each application must first achieve consensus to provide a vote in each global consensus phase. In this case, CAPER uses an *aBFT* protocol. This process is the instance shown previously in **??**, which is expensive, as all local consensus protocols must run at a specific pace. The need for *aBFT* in this global consensus approach comes from possible liveliness imbalances for messages distributed between applications due to differences in their local consensus protocols.

There is a third option, in which the ordering of the local and cross-application transactions is merged. This requires ensuring that the majority of nodes on each application agree on the order of transactions. However, the number of nodes and consensus algorithm can vary per application, so it is required to take into account how *local majority* threshold is measured by each application. This approach entails the highest number of implementation obstacles in terms of technical compatibility, performance and governance.

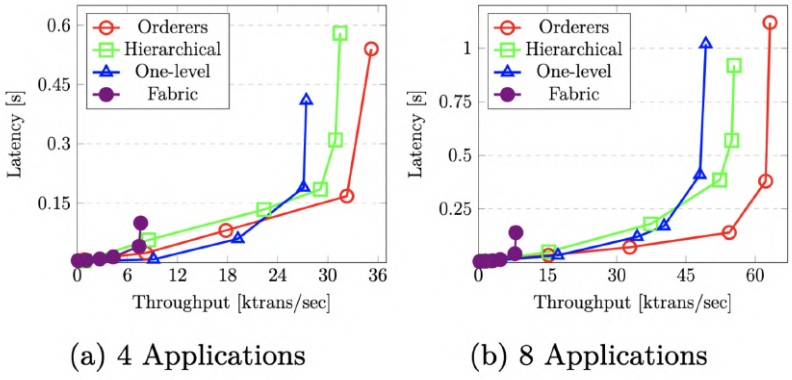


FIGURE E.2: Consensus performance for: (a) 4 applications and (b) 8 applications [9].

The transaction throughout (thousand transactions per second) and latency (time needed to add a record on the registry) for the three proposed consensus approaches and an equivalent application on *Hyperledger Fabric* was simulated by *Aimiri et al.* [9] (Figure E.2). The results indicate that *Hyperledger Fabric* tends to outperform CAPER in terms of latency. CAPER shows the possibility to increase the system throughput significantly by increasing the number of applications, while increasing the number of applications (channels) in *Hyperledger Fabric* does not enhance performance considerably. From a design perspective this means that a trade-off must be made between desired throughout and latency requirements based on an estimated number of applications and percentage of cross-application transactions, which falls outside the research scope at the current stage.

There are reasons to consider the use of independent nodes to order the transactions. Experiments indicate that *DAG*-based consensus services can achieve quantum immunity [27, 125] and provide proofs of asynchronous byzantine fault tolerance [27, 11]. There is skepticism in the academic community and more research to validate these claims is needed [59]. Nonetheless, it is widely accepted that the technology will at least perform better than traditional consensus algorithms against quantum attacks [189].