Towards a Reference Enterprise Architecture to enforce Digital Sovereignty in International Data Spaces

Danniar Reza Firdausy
Faculty of Behavioural, Management and Social Sciences
University of Twente
Enschede, The Netherlands
d.r.firdausy@utwente.nl

Patrício de Alencar Silva
Faculty of Electrical Engineering, Mathematics & Computer Science
University of Twente
Enschede, The Netherlands
p.dealencarsilva@utwente.nl

Marten van Sinderen
Faculty of Electrical Engineering, Mathematics & Computer Science
University of Twente
Enschede, The Netherlands
m.j.vansinderen@utwente.nl

Maria-Eugenia Iacob
Faculty of Behavioural, Management and Social Sciences
University of Twente
Enschede, The Netherlands
m.e.iacob@utwente.nl

***Abstract*— An International Data Space (IDS) aims to facilitate sovereign data sharing in business ecosystems. The GAIA-X project and the International Data Spaces Association (IDSA) lead initial European efforts to create such information systems. These institutions' high-level business rules and architectural guidelines are essential to attract companies interested in joining the IDS vision. However, companies may interpret these guidelines differently and derive implementations that have interoperability issues. This paper addresses this issue by reconciling data sovereignty and Enterprise Interoperability requirements into a Reference Enterprise Architecture for IDS. It aims to help companies create instantiations or specializations of organizational and software components to meet specific business cases' needs whil e preserving essential IDS principles. Representatives of two Enterprise Integration software companies interested in exploring the IDS vision helped refine the architecture through Technical-Action Research. An expert panel of representatives from the Dutch Logistics sector evaluated the architecture regarding its potential acceptance by small and medium enterprises (SMEs).**

Keywords—Enterprise Architecture, Enterprise Interoperability, Digital Sovereignty, International Data Spaces

I. Introduction

An International Data Space (IDS) comprehends a trusted environment where companies can share operational data to optimize core competencies [1]. For instance, in the Logistics sector, access to near real-time information about routes, transport orders, or vehicles can help enterprises optimize service delivery [2]. The GAIA-X project and the International Data Spaces Association (IDSA) are the two main initiatives setting up the European IDS vision. While the former provides technical specifications for deploying security-enforcing communication infrastructure [3], the latter recommends flexible business roles and application development guidelines that companies should adopt upon joining an IDS ecosystem [3, 4].

On the one hand, those guidelines seem flexible enough to attract as many companies as possible to the IDS vision. On the other hand, companies may interpret them distinctively: they might either reject the IDS vision or extend its technical specifications to cope with particular requirements demanded by emergent business cases. For instance, *data sovereignty* is an essential requirement in IDS, granting enterprises primary control on which data to share with whom and how [5]. Data sovereignty can also presume *digital sovereignty*, which translates into an enterprise's autonomy to adapt or create organizational assets and software components to share sensitive data [6]. However, enterprises striving for digital sovereignty may create IDS architecture implementations that are challenging to interoperate. Such phenomenon may limit the achievement of cross-border and cross-sector services delivery as envisaged by the European Interoperability Framework (EIF), which comprehends interoperability as the organizations' ability to interact with each other to realize shared goals through information sharing and business processes integration by employing data exchange between their ICT systems [7]. As a result, digital sovereignty and Enterprise Interoperability (EI) can become conflicting requirements in typical IDS implementations, which leads to the main research question addressed in this paper:

How could companies meet digital sovereignty and Enterprise Interoperability requirements while implementing an IDS ecosystem infrastructure?

The motivation to treat this problem is two-fold. First, there is an abstraction gap between the high-level IDSA architectural guidelines and the specific demands of sovereign data-sharing business cases. Those declarative guidelines are helpful but could be faster and easier to use if communicated in a standard architecture language. Second, a more concrete architectural model could help companies decide what components to instantiate or specialize in when designing and deploying private IDS architecture implementations. Such a model could facilitate communication among companies willing to interoperate in IDS ecosystems.

This paper addresses this problem by proposing a Reference Enterprise Architecture for IDS ecosystems to make the IDS Reference Architecture Model (RAM) easier to accept and use, especially by small and medium enterprises (SMEs). The problem-solving approach adopted in this work is Design Science [8]. The planned research methodology combined different research methods. First, a literature review clarified how digital sovereignty in IDS could translate into more specific requirements. Second, Technical-Action Research (TAR) helped design and refine an Enterprise Architecture to enforce digital sovereignty in IDS. Two software companies experienced in Enterprise Integration for

2378-1971/22/$31.00 ©2022 IEEE
DOI 10.1109/CBI54897.2022.00020

the Dutch Logistics sector and interested in joining the IDS vision cooperated with this study. Finally, an expert consultation panel of representatives from the Dutch Logistics sector evaluated the architecture based on the six criteria of technology acceptance for SMEs proposed by Bernaert, et al. [9].

The rest of this paper elaborates as follows. The next section brings a theoretical background on digital sovereignty in IDS and translates this generic concept into more specific software requirements. **Section III** describes the proposal of a Reference Enterprise Architecture to enforce digital sovereignty in IDS, structured in three viewpoints: *data and metadata exchange*, *certification and evaluation*, and *infrastructure*. **Section IV** discusses a preliminary architecture assessment done with an expert consultation panel. **Section V** outlines the contribution of this work regarding the closest related companion research. Lastly, a summary of this research's main achievements, limitations, and future steps closes this paper.

II. Defining Digital Sovereignty in IDS

The literature in International Data Spaces often refers to *digital sovereignty* and *data sovereignty* interchangeably, but these terms concern different levels of autonomy. In February 2020, the President of the European Commission, Ursula von der Leyen, launched the European vision of digital sovereignty to balance the flow and wise use of data while preserving high privacy, security, safety, and ethical standards. Based on its own rules and values [10]. According to the Internet Society, however, digital sovereignty primarily (but not totally) manifests itself in Europe in the form of data [6], converging to the perspective of the GAIA-X project and the IDSA, which consider digital sovereignty a much broader scope than data sovereignty [11].

Digital sovereignty, therefore, comprehends the control over the digital assets necessary to disclose valuable information safely [5]. The complexity of this subject has motivated the separation of duties between the GAIA-X project and the IDSA initiative [11]. While the former leads research in sovereign cloud computing for IDS, the latter recommends business rules and architectural guidelines to promote trusted data exchange in IDS ecosystems. According to the Fraunhofer Institute, *data sovereignty* is the capacity of exclusive self-determination of a natural person or corporate entity concerning data assets [1]. The Chairman of the Board of IDSA, Reinhold Achatz, reinforced this definition by stating that companies understand that data is a valuable source for optimizing their processes. Still, this benefit is currently happening to a minimal extent because companies fear losing control over sensitive data [12].

Braud et al. define *data sovereignty* as the right to determine who is allowed to do what in which context with the data owner's data [5]. Bader et al. advanced on the subject with an ontology to describe an IDS ecosystem. The ontology has six primary partitions, with two directly related to the definition of data sovereignty: the community of trust partition describes data sovereignty from an organizational aspect relating it to concepts such as participant, data connector, certification, and contract, while the commodity partition represents the business aspect of data regarding provenance, quality, access policy, and pricing schemes [13].

In summary, *digital sovereignty subsumes data sovereignty in IDS*, standing for *a company's autonomy to use its digital assets to share private and sensitive data*. Still, dissonant interpretations of this subject can cause at least two practical problems. From one extreme, companies may find the original IDS vision too generic or lacking practice elements to reject it upfront. From another extreme, in filling the demands of specific business cases, companies may extend the IDS guidelines with particular business case requirements, thereby forming siloed implementations with interoperability problems.

III. A Reference Enterprise Architecture for Sovereign Data Sharing in IDS

According to vom Brocke [14], there are at least five design principles to guide the development and reuse of reference models: (1) *configuration*, which consists of specifying the model components enactable by different arrangements of possible interconnections; (2) *instantiation*, which brings giving the user freedom to allocate different resources to implement the components of model; (3) *aggregation*, which comprises designing a model with well-defined interfaces for modularity and interoperability with other models; (4) *specialization*, which relates to the facility to extend the model according to the requirements of specific application domains; and (5) *analogy*, that involves proposing or using a reference model as a design pattern. The reference architecture proposed in this research was designed to promote reuse by instantiation and specialization. However, its application in multiple business cases may unveil reuse possibilities by aggregation, analogy, and configuration.

The Enterprise Architecture model presented in this section conforms to the ArchiMate modeling language, representing the alignment of structural, behavioral, and informational aspects among business, application, and technological layers to realize the defined requirements [15] [16]. Four architectural viewpoints are presented in the following subsections to identify the different stakeholders' perspectives and address other concerns. In the following architectural viewpoints, especially in **Fig. 2**, **Fig. 3**, and **Fig. 4**, the notation for business actors and business roles appears colored in orange to indicate the starting point for the readers in inspecting the model. Another reason is to suggest to the stakeholders what business behaviors they will perform in the IDS-based data-sharing ecosystem.

A. Reconciling Digital Sovereignty and Enterprise Interoperability in IDS

This work takes an Enterprise Architecture perspective to treat digital sovereignty and Enterprise Interoperability requirements seamlessly in IDS. It is motivated by the possibility of making the declarative architectural guidelines offered in the IDS RAM more palatable for companies. The ArchiMate language [17] is suitable for this purpose. It promotes communication between business analysts and IT developers about matching companies' high-level business requirements with underlying software applications and communication infrastructure resources. An ArchiMate specification starts with a motivation viewpoint that associates business requirements with stakeholders, goals, assessments, drivers, and outcomes.

Fig. 1 depicts the elements of the architecture motivation viewpoint. The main stakeholders are (1) product owners of two software companies experienced in Enterprise Integration for the Dutch Logistics sector (and willing to enter IDS ecosystems); (2) representatives of the Dutch Logistics sector,

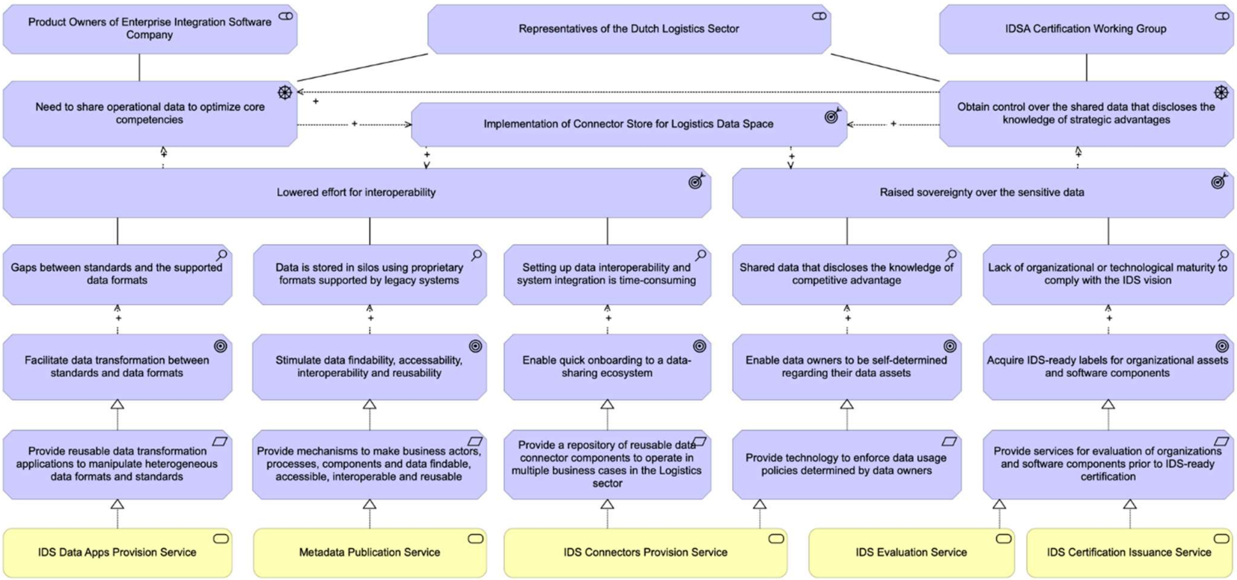
Product Owners of Enterprise Integration Software Company
Representatives of the Dutch Logistics Sector
IDSA Certification Working Group
Need to share operational data to optimize core competencies
Implementation of Connector Store for Logistics Data Space
Obtain control over the shared data that discloses the knowledge of strategic advantages
Lowered effort for interoperability
Raised sovereignty over the sensitive data
Gaps between standards and the supported data formats
Data is stored in silos using proprietary formats supported by legacy systems
Setting up data interoperability and system integration is time-consuming
Shared data that discloses the knowledge of competitive advantage
Lack of organizational or technological maturity to comply with the IDS vision
Facilitate data transformation between standards and data formats
Stimulate data findability, accessability, interoperability and reusability
Enable quick onboarding to a data-sharing ecosystem
Enable data owners to be self-determined regarding their data assets
Acquire IDS-ready labels for organizational assets and software components
Provide reusable data transformation applications to manipulate heterogeneous data formats and standards
Provide mechanisms to make business actors, processes, components and data findable, accessible, interoperable and reusable
Provide a repository of reusable data connector components to operate in multiple business cases in the Logistics sector
Provide technology to enforce data usage policies determined by data owners
Provide services for evaluation of organizations and software components prior to IDS-ready certification
IDS Data Apps Provision Service
Metadata Publication Service
IDS Connectors Provision Service
IDS Evaluation Service
IDS Certification Issuance Service

Fig. 1. Motivation Viewpoint of the Reference Enterprise Architecture for IDS

due to the domain's immediate demands for the IDS vision; and (3) the IDSA Certification Working Group – currently the only organization responsible for the certification scheme that will label organization and software components as trusted in IDS ecosystems. The list of stakeholders is currently limited but authentic and representative.

The stakeholders' main drivers are (1) sharing operational data to optimize core competencies and (2) control over competitive advantage knowledge disclosure, which can contribute to implementing a data connector store, i.e., a repository of descriptions of IDS Connectors [3]. This artifact can lead to practical outcomes such as lowering companies' efforts to interoperate and enforce sovereignty over sensitive data. These outcomes relate to gaps identified in the literature, which correspond to state-of-the-art IDS capabilities assessments. Specific goals derive from these challenges, realized by five types of requirements. The first three requirements relating to the three Enterprise Interoperability barriers: (1) conceptual, i.e., syntactic and semantic issues on information exchange; (2) organizational, i.e., distribution of responsibilities necessary to enforce interoperability; and (3) technological, e.g., industrial standards for Enterprise data exchange [18]. In terms of the four interoperability levels suggested by the EIF to be implemented in a particular data sharing ecosystem [19], these first three requirements align with: (1) technical interoperability, i.e., providing applications and infrastructures for secure communications; (2) semantic interoperability, i.e. ensuring the format and meaning of the shared data are preserved and understood; and (3) organizational interoperability, i.e. enabling organizations to discover and connect to achieve shared goals. Meanwhile, the fourth interoperability level, the legal interoperability, is associated with the last two requirements that highlight the

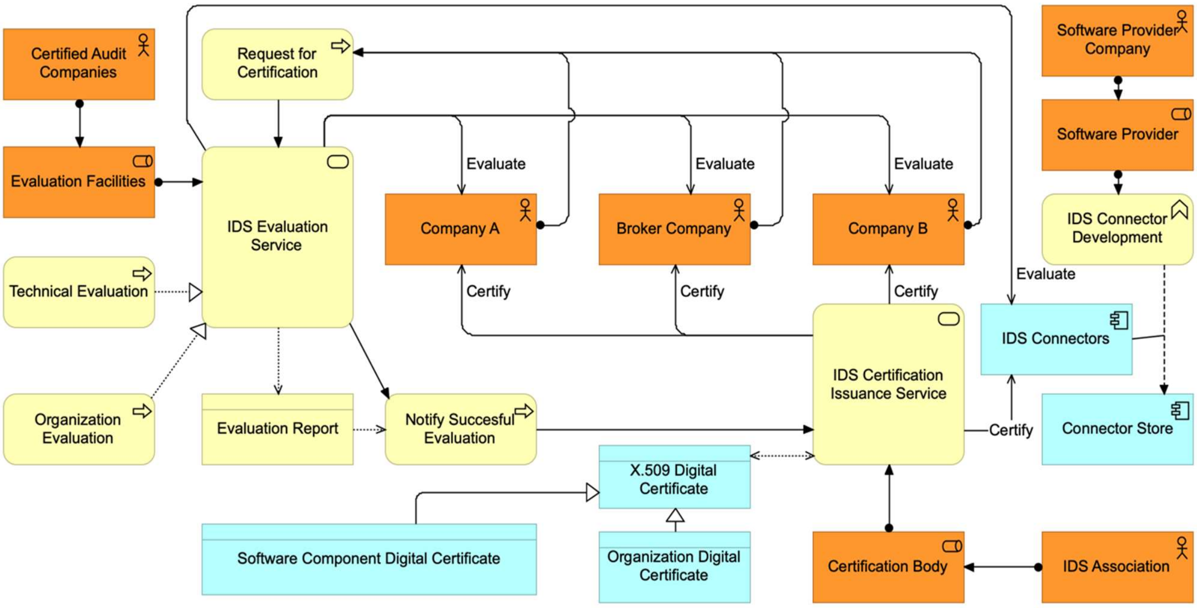

Fig. 2. Certification Viewpoint of the Reference Enterprise Architecture for IDS

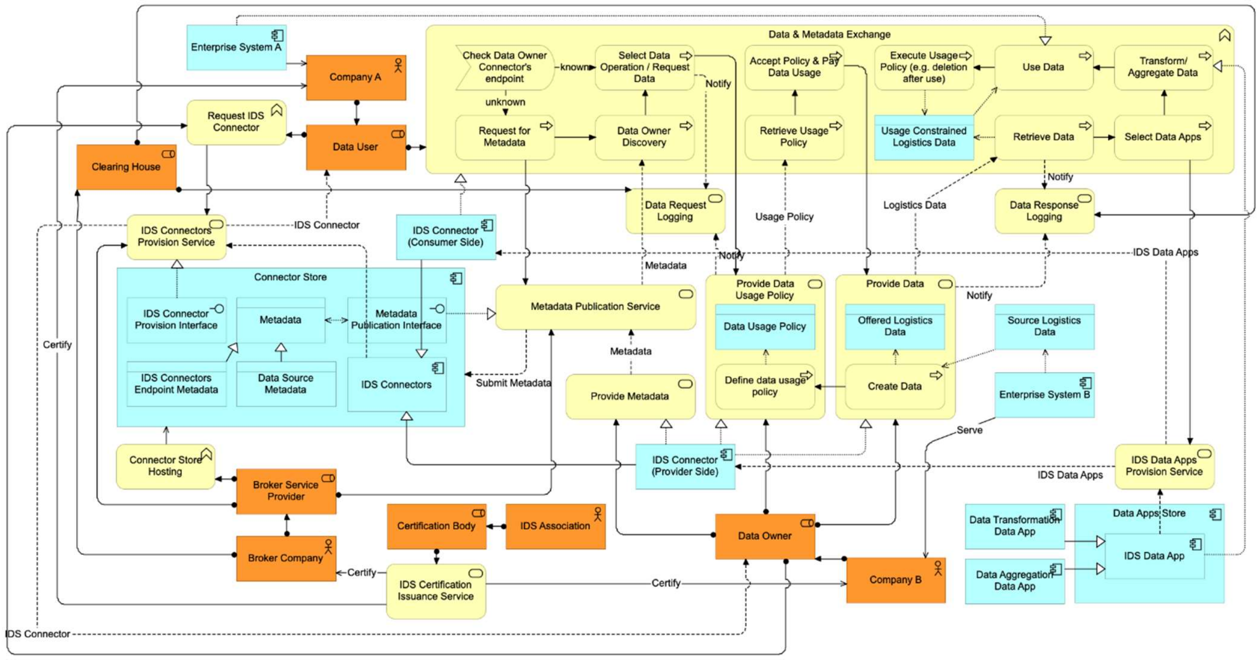

Fig. 3. Data and Metadata Exchange Viewpoint of the Reference Enterprise Architecture for IDS

digital sovereignty vision in IDS. The assessment-goal-requirement triples trace how the architecture shall address these criteria, which are summarized as follows:

- **Req. 1:** *Provide reusable data transformation applications to manipulate heterogeneous data formats and standards* [3] – a data connector can orchestrate multiple data transformation apps differently;
- **Req. 2:** *Provide mechanisms to make business actors, processes, components, and data findable, accessible, interoperable, and reusable* [13] – this principle applies to software components and organizational assets;
- **Req. 3:** *Provide a repository of reusable data connector components to operate in multiple business cases in the Logistics sector* [3, 20] – multi-sided data-sharing may demand a choreography of data connectors;
- **Req. 4:** *Provide technology to enforce data usage policies determined by data owners* [13, 20] – data analytics applications can readily disclose knowledge of competitive advantage from ungoverned data;
- **Req. 5:** *Provide services for evaluating organizations and software components before IDS-ready certification* [4, 16, 21] – IDSA partially delegates this task to independent audit companies.

These requirements comprehend the types of demands that specific business cases shall extend. The motivation viewpoint of the architecture also indicates the types of business services that should realize the requirements (see the bottom part of **Fig. 1**). Such services include data apps, metadata publication, data connectors, (pre-) evaluation of IDS certification compliance, and final IDS certification issuance. The following section describes this service in detail.

B. Certification Viewpoint

The IDS RAM describes the candidate participants as being evaluated and certified before participating in a data space [3]. **Fig. 2** depicts the core competencies assigned to the participant actors. These activities comprehend two business functions: the IDS evaluation service provided by the evaluation facilities and the IDS certification issuance service. These business functions trace back to **Req. 5** above since, in sovereign data space, data exchange is only allowed to take place between participants who are certified and using software components that are certified as well [3, 4].

Upon request, the certification process starts from a candidate participant to an evaluation facility, e.g., an independent audit company. The evaluation facilities evaluate the candidates' organizational assets and software components regarding their maturity to operate in an IDS ecosystem. After a successful evaluation., the facility sends a pre-evaluation report to the IDSA certification body, granting the candidate company an X.509 digital certificate. The evaluation facilities and the certification body are also responsible for evaluating and issuing certificates for the IDS Connectors before being published in the connector store. That will ensure that the IDS Connectors used in the data space comply with the IDS specifications.

C. Data and Metadata Exchange Viewpoint

The viewpoint illustrated in **Fig. 3** presents the essential roles, activities, and components enacting data sharing among the participants of an IDS ecosystem. Certified IDS Connectors enforce data sovereignty based on data usage policies. The actor roles defined in the IDS RAM found essential to this viewpoint are the *data owner*, the *data user*, and the *broker service provider* [3]. A *data owner* creates and publishes data. A *data user* is an entity that requests the data from the data owner, using it in compliance with the data usage policy. A *broker service provider* provides the interface to receive, maintain and publish metadata describing the participant's IDS Connectors and catalogs of data sources. The provisioning of this service, defined as the *Metadata Publication Service* (provided by the *broker service provider*),

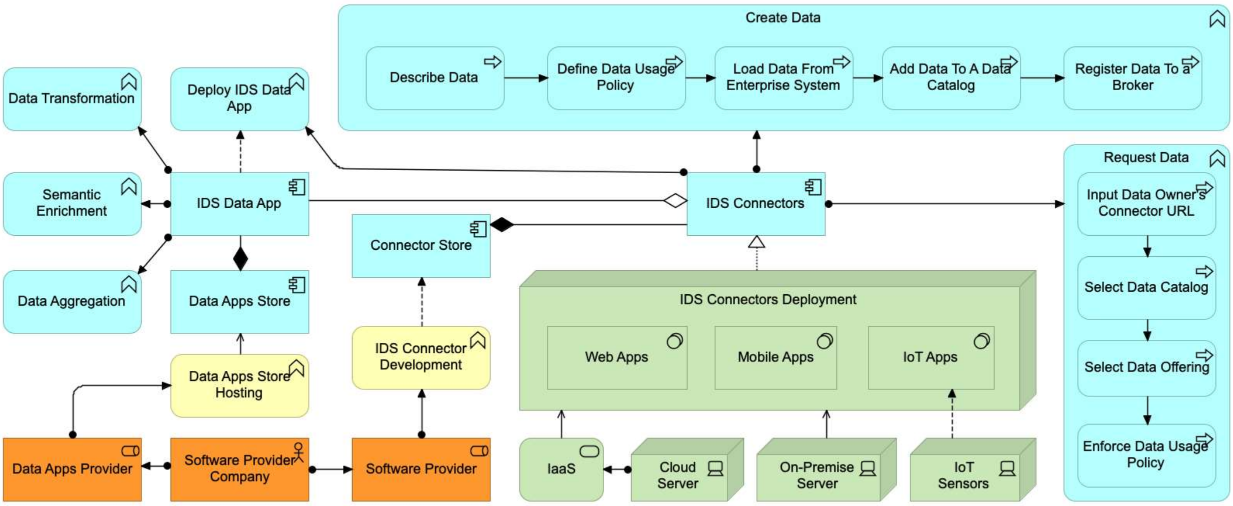

Fig. 4. Infrastructure Viewpoint of the Reference Enterprise Ar chitecture for IDS

aims to realize the **Req. 2** to make the participants, software components, and data sources discoverable.

According to the IDS RAM, the broker service provider may accumulate business roles, e.g., a clearinghouse [3]. This role accounts for logging all activities related to data exchange in an IDS ecosystem. A *clearinghouse* can resolve conflicts related to failures or inconsistencies in data transactions.

After completing the certification process, *data owners* and *data users* can use one or more IDS Connectors from a *connector store* maintained by a particular *broker service provider* to exchange private data in an IDS ecosystem. This *connector store* acts as a repository that provides the participants with IDS connectors suitable to their demands, realizing **Req. 3** and **Req. 4** at the same time. Different connector types demand specific deployment configurations, e.g., base connector, trusted connector, IoT connector, or mobile connector, and deployed on-premise or in the cloud [3]. After setting up the connector's deployment environment, the peers can start sharing data. Therefore, the data owner needs to create catalogs of data offerings pulled out of internal information systems and define their usage policies. The following IDS usage control directives [22] may apply to modeling these policies:

- **Connector restriction:** allows data usage through a specific connector;
- **Duration restriction:** allows data usage for a specified period;
- **Some usage restrictions:** allows data usage multiple times;
- **Security level restriction:** allows data access only for connectors with a specified security level (i.e., a trusted connector);
- **Usage and deletion after:** allow data usage within a specified time interval with the restriction to delete it at a limited time stamp.

Data users may start requesting data when the *data owner* defines the data usage policies. The *data user* checks the identity of the participants who want to share data and requests access to their IDS Connector's endpoints. Assuming they have not shared data before, the *data user* can request historical metadata from the *broker service provider*. After that, the *data user* selects the data (or data operation) of interest, notifying the *clearinghouse* about the request. Therefore, the *data user* must accept the *data owner's* usage policy before requesting data.

The *data user* may need to transform the data received into a format compatible with its enterprise systems to use it. An *IDS Data App Store* can serve this purpose by providing reusable data transformation applications, facilitating **Req. 1** in the process. An IDS Connector may orchestrate different data apps internally. If a data transformation is successful, the requested data becomes available to the *data user*.

D. Infrastructure Viewpoint

This viewpoint focuses on the functional capabilities and deployment environment of the IDS Connectors and the IDS Data Apps. According to IDSA, a data space aims to facilitate data transfer to and from participants' systems, be it enterprise systems (e.g., CRM, ERP, etc.) or cyber-physical systems (i.e., IoT-enabled systems), by using a system adapter that supports necessary data format transformation and data usage policy enforcement [23]. An IDS Connector constitutes this system adapter, which should help different implementation types (e.g., web apps, mobile apps, or IoT apps) and deployment environments, such as on-premise or cloud environments, to serve multiple cases. Therefore, a software provider is needed in the data space to develop and provide the IDS Connectors to support the participants with several types of connectors [3].

As highlighted in **Fig. 4**, a set of application processes support the IDS Connector in creating data. The first step is to describe the data offered. Next, the IDS Connector has to facilitate the user to define and attach the usage policy to the data. Then, the IDS Connector must load the data from the data sources. The data generated is pulled out of the company's internal information systems (e.g., Enterprise Resource Planning, Production Planning System, Transport Management System) and retrieved from a database directly or through a REST API. The final step comprises registering the newly uploaded data to a particular broker service provider, which will make the data offering discoverable by the other participants of the data space.

Another set of business processes supports the IDS Connectors in requesting data. Data users perform this function. Its first step comprises accessing a data owner's data connector through its endpoint URL. Next, the data user must select a data provider's data catalog. The IDS Connector returns the data usage policy attached to a particular data offering to the data user. After the approval, the IDS Connector downloads the selected data from the data owner and starts enforcing and monitoring data usage based on its corresponding access policy. An IDS Connector also supports the deployment and execution of IDS Data Apps. According to the IDS RAM, data apps provide the IDS participants with functionalities of data transformation, aggregation, or semantic enrichment [3]. Such applications are developed and delivered by a data apps provider, who is also responsible for describing them with metadata to make them discoverable and trusted.

E. Discussion

The requirements presented in the motivation viewpoint synthesize core demands of Enterprise Interoperability and digital sovereignty for IDS. However, these requirements shall extend to accommodate needs from specific business cases. Likewise, the ArchiMate representation of the architecture aims to simplify the declarative guidelines provided by the IDS RAM and serve as a communication basis to guide the instantiation and specialization of its organizational and software components.

It is also worth classifying the architecture proposed in this work to promote clarity and reuse. Fettke and Loos [24] provided a framework to organize reference models, elaborating on a model's *domain-independent* and *domain-dependent* characteristics. *Domain-independent* characteristics comprehend the language of description, views, model size, computational complexity, performance, and qualitative assessments. *Domain-dependent* ones include the economic activity, industrial sector, types of enterprises involved, functional area, phase of a transaction, and categories of target application systems. Based on this framework, the classification of the architecture proposed here is summarized in **Table 1**.

TABLE I. CLASSIFICATION OF THE REFERENCE ENTERPRISE ARCHITECTURE FOR IDS BASED ON THE FRAMEWORK OF *FETTKE AND LOOS [24]*

Domain-Independent Characteristics	***Modeling language***	ArchiMate
	Architectural views	Structural views: strategy, business, application, and infrastructure
	Model size	Medium size: 100-300 constructs
	Computational complexity	Not evaluated
	Performance	Not evaluated
	Qualitative assessment	Feasible for a DevOps team
Domain-Dependent Characteristics	***Economic activity***	Transport Logistics
	Industrial sector	European IDS, the Dutch Logistics sector
	Types of enterprises involved	Transport Logistics, Enterprise Integration software companies, IDSA (certification agency), App providers, clearinghouses
	Functional area	Production
	Phase of transaction	Information exchange, billing, and payment for data usage
	Target application systems	Reporting and controlling systems

Goel, et al. [25] provided the types of architectural views considered here as a domain-independent characteristic of a reference model. According to them, architectural views may unveil: (1) *structure*, e.g., components and connections between them; (2) *function*, i.e., a scheme specifying preconditions and postconditions of operations changing the state of the architecture (or system invariants); and (3) *behavior*, comprehending a sequence of states and transitions between them. The views of the architecture proposed here are essentially *structural*. Although the data/metadata exchange view brings data flow elements, a formal specification of the architecture's dynamics is part of future work.

The architecture classification summarized in **Table 1** also aims to promote the benefits of reference model classification as identified by [24]. Such benefits include (1) *model standardization*, as the architecture translates some of the critical technical guidelines promoted by the International Data Spaces Association (IDSA); (2) *model integrity and consistency*, as it conforms to the syntax of the Archimate, which is a widely used business architecture notation; (3) model discoverability, as its characteristics can guide business experts and IT developers on instantiating or specializing its components. The following section reports on a preliminary architecture assessment with an expert consultation panel from the Dutch Logistics sector, which elaborates on the qualitative evaluation of its technical feasibility.

IV. VALIDATION

An expert panel assessed the architecture proposed in this work regarding its potential acceptance by SMEs. The experts represented organizational stakeholders heading the motivation viewpoint of the architecture presented in **Fig. 1**. However, the consultation did not include an IDSA Certification Working Group expert. This choice minimizes research bias, as the success of the IDS vision may depend not only on the idealized requirements proposed by IDSA but mainly on the feedback from companies that will provide the organizational and software assets to implement such a vision. A short description of the organizations and corresponding experts follows:

- **EMONS Group** - a privately owned group of companies with expertise in the logistics of glass, non-stackable goods, and humus-rich soil improvers. This organization believes delivering value to customers depends not only on offering the lowest price but also on caring about environmental aspects such as CO2 emissions per transport unit. In this research, the organization is interested in innovative and sustainable solutions to optimize its internal operations. The expert from EMONS is a business consultant with more than twenty years of experience in Enterprise Transformation for SMEs.
- **SUTC** - the *Uniform Transport Code Foundation* (helps logistics companies share data securely and efficiently. In this research, SUTC is interested in implementing IDS data connectors from reusable data transformation applications. The expert from SUTC is a policy advisor with more than fifteen years of experience in interfacing ICT and Transport Logistics,

responsible for the development, adoption, and implementation of the Open Trip Model (OTM) and involved in DALTI, TransFollow (e-CMR), i-SHARE, DefLOG, and Basic Data Infrastructure (BDI) initiatives.

- **TNO** - the *Netherlands Organisation for Applied Scientific Research* (free translation) supports a project named DASLOGIS, which aims to leverage the Dutch Logistics Data Space (DLDS) to federated data spaces [19]. In this research, TNO is interested in sharing experiences in implementing IDS Connector stores. The expert from TNO is a business consultant with more than seventeen years of experience in telecommunications and large-scale ICT architectures, currently leading several data-sharing research projects.
- **CAPE Group** - a consultancy company that combines Enterprise Integration tools (e.g., Mendix, eMagiz, and Power BI) to promote Enterprise Transformation. The company has considerable experience customizing Enterprise Integration for the Dutch Logistics sector. The experts from CAPE are two business consultants with five and eight years of experience, respectively. They own a *control tower* solution that can be used as a proof-of-concept environment to deploy a *clearinghouse* in the IDS reference architecture.
- **eMagiz** - an Enterprise Integration software provider specialized in delivering data transformation applications to the Dutch Logistics sector. The company is interested in combining open-source IDS components with its private software components to offer the best cost-benefit solutions for SMEs to join IDS. The eMagiz expert is a project manager with almost ten years of experience in data integration for the Dutch Logistics sector.

The experts assessed the architecture proposal in separate bilateral workshops. The researchers requested the participants' permission to record the meeting to enable transcript analysis. After that, the researchers explained the architecture in detail and invited the participants to answer a questionnaire adapted from the six questions on technology acceptance by SMEs proposed by Bernaert et al. [9]. The researchers considered these questions open and straightforward enough to promote discussion among business experts. Besides, as SMEs comprise approximately 90% of businesses worldwide, practical research should benefit this sector [8]. The questionnaire follows with a synthesis of the answers provided by the experts:

Question 1: Are the architecture requirements complete and consistent?

Experts: *Simplification is needed, but treating Enterprise Interoperability and digital sovereignty requirements separately will change the rationale of the architecture. Conditions seem complete and consistent enough to stimulate initial discussion. The expert from EMONS noted that the industry is heading towards an ecosystem where value creation and service offerings are executed through collaborative systems, and this proposed architecture embraces this vision. Another remark for establishing such a vision is that the gaps in organizational maturity and technological capabilities of participating companies need to be kept at a bare minimum. Therefore, realizing such Enterprise Interoperability is resource-consuming and requires leadership perseverance in executing the change management process.*

Question 2: How could the architecture motivate a company to enter an IDS?

Experts: *The architecture could make the IDSA guidelines more tractable and palatable for SMEs, but it may demand organizational and technological maturity. The expert from EMONS stated that Enterprise Transformation and Enterprise Interoperability projects such as the ones proposed in this architecture are always challenging, especially for SMEs with limited organizational resources and IT capabilities to implement the elements indicated in the architecture. They not only ask for the participating companies to keep the gaps in organizational maturity and technological capabilities at a bare minimum but also demand leadership perseverance in executing the change management process to embrace the project outcome envisioned by the IDS (indicated as the outcomes in **Fig.1**).*

To approach these challenges, the experts from CAPE Groep and eMagiz recommend starting the project small by first identifying the biggest problem in the organization and identifying its properties. Therefore, the architecture can help by providing a company with a roadmap to solve specific data-sharing issues. Additionally, the operationalization of a business case that addresses a particular set of pain points companies face in establishing a data-sharing ecosystem is also necessary to motivate the use of this architecture and support proof-of-concept implementations to prospect the architecture's feasibility and immediate benefits.

Question 3: What kind of IT skills are required to use the architecture?

Experts: *The expert from eMagiz stated that a team consisting of Software Developers to build the applications and DevOps to prepare the infrastructure for the application deployment on different deployment environments and configurations would be required. Additionally, knowledge of IoT-related software and hardware integration will also be necessary if the use of IoT-Connector serves the company's use case. It will not be an easy task for most SMEs, but it is not impossible.*

Question 4: How easy would it be to use or implement the architecture without assistance from external experts?

Experts: *Judging by the reactions of the experts, the need to have external experts are not necessary. The expert from EMONS stated that many companies have become acquainted with the ArchiMate modeling language in recent years. Hence, familiarity with the ArchiMate models of the proposed Reference Architecture Model for IDS and knowledge of the IDSA Reference Architecture Mode is required.*

Question 5: To what extent could the architecture help a company re-engineer its business processes to share data with its business partners?

Experts: *It could help optimize and refactor business processes. The experts from TNO, SUTC, and CAPE Groep mentioned that, in a business context, ideas grow from small situations that evolve into oversized cases. In that sense, the architecture helps answer questions: How should a company use a clearinghouse? How do you share data in IDS in a well-structured way? Especially in the Dutch Logistics domain,* *much of the information is still transferred by email, paper documents, or even phone calls. That is typical for SMEs, but large frontrunner companies also have complex information systems. In the end, the architecture could facilitate the transfer of information from legacy systems to IDS ecosystems.*

Question 6: Would a company's CEO be involved in using the architecture?

Experts: *The architecture may look over-engineered for CEOs, quoted the expert from eMagiz. However, business analysts or architects with experience in business process modeling might be efficient in interpreting and using it. In addition, experts from EMONS Group also mentioned that people tend to accumulate roles in companies with less than 50 employees. In that case, the chance of finding senior people involved in daily operation tasks is high. They might be interested in using the architecture as well.*

Question 7: Would the architecture's benefits be higher than its costs and risks?

Experts: *As pointed out by the experts from CAPE Groep and eMagiz, this question claims for business cases and how the companies would approach them in IDS ecosystems. The expected benefits can be high, but that would demand strategic business analysis. A company might not adopt the whole architecture at first glance. Still, it could build proof-of-concept implementations, evaluating their immediate return on investment to decide on adopting the architecture to a fuller extent. Simplification requires generalization that demands reasoning over multiple cases. However, the architecture must overfit and not underfit business cases.*

In summary, according to the experts, balancing Enterprise Interoperability and Data Sovereignty requirements is sufficient to demonstrate the current feasibility of the architecture. Attempting to reconcile adaptability, openness to standards, or scalability in the architecture might turn it cumbersome in the short term. Besides, the ArchiMate specification facilitates business experts' communication about the architectural elements. However, real-world business cases can help make its return on investment clearer for SMEs interested in deploying it, if not totally, but enough to enact minimally sovereign data sharing in IDS. Moreover, a company willing to adopt the architecture should first assess its organizational and technological maturity to embrace the IDS vision and assume risks related to Enterprise Transformation.

V. Related Work

The research initiatives most closely related to this work are: (1) the IDS Information Model proposed by Bader, et al. [13]; (2) the +CityxChange architecture framework of Petersen, et al. [26]; (3) the Federated Network-Model Approach for Multilateral Data Sharing proposed by Bastiaansen, et al. [19]; the Governance Structure for Federated Digital Platforms of Nübel, et al. [27]; and the Smart Factory Web of Usländer, et al. [28].

The IDS Information Model proposed by Bader, et al. [13] is a domain ontology describing an IDS ecosystem. The extensively documented and verified OWL model contains the central organizational roles, system components, and interactions of a data space, thereby promoting conceptual interoperability. The authors refer to the ontology as a *cornerstone for any IDS-related implementation*, prospecting the impact of its use in industrial platforms. However, they focus more on providing guidelines to develop technical implementations compliant with the ontology than reporting how feasible that would be for companies in practice.

The +CityxChange architecture framework of Petersen, et al. [26] aims to guide the designers of smart cities to identify types of services that could create value out of citizens' data. It suggests treating data sovereignty in layers corresponding to an architecture viewpoint to treat data access, interoperability, privacy, regulations, and ownership issues. However, the authors provide no technical guidelines for extending the architectural layers or any preliminary assessment from business stakeholders.

The Federated Network-Model Approach for Multilateral Data Sharing proposed by Bastiaansen, et al. [19] assumes that interoperability in IDS develops on four levels: technical, semantic, organizational, and legal. The authors address problems companies may face when reviewing legal contracts before sharing data in federated data spaces. Nevertheless, the authors do not elaborate on how changes in legal contracts could impact companies' existing technical solutions for data sharing or vice-versa.

The Governance Structure for Federated Digital Platforms of Nübel, et al. [27] complements the IDSA RAM [3] by proposing integrating requirements from the most promising cases to customize federated data-sharing infrastructures. Despite its theoretical soundness, the work does not give a practical example of how a federated IDS architecture could reconcile the business demands of disparate business cases.

Finally, the Smart Factory Web of Usländer, et al. [28] proposes an architecture for sharing data in open marketplaces for the domain of Industrial Production. Its stakeholders fit the business roles defined in the IDSA RAM [3]. It aims to cope with adaptability, interoperability, openness to standards, sovereignty, scalability, and security requirements. However, the authors do not indicate how a company could use the architecture's complex Petri Nets and ontologies to derive a technically feasible IDS architecture. Besides, the academic projects validating the architecture represent idealizations rather than real-world business cases.

There are similarities and differences between these approaches and the one proposed in this paper. First, they also attempt to treat sovereignty and interoperability requirements in conjunction (but not digital sovereignty or Enterprise Interoperability, specifically). Second, they use the IDSA RAM as a starting point to explore specific requirements (except for the work of Nübel, et al. [27] that proposes to start eliciting requirements from business cases). Yet, the main difference is that those approaches lack feedback from companies interested in deploying their perspectives on implementing the IDS vision.

VI. Conclusion and Research Outlook

This paper introduced an Enterprise Architecture to help companies decide which organizational and software components to deploy before entering an IDS ecosystem. It also aims to make the guidelines provided by the IDS RAM more understandable to companies interested in joining IDS ecosystems soon. By customizing or reusing its elements, a company could strive for digital sovereignty without distancing too much from the IDS vision or compromising Enterprise Interoperability with other IDS actors envisioning the same principles. Besides, the ArchiMate specification of the architecture supports communication between business analysts and architects and advances state-of-the-art IDS technical specifications.

According to the framework of Fettke and Loos [24], this architecture has the characteristics of a specific and domain-dependent reference model. It is classified as specific since it is expressed in a specific architectural modeling language with precise semantics, i.e. ArchiMate. At the same time, domain-dependent characteristic comes from how the architecture identifies: (1) IDS-compliance Dutch Transport Logistics as the prominent economic activity and industrial sector; (2) information exchange, billing, and payment for data usage as the phase of transaction; and (3) reporting and controlling system as the highlighted target application systems, etc.

Three limitations threaten the validity of this work. First, the data sovereignty and EI requirements considered in the motivation viewpoint of the architecture are considerably generalized. However, emergent business cases could help decompose these requirements into more concrete functional and non-functional properties, impacting the composition of the architecture, e.g., by the inclusion of new components or rearrangement of its internal relationships. Second, the expert panel consultation essentially represented the interests of the Dutch Logistics sector in deploying future IDS infrastructures, but representatives from different domains and nationalities could extend the panel. Third, SMEs' six questions about technology acceptance are relatively broad and demand solid expertise from the respondents.

This work shall continue in three directions. First, a business case will help refine the architecture requirements into more specific taxonomies of data sovereignty and Enterprise Interoperability properties and restrictions. Second, an ontology to describe IDS data connectors could make them findable, accessible, interoperable, and reusable. Last, a proof-of-concept implementation of the architecture could better demonstrate its feasibility by combining reusable and open software components provided by IDSA with privately-owned Enterprise Integration solutions.

VII. Acknowledgment

This research is part of the CLICKS project, which is financially supported by the Dutch Ministry of Economic Affairs via TKI DINALOG and NWO under Accelerator 2019 grant Nº 439.19.633. CLICKS is the acronym for Connecting Logistics Interfaces, Converters, Knowledge and Standards.

References

[1] B. Otto, M. ten Hompel, and S. Wrobel, "International data spaces," in *Digital Transformation*: Springer, 2019, pp. 109-128.

[2] K. Spanaki, E. Karafili, and S. Despoudi, "AI applications of data sharing in agriculture 4.0: A framework for role-based data access control," *International Journal of Information Management,* vol. 59, p. 102350, 2021.

[3] IDSA, "IDSA Reference Architecture Model Version 3.0," International Data Spaces Association, Berlin, White paper April 2019 2019.

[4] IDSA, "IDSA Rule Book Version 1.0," International Data Spaces Association Berlin, White Paper of the IDS Association December 2020 2020.

[5] A. Braud, G. Fromentoux, B. Radier, and O. Le Grand, "The Road to European Digital Sovereignty with Gaia-X and IDSA," *IEEE Network,* vol. 35, no. 2, pp. 4-5, 2021.

[6] K. Komaitis, "Europe's ambition for digital sovereignty must not undermine the Internet's values," *Computer Fraud & Security,* vol. 2021, no. 1, pp. 11-13, 2021.

[7] V. Kalogirou and Y. Charalabidis, "The European union landscape on interoperability standardisation: status of European and national interoperability frameworks," in *Enterprise Interoperability VIII*: Springer, 2019, pp. 359-368.

[8] R. J. Wieringa, *Design science methodology for information systems and software engineering*. Springer, 2014.

[9] M. Bernaert, G. Poels, M. Snoeck, and M. De Backer, "Enterprise architecture for small and medium-sized enterprises: a starting point for bringing EA to SMEs, based on adoption models," in *Information systems for small and medium-sized enterprises*: Springer, 2014, pp. 67-96.

[10] U. Von der Leyen, "A Union that strives for more," *My agenda for Europe. Political guidelines for the next European Commission,* vol. 2024, p. 2019, 2019.

[11] "GAIA-X and IDS," Fraunhofer Institute for Software and Systems Engineering ISST Berlin, Position Paper of members of the IDS Association January 2021 2021.

[12] B. Otto, "Interview with Reinhold Achatz on "data sovereignty and data ecosystems"," *Business & Information Systems Engineering,* vol. 61, no. 5, pp. 635-636, 2019.

[13] S. Bader *et al.*, "The International Data Spaces Information Model– An Ontology for Sovereign Exchange of Digital Content," in *International Semantic Web Conference*, 2020: Springer, pp. 176-192.

[14] J. vom Brocke, "Design principles for reference modeling: reusing information models by means of aggregation, specialisation, instantiation, and analogy," in *Reference modeling for business systems analysis*: IGI Global, 2007, pp. 47-76.

[15] M. Lankhorst, T. Halpin, J. Hoogervorst, M. O. t. Land, R. G. Ross, and R. Winter, Eds. *Enterprise Architecture at Work*, Fourth Edition ed. (The Enterprise Engineering Series). Enschede: Springer, 2016, p. 360.

[16] M. E. Iacob, L. O. Meertens, H. Jonkers, D. A. C. Quartel, L. J. M. Nieuwenhuis, and M. J. van Sinderen, "From enterprise architecture to business models and back," *Software & Systems Modeling,* vol. 13, no. 3, pp. 1059-1083, 2014/07/01 2014, doi: 10.1007/s10270-012-0304-6.

[17] A. Josey, *ArchiMate® 3.0. 1-A pocket guide*. Van Haren, 2017.

[18] N. Daclin, D. Chen, and B. Vallespir, "Methodology for enterprise interoperability," *IFAC Proceedings Volumes,* vol. 41, no. 2, pp. 12873-12878, 2008.

[19] H. Bastiaansen, S. Dalmolen, M. Kollenstart, and T. M. van Engers, "User-Centric Network-Model for Data Control with Interoperable Legal Data Sharing Artefacts," 2020.

[20] J. Zrenner, F. O. Möller, C. Jung, A. Eitel, and B. Otto, "Usage control architecture options for data sovereignty in business ecosystems," *Journal of Enterprise Information Management,* 2019.

[21] M. Bartsch *et al.*, "Framework for the Industrial Data Space Certification Scheme," Berlin, 2019.

[22] A. Eitel *et al.*, "Usage Control in the International Data Spaces 3.0," 03/31 2021.

[23] L. Nagel and D. Lycklama, "Design Principles for Data Spaces - Position Paper (1.0)," 2021.

[24] P. Fettke and P. Loos, "Classification of reference models: a methodology and its application," *Information systems and e-business management,* vol. 1, no. 1, pp. 35-53, 2003.

[25] A. K. Goel, S. Rugaber, and S. Vattam, "Structure, behavior, and function of complex systems: The structure, behavior, and function modeling language," *Ai Edam,* vol. 23, no. 1, pp. 23-35, 2009.

[26] S. A. Petersen, Z. Pourzolfaghar, I. Alloush, D. Ahlers, J. Krogstie, and M. Helfert, "Value-added services, virtual enterprises and data spaces inspired Enterprise architecture for smart cities," in *Working Conference on Virtual Enterprises*, 2019: Springer, pp. 393-402.

[27] K. Nübel, M. M. Bühler, and T. Jelinek, "Federated Digital Platforms: Value Chain Integration for Sustainable Infrastructure Planning and Delivery," *Sustainability,* vol. 13, no. 16, p. 8996, 2021.

[28] T. Usländer *et al.*, "Smart Factory Web—A Blueprint Architecture for Open Marketplaces for Industrial Production," *Applied Sciences,* vol. 11, no. 14, p. 6585, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/14/6585.