Discussion paper:

A new level of data synchronization for efficient logistics data exchange

Authors: Christophe Maurin (GEODIS), Eric Ballot (Mines Paris-Tech), Benoit Tiers (GEODIS), Tomas Ambra (ALICE) & Fernando Liesa (ALICE)

Contact: info@etp-alice.eu ref : Shared Data Language

Motivation

Paper-based documents, fragmented information and various system silos contribute to inefficiencies in freight transport and logistics. These inefficiencies relate to challenges to proper planning that turn out in delays, uncertainties, higher administrative costs as well as unnecessary emissions. Many initiatives have been launched to tackle such problems from both governmental and industry perspectives. The former relates to the Digital Transport and Logistics Forum[1] (DTLF) which supports and contributes to the development and implementation of the new eFTI regulation[2] and corridors information systems. The Common European Mobility Data Space[3] (EMDS) initiative aims to facilitate data access, pooling and sharing for more efficient, safe, sustainable, and resilient transport. Additionally, and transversal to all industries the Data Act[4] aims to address the challenges and unleash the opportunities presented by data in the European Union, emphasizing fair access and user rights, while ensuring the protection of personal data. The industry perspective covers companies and organizations who build data sharing solutions within their own processes and systems for their daily operations. Based on top-down government initiatives, it is crucial to understand how companies can leverage their existing assets while tangibly aligning with DTLF and EMDS principles when connecting with various parties, including new ones. All at minimal cost and to benefit from the implementation of portals requiring EU governments to accept paperless information sharing.

The main goal of this discussion paper is to propose a solution from the industry perspective to support the creation of more efficient logistics processes and facilitate the exchange of data while improving its accuracy between business partners and with governments. This document offers a framework for data exchange during a business partnership while retaining control and ownership of the data during the whole process. This enables parties involved in logistics chains to share relevant data related to operational processes and their execution. It is based on a fully decentralized approach and relies on two main contributions: an abstract language used as a pivotal semantic level and a suite of protocols to realize data exchange in a secure and timely manner. The whole is illustrated on a use case.

ALICE aims to raise awareness on the relevant government initiatives towards companies and industry and gather input from industry raising consensus on frameworks and solutions to lever the potential value from digitalization and data exchange. ALICE welcomes additional use cases from industry and in practice that could support a bottom-up approach.

[1] https://transport.ec.europa.eu/transport-themes/digital-transport-and-logistics-forum-dtlf_en
[2] https://eur-lex.europa.eu/EN/legal-content/summary/electronic-freight-transport-information.html
[3] https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13566-Transport-data-creating-a-common-European-mobility-data-space-communication-_en
[4] https://digital-strategy.ec.europa.eu/en/policies/data-act

What does data sharing mean?

Data sharing is mentioned in many contexts and could be seen as a threat to data value and business confidentiality. Nevertheless, the flow of data between companies involved in logistic activities is a must and is core in current logistics business processes to ensure seamless operations in the supply chains. In this discussion paper, data sharing should be understood as a controlled process by the data owner, to serve its processes with partners, such as customers, suppliers, authorities… This process could be open to different degrees, but it is always supported by a legal framework (a contract, a law…) therefore with limitations and based on the will of the parties involved. Data sharing has two different angles that may be intertwined:

- Data sharing with the purpose of planning and executing collaborative logistics processes seamlessly and efficiently providing visibility in the supply and value chain.
- Data sharing with the purpose of providing access to information for regulatory verifications and or giving access to data sets for monetization.

Data sharing for logistics activities

Logistics operations are by nature distributed between a multitude of areas of activity, actors, locations, which require up-to-date data as they evolve over time for all parties, to synchronize planning, monitor the execution, including the needs for replanning, and controls by regulators. As shown in the following figure, in order to execute their processes and address their challenges, stakeholders from various relevant domains need to exchange data and this need will just grow in the forthcoming years.

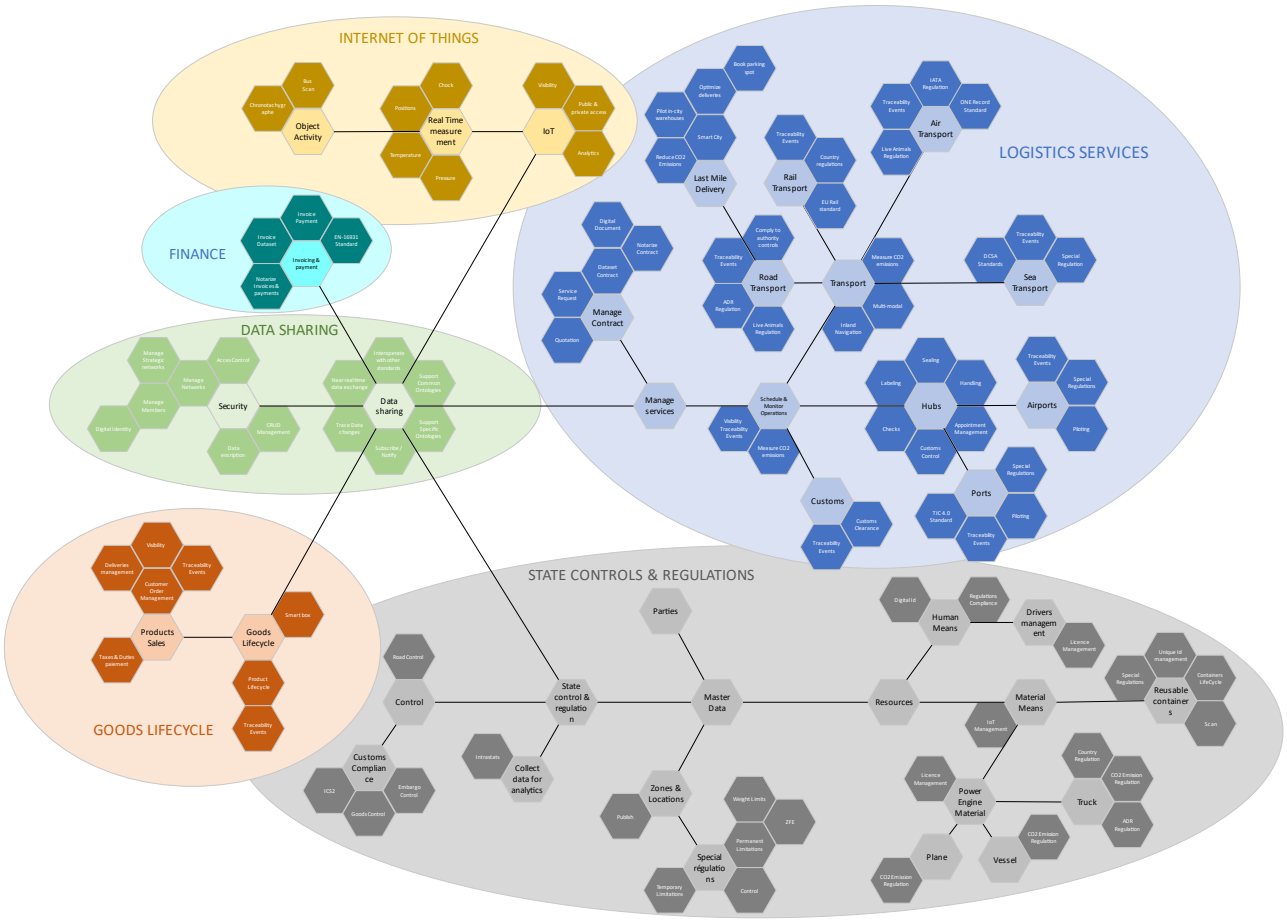

Figure 1: data exchange between logistics stakeholders

Given the above setting, there are several important barriers to overcome.

- **Logistics is executed by a large majority of small and very small players (SMEs**), except rail, air and sea modes, mainly represented by a limited number of companies. As a result, there is a need to connect many companies to many others in a dynamic way.
- The **fragmentation of the market** by transportation modes, service types, IT systems etc., **implies a huge diversity in data and multiple standards**. Most of the logistics providers are serving their customers in a multimodal environment based on capacity availability. When it comes to data sharing, it drives huge complexity and inefficiencies in processes. Despite all efforts done, monolithic shippers, freight forwarders, logistics service providers or authorities platforms prevail[5] creating burden and complexity to engage in a digital coordination with affordable cost and ease, in particular for SMEs.
- The **capability of most operators to invest in IT systems is very limited**, despite their need to own, control and process their data. The transaction logic found in EDI and other tools is not relevant anymore to represent the dynamic environment, yet dominant, as well as to lever opportunities brought forward by technologies such as AI, DT, 5G, IoT and others.
- **Small logistics providers do not possess the capacity and capability to adapt to each single customer digital requirements** and each transport mode standard to improve their process synchronization; this proves to be the bottle neck of digitalization. It is therefore necessary to develop and adapt universal plug & play solutions based on open-source components for players of all sizes.
- **Digital logistics platforms and dataspaces are their infancy and emerging with several solutions.** If not enough attention is paid to crucial points, the risk is to continue creating silos, monolithic platforms and data spaces that are not interconnected with each other.

To get around these barriers, a commonly accepted approach consists of addressing the subject of data sharing by use cases. While this helps meet immediate and one-to-one needs, one can fear counterproductive effects. The benefits of digitalization risk being wiped out by the multiplicity and complexity of digital solutions. For example, a road driver will need to have an application to show digital transport contracts to the authorities, another one to declare the information necessary to calculate greenhouse gas emissions and many others for customs, to report to its customers that the goods have been delivered.

In the same way, players in logistics fields, such as air, maritime or even rail transport, organize themselves within their sector to define their own data exchange standards. This is facilitated by the limited number of representatives in their sector. If this can solve questions of data sharing within each of these sectors, this approach does not contribute to an overall improvement of the logistics chain neither facilitate eco-responsible modal shift. Finally, most actors carrying out road transport or handling, do not have a standardized solution for sharing the data that they process.

To solve this rising and known issue, platforms appear as a solution: a data hub. It is proposed at country level for legal data (e-cmr) but also by other EU initiative like E-federated or players on the market. This one connection data spot, except for legal data, rises also questions among players in the sector: dependency, value sharing, anti-trust regulations, too few or too many platforms. All these aspects are hampering the digitalization of the logistics sector.

[5] https://www.etp-logistics.eu/horizontal-collaboration-by-collaborative-platforms-workshop/

One principle and two proposals to synchronize logistics' data and improve processes digitalization

Logistics data should be shared to any stakeholder related to it: economic actors, authorities, consignees regardless of their size and under the full control of its owner in a peer-to-peer manner. This data sharing environment is common and should be open source to be adapted and improved over time and not the property of a particular provider. Providers will operate in this digital ecosystem providing value added services to participants.

Proposal #1. A common high-level language

To avoid the well-known problem of interconnection needs from many systems to any of the others, a common language is required. The proposition is to build a high-level conceptual model (a shared language) that will describe all logistics operations starting with highest level of genericity and going into the details when needed. The knowledge description carried out by such tools ease the matching of fields from any application to any other. An application developed in accordance with the common language specification has a data model consistency regardless of the other provider as long as it also follows the same open specifications. Therefore, all customers or suppliers could appear integrated in the same tool. In this framework the integration of data coming from another party is as easy as granting an access. At lowest level of data, it could refer to any kind of standards for a shipment, a location and so on. The investment made in standards like IATA, DSCA, TIC4.0, UN/Cefact, GS1 (RAIL)… is then preserved and remains key elements to build this high-level language.

While information exchange is crucial to ensure discoverability and identification of underutilized assets, infrastructures and supply chain events, it is even more important to implement a standardized approach to facilitate data and information exchange, enabling systemic process coordination between players. In that regard, open and universally adopted standards set-up a pillar to enable seamless and end-to-end (digital) coordination. This is the backbone of the Physical Internet which comprises of key building blocks that connect fragmented solutions and system silos via open standards and protocols. One of such building blocks is interoperability between various stakeholders within the freight transport & logistics ecosystem, but also harmonization of existing standardization groups listed above. Such a universal language could be seen as a utopia until it is developed and tested successfully.

Proposal #2. Adoption of a set of communication protocols

The implementation of software compliant to a common language is key but not sufficient if we still rely on data exchange protocols designed in the seventies. Modern secured protocols already described by the W3C, are key enablers to share standardized data in real-time to synchronize business processes.

Data spaces need the same type of technology to grant access to their functionalities. Therefore, there is a need to standardize access to all parties. This goes through a selection of interoperable protocols for a seamless integration of the different applications needed for any logisticians.

How will an ontology help to share data in logistics activities?

An ontology is a tool that is built on a high level of abstraction like an object in object-oriented coding. With the use of an ontology a logistics service can be defined with universal attributes but can also be refined according to each sector needs as an example. Within the common logistics language, the air sector can be specified in compliance with the International Air Transport Association - IATA's specifications, Digital Container Shipping Association (DCSA) for maritime sector, Terminal Industry Committee (TIC4.0) for terminals and so on. At the same time the service is still accessible with properties for LSP coming from other sectors. Like carbon emissions it is possible to report them with components of different levels of accuracy. The ontology mechanisms make possible the link between services coming from different providers even if they don't have the same level of description, thanks to its generic approach. This approach will drastically easy all computations related to more transversal topics such as carbon emissions or dangerous goods. Instead of specific add-ons and multiples interfaces, they are considered natively if the topic is covered for all sectors.

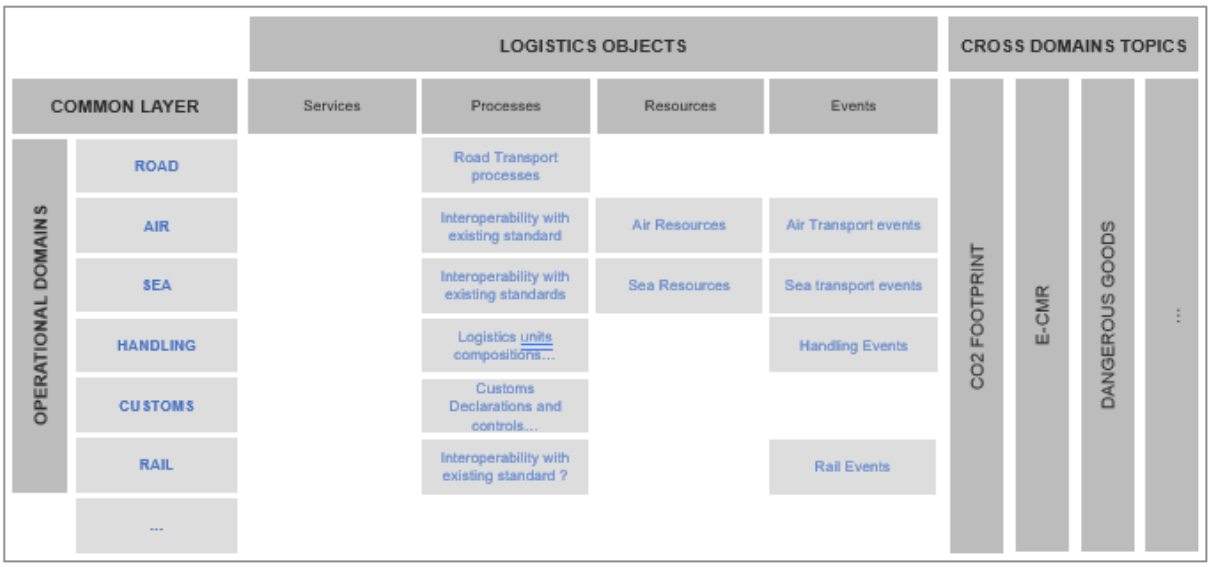

Figure 2: Creating ontologies from reusable commonality to specializations.

Lastly, such a generic approach is also a way to support interaction with other domains such as urban logistics. In a smart city, delivery spaces will be accessible online. They are city's resources that logisticians should be ready to connect with and without customization of each app for each city. This

kind of solution was already successfully tested in an experimentation and is only an illustration of the power of the proposed approach.

Transition phase

From a logistician point of view, it implies as depicted in the next figure to build a single interface to access the different services. The grey block represents the investment required by any node to access all shared services compliant with the two above mentioned principles.

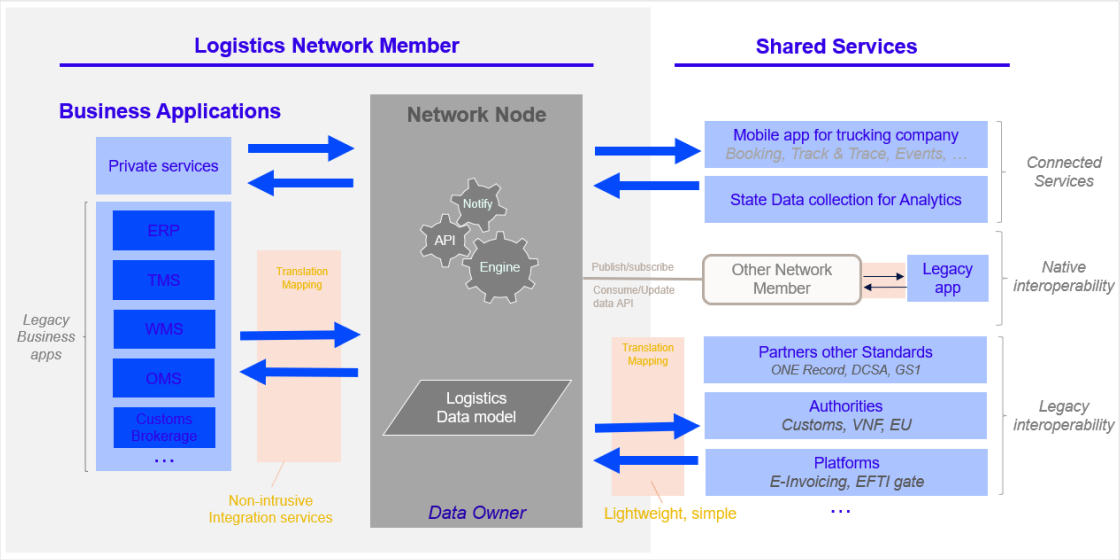

Figure 3: Example of architecture design for a network member

As shown, it can be started regardless of the legacy systems in place. Of course, the better the business applications, the better it is. The proof of concept below validates the architecture above.

In such an environment, the competition applies for business applications. But even a small carrier with an application developed in accordance with the network node specifications could benefit at its level of the approach's potential. To ensure that the network node that makes the collaborative space a reality, designed to the needs of logisticians, we recommend its specification by logisticians in collaboration with logistics stakeholders.

A proof of concept: shared data language by GEODIS

The GEODIS company, a French global logistics specialist, launched a working group in 2021 on the topic of the universal language of logistics, coming from common work with the International

Transport Association (IATA), engaged in the redesign of its data sharing standards for the Air Transport domain.

The objective was to carry out a proof of concept for a universal logistics language applicable to all stakeholders.

The complexity of logistics often comes from the diversity of operations performed, the very regular use of subcontracting to solve the difficult balance between resource needs and their availability in the right place at the right time.

This is compounded by the risks of information loss due to the low level of digitalization among actors. As a result, delayed knowledge of information can lead to urgent decision-making, opposite to the goal of an efficient and eco-friendly logistics.

The scheme below shows a regular example of intercontinental transport chain.

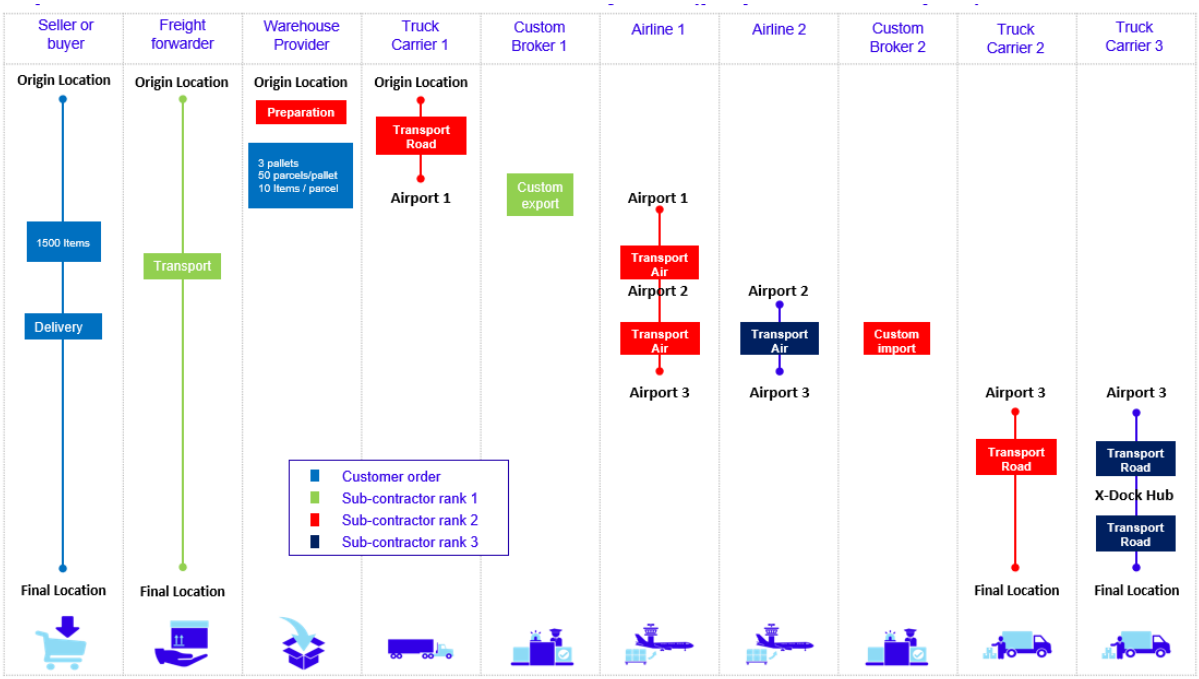

Figure 4: example of international transport flow

The solution proposed in the demonstration is based on two foundations:

- A data model organized in layers, initially generic and then specialized for domains (road transport, customs...) or subjects (dangerous goods, perishable goods, greenhouse gas emissions...)
- A secured communication protocol allowing data sharing respecting internet standards for process synchronization.

The data model

To be as generic as possible, the model abstracts from the specificities of logistic domains by considering that a logistics chain is an aggregate of services performed by providers on behalf of their

clients. Process activities are scheduled and executed in each service, and an activity can itself be delegated to another provider, triggering subcontracting mechanisms.

The generic layer is a service and a process driven model, and it is the same end-to-end.

In its generic layer, the model also describes goods orders, goods packaging, and resources used (containers, engines, human means...). It also allows events triggered by operational processes, documents, and references from legacy IT Systems to be added to any object.

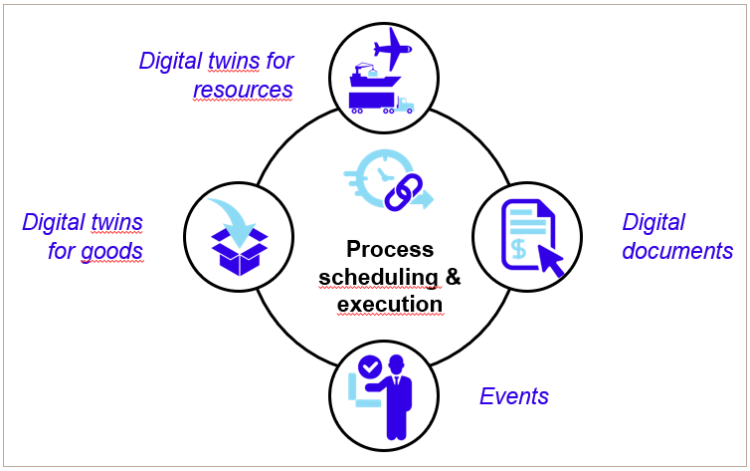

Figure 5: service and process driven data model to apply to all logistics providers.

The model doesn't try to compromise on all the existing data models. By the abstraction exercise, new objects and new vocabulary are created to cover the gaps and make the model universal.

The design of the model does not define the processes themself and business rules to be applied to each domain or even to each client-supplier relationship. It allows them to make them consistent in a universal language to facilitate sharing and holistic views.

The protocol for data exchange.

Having a generic data model is not enough to streamline the flow of logistics data.

Stakeholders need to be connected in a standardized network. It is mandatory to define a protocol for distributed network operations, secured and allowing each data owner to retain sovereignty without having to transfer them to centralized proprietary platforms.

Thanks to proven internet technologies, the protocol is based on state-of-the-art WEB mechanisms defined by the W3C, namely:

- Web APIs accessible via URIs.
- Security rules of verifiable credentials to guarantee machine-to-machine exchanges.
- Data management in wallets coupled with verifiable credentials for secure data sharing during interactions between physical actors (for example, checking the digital identity of a carrier when arriving to the shipper's location)

This protocol simplifies sharing of logistics data while guaranteeing their integrity by authorized access to the source.

In addition to the application of W3C standards, the protocol integrates mechanisms to bring value to the stakeholders:

- Keeping track of changes on shared objects (audit trail)
- Access rights management by the data owner to ensure sovereignty.
- Real time notifications to facilitate processes.

The solution developed must be non-intrusive to anticipate a progressive implementation:

- either by connecting compatible actors each with their own instance
- or by serving as a vector of interoperability with existing standards
- or by allowing small carriers not connected to the network to still have plug & play tools to feed their clients' systems, or even the authorities.

By complying to the data model and the protocol rules, each member decides how to expose the data he owns to his counterparts via his Logistics network node.

Implementation example:

The language was used to demonstrate how to synchronize, in real time, processes between parties and systems. The show case was an End To End transport flow driven by a customer to deliver the goods from a warehouse to all Europe.

One part was about showing how to use the language to synchronize the processes between different logistics players on the same logistics network.

A second part was about showing how to provide a mobile app to a small carrier connected to multiple customers, thanks to a standardized API language.

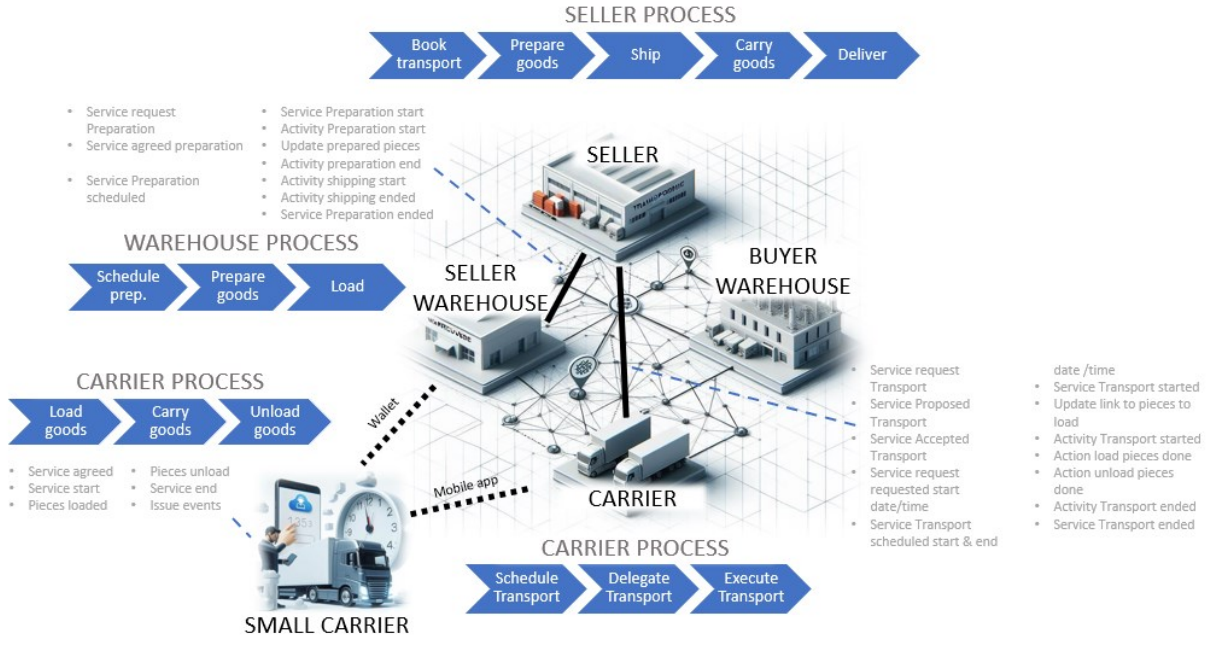

Figure 6: synchronizing real time processes using the same logistics language

This proof of concept serves to pragmatically illustrate the importance of adopting a common language by exploring the appropriate levels of abstraction in exchange models, rather than justifying their differences. Ultimately, regardless of sectors or stakeholders involved, digitalization primarily

impacts intra and inter-company processes, which must be addressed for efficient and autonomous logistics."

How to engage on the topic of shared data language? (proposal)

European Union has defined the concept of data spaces as an answer to support GDPR, Data Act and more recently Data Market Act. Current data spaces developments are facing difficulties to implement interoperability inside and between data spaces. One main reason is the “Use case” approach which may not be the most relevant to build structural components.

Additionally, there is no data space for logistics stakeholders who are facing governance limitations to address the topic of interoperability.

Isn’t it an opportunity to create a Logistics Data space governed through the rules of the EDIC and IDSA having a shared logistics language to structurally manage interoperability at a semantic level, using common and secured access protocols for stakeholders and to be reused by the other data spaces as soon as they need logistics services?

The benefits could be twofold.

1. A cheaper and easier access to the functionalities needed to operate as a modern logistician.
2. An opportunity for software developers to provide better and scalable applications with function that will take advantage of data integration: direct emissions reporting, appointment booking integrated in vehicle route, etc.

Why logisticians should being involved?

The developments underway will structure how data will be exchanged in the next decades or so like EDI. If major logisticians are not engaged in the specifications of the future data exchange mechanisms, most likely too many solutions will co-exist replicating current sector fragmentation in the digital age or will be imported with a loss of control and major cost for the logistic sector. There is a unique time window in the forthcoming year to define standards by and for logisticians.

For whom this initiative is relevant and for whom is not

Those organizations that believe a common language, interoperability and focus on commonalities is critical for future competitiveness of their companies and logistics ecosystem and companies that wish to break silos in data sharing governance, frameworks and structure.

Monolithic platforms that perceive competitive advantage in operating as such.

Next steps and how to be involved in the discussion

We are looking for additional use cases and implementations of open and universal data languages in the context of logistics as well as any other private initiative related to this discussion paper.

We are seeking for transportation, logistics and shipper companies to reflect on this discussion paper and eventually engage in a broader discussion to define adoption pathways.

If you would like to know more and get engaged, please contact ALICE at info@etp-alice.eu ref: Shared Data Language.