Metadata of the chapter that will be visualized in SpringerLink

Book Title	Cooperative Information Systems	
Series Title		
Chapter Title	A Data Connector Store for International Data Spaces	
Copyright Year	2022	
Copyright HolderName	The Author(s), under exclusive license to Springer Nature Switzerland AG	
Author	Family Name	**Firdausy**
	Particle	
	Given Name	**Danniar Reza**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Twente
	Address	Drienerlolaan 5, Enschede, 7522 NB, The Netherlands
	Email	d.r.firdausy@utwente.nl
	ORCID	http://orcid.org/0000-0001-9743-9754
Corresponding Author	Family Name	**Alencar Silva**
	Particle	**de**
	Given Name	**Patrício**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Twente
	Address	Drienerlolaan 5, Enschede, 7522 NB, The Netherlands
	Division	Graduate Program in Computer Science
	Organization	UERN/UFERSA, Rio Grande do Norte State University (UERN), Federal University of the Semi-Arid Region (UFERSA)
	Address	Mossoró, RN, Brazil
	Email	p.dealencarsilva@utwente.nl
	ORCID	http://orcid.org/0000-0001-6827-1024
Author	Family Name	**Sinderen**
	Particle	**van**
	Given Name	**Marten**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Twente
	Address	Drienerlolaan 5, Enschede, 7522 NB, The Netherlands
	Email	m.j.vansinderen@utwente.nl

	ORCID	http://orcid.org/0000-0001-7118-1353
Author	Family Name	**Lacob**
	Particle	
	Given Name	**Maria-Eugenia**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Twente
	Address	Drienerlolaan 5, Enschede, 7522 NB, The Netherlands
	Email	m.e.iacob@utwente.nl
	ORCID	http://orcid.org/0000-0002-4004-0117

Abstract	The International Data Spaces Association (IDSA) has promoted the idea of International Data Spaces as a place for companies to share data with trust and security enforced by software and organizational competence. There has been considerable progress in delivering corporate guidelines, technical specifications, and software components available for testing and deploying applications to support IDS-based ecosystems, such as the IDS data connectors classified by the Fraunhofer Institute. However, full implementation of IDS applications seems still complex and expensive for small and medium enterprises (SMEs). A possible strategy to deal with such an issue is to break the IDSA specification's complexity into smaller pieces and build small IDS ecosystems formed by its core business roles (e.g., data owners, users, and broker service providers). In this context, this paper addresses the problem of designing an application to support the broker service provider's role in operating in an IDS-based ecosystem. This research, therefore, follows a Design Science approach in a three-step process. First, it investigates problems of practical relevance elicited from the IDSA guidelines in combination with requirements provided by representatives of the Dutch Logistics sector. Second, it gives design to tackle the problem by combining Semantic Web, Linked Data, and Enterprise Architecture modeling artifacts. Last, it validates the architecture of the broker service provider's application by demonstrating its technical feasibility, innovation, and software integration.

Keywords (separated by '-')	Broker service provider - Enterprise architecture - International Data Spaces - Linked data - Ontology - Semantic Web

A Data Connector Store for International Data Spaces

Danniar Reza Firdausy[1], Patrício de Alencar Silva[1,2](✉), Marten van Sinderen[1], and Maria-Eugenia Iacob[1]

[1] University of Twente, Drienerlolaan 5, Enschede 7522 NB, The Netherlands
{d.r.firdausy,p.dealencarsilva,m.j.vansinderen,m.e.iacob}@utwente.nl

[2] Graduate Program in Computer Science, UERN/UFERSA, Rio Grande do Norte State University (UERN), Federal University of the Semi-Arid Region (UFERSA), Mossoró, RN, Brazil

Abstract. The International Data Spaces Association (IDSA) has promoted the idea of International Data Spaces as a place for companies to share data with
AQ1 trust and security enforced by software and organizational competence. There has been considerable progress in delivering corporate guidelines, technical specifications, and software components available for testing and deploying applications to support IDS-based ecosystems, such as the IDS data connectors classified by the Fraunhofer Institute. However, full implementation of IDS applications seems
AQ2 still complex and expensive for small and medium enterprises (SMEs). A possible strategy to deal with such an issue is to break the IDSA specification's complexity into smaller pieces and build small IDS ecosystems formed by its core business roles (e.g., data owners, users, and broker service providers). In this context, this paper addresses the problem of designing an application to support the broker service provider's role in operating in an IDS-based ecosystem. This research, therefore, follows a Design Science approach in a three-step process. First, it investigates problems of practical relevance elicited from the IDSA guidelines in combination with requirements provided by representatives of the Dutch Logistics sector. Second, it gives design to tackle the problem by combining Semantic Web, Linked Data, and Enterprise Architecture modeling artifacts. Last, it validates the architecture of the broker service provider's application by demonstrating its technical feasibility, innovation, and software integration.

Keywords: Broker service provider · Enterprise architecture · International Data Spaces · Linked data · Ontology · Semantic Web

1 Introduction

There has been a growing discussion around data sovereignty for people and companies in Europe. According to Braud et al. [1], data sovereignty means *providing data owners with complete control over their data and digital identities*, which demands *defining who is allowed to do what in which context with the data shared by the data*

M. Sellami et al. (Eds.): CoopIS 2022, LNCS 13591, pp. 1–17, 2022.
https://doi.org/10.1007/978-3-031-17834-4_14

owner. Although enforcing data sovereignty on an individual level is still a long-term goal, there has been considerable progress in the corporate instance. For instance, the International Data Spaces Association (IDSA) has delivered comprehensive organizational guidelines defining the business roles and responsibilities involved in IDS-based ecosystems [1]. From a more technical perspective, the GAIA-X project offers further guidance to build, configure and deploy the infrastructure necessary to ensure secure data transfer among companies, including specialized recommendations for cloud and network service providers to operate in an IDS-based ecosystem [2]. These representative organizations are therefore promoting research in technology to help companies and individuals (in the longer term) enforce their data sovereignty rights.

There is a "space" to enforce data sovereignty called *data space*. Initially referred to as *Industrial Data Space*, the term was recently updated to *International Data Space* (IDS) to reflect the vision of building data-sharing ecosystems crossing national boundaries [3]. According to Otto and Jarke [4], *the IDS initiative is a joint effort of various international research institutes and industrial enterprises to establish a decentralized platform for secure and trusted data sharing*. Braud et al. [1] also state that an *IDS aims to allow the building of data-driven ecosystems in which independent partners (from different sizes, ecosystems, and financial power) trust how external parties handle their data while allowing the innovative data services to be constructed cooperatively* [1]. Hence, there seems to be an association between data sovereignty and trust, which can manifest in IDS as a result of long-term successful cooperation (i.e., proven trust) or acquired competency (i.e., enforced trust). Moreover, a business ecosystem supported by organizational and technical guidelines of IDS could be called an IDS-based business ecosystem.

Proven trust takes time, and there is an urgency to attract companies to join IDS-based ecosystems. The IDS Rule Book [5] and the IDSA Reference Architecture Model (IDS RAM) [6] provide the initial guidance for that purpose. While the former describes the business roles and responsibilities of the parties willing to cooperate in an IDS-based business ecosystem, the latter specifies the technical capabilities of software components necessary to implement an IDS application to support the ecosystem. An essential software component described in both documents is the *data connector* – a software application composed of data transformation applications that can automatically enforce a company's data sovereignty requirements expressed in a machine-readable data policy [6, 7]. IDSA has paid considerable effort to deliver different types of data connectors for testing [7, 8], and there is an optimistic expectation that private companies will start offering their solutions shortly.

Although necessary, data connectors are not sufficient to implement an IDS-based application. Implementing a complete IDS ecosystem, with its organizational roles and technical mechanisms, is somewhat complex and not yet economically attractive, especially for small and medium enterprises (SMEs). Reinhold Achatz, the chairman of the IDSA board, has recently issued a call for business cases and applications to demonstrate the organizational and technical feasibility of building IDS-based business ecosystems in Europe [9]. However, in earlier work, representatives from the Dutch Logistics sector and Enterprise Integration software companies have reported that implementing even elementary viewpoints of the IDS RAM model could be considerably challenging both in the organizational and technical effort. Therefore, a possible direction to address the IDSA call for arms is to show the feasibility of the IDSA vision by parts, starting from the core business roles and software components of an IDS-based ecosystem and progressing as companies and IDS technologies become mature [10].

According to the IDSA Rule Book [5], the core business roles of an IDS-based ecosystem include *data owners*, *users*, and *broker service providers*. The latter has the responsibility to help the former to find the data resources and the suitable data connectors to send and receive data from and to the IDS ecosystem [7]. Hence, the research problem addressed in this paper is *to design an application to support the broker service provider's role in an International Data Space*. The research methodology adopted to approach this question is Design Science, as structured by Wieringa [11]. Thence, the main practical research question splits into (1) a *knowledge question* about *what architectures currently guide the design of a broker service provider application*; and (2) a *technical question* about *how a company could implement such an application*. The motivation to treat this problem is threefold. First, it is an effort to motivate companies to join IDS-based ecosystems and invest in its supporting technology. Second, it will demonstrate technological feasibility and maturity. Last, it will help identify gaps for future development to pave the IDS vision.

The execution of this research follows the Design Cycle proposed by Wieringa [11], which defines three main phases: (1) problem investigation, (2) treatment design, and (3) design validation. The first phase elicited research questions and requirements of relevance for companies and organizations interested in the development of IDS. This research has an active engagement of representatives from the Dutch Logistics sector (e.g., SUTC[1] and EMONS[2]), Enterprise Integration software companies (e.g., eMagiz[3] and CAPE Group[4]), and TNO[5] – the Dutch representative of IDSA. The requirements raised by these organizations specifically concern Enterprise Interoperability and data sovereignty issues that may hinder the adoption of the IDS vision by companies and have been published and partially addressed in earlier work [12]. The second phase focused on designing a reference architecture to guide the implementation of IDS-based ecosystems with a direct application in the Logistics sector. The first architecture model and its components have been introduced in previous work [12, 13], but the current paper opens the black box of broker service provider's infrastructure. Finally, the third phase comprehends the validation of the architecture, which is partially achieved through the application prototype described in detail in this paper.

The continuation of this paper is organized as follows. The next section discusses the role of a broker service provider in IDS, by providing an Enterprise Architecture viewpoint on its internal components and how they relate to the core business roles of an IDS-based ecosystem. Section 3 describes how the combination of Semantic Web and Linked Data technologies leveraged the internal structure of the broker service provider. Section 4 reports on the development of the prototype of a Data Connector Store – an application to help companies discover and select data connectors suitable to enforce

1 https://www.sutc.nl/en_US.
2 https://www.emons.eu/.
3 https://www.emagiz.com/en/en-home/.
4 https://capegroep.nl/en/.
5 https://www.tno.nl/en/.

their data sovereignty needs. A discussion about how this work advances companion research follows in Sect. 5. Last, the paper closes with a summary of contributions, threats to validity, and research outlook.

2 The Role of a Broker Service Provider in IDS: An Enterprise Architecture Viewpoint

According to the IDS Reference Architecture Model (RAM), the *broker service provider* (hereafter referred to as BSP) is an intermediary entity that registers, publishes, and supports the search for metadata about data sources and services available in an IDS ecosystem [6]. A BSP adds value to a data space by providing services to leverage the discoverability of IDS connectors and resources offered by other participants [14]. To ensure the core functionalities of an IDS ecosystem, the International Data Spaces Association (IDSA) prescribes four essential business roles: *core participants*, *intermediary participants*, *software providers*, and *governance bodies* [6]. The BSP belongs to the second group, representing the trusted entities whose business involves managing and providing metadata to the other ecosystem participants.

It is necessary to have at least one BSP operating per business domain (e.g., the Logistics sector). Thus multiple BSPs could simultaneously serve a cross-domain application [6, 14]. According to the IDS RAM [6], a BSP may also assume other business roles, e.g., a *clearinghouse* responsible for keeping logs of all activities related to data exchange in an IDS ecosystem. However, the BSP's responsibilities are somewhat limited to supporting data users and owners with managing the metadata about a particular resource. Therefore, the direct data exchange and usage negotiation processes involving only data users and owners are not part of the responsibilities of a BSP [14].

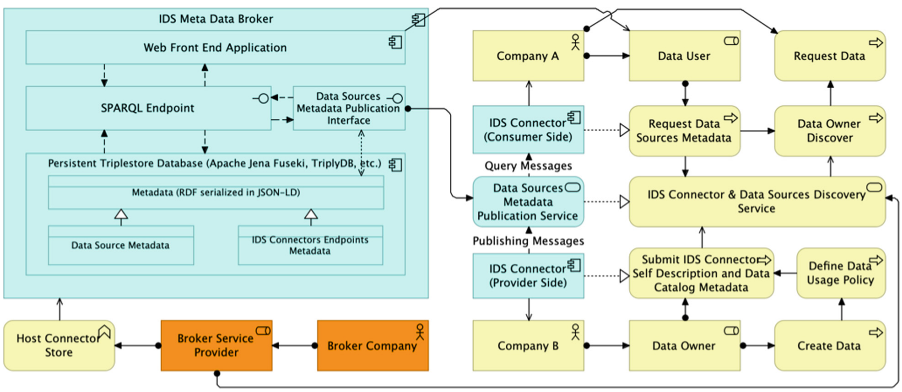

Fig. 1. Enterprise Architecture model of a broker service provider's infrastructure in an International Data Spaces Ecosystem

Figure 1 depicts an Enterprise Architecture model of a BSP, specifying how this entity interacts with other actors and components of an IDS ecosystem through an IDS metadata broker. The architecture conforms to the technical specifications of the IDS RAM and the IDS metadata broker component [6, 15]. The ArchiMate modeling language used to specify the model emphasizes the interplay between concepts in the business and application layer [16, 17]. The BSP appears orange as a business role that develops, hosts, and maintains the IDS metadata broker, which acts as a metadata repository that exposes GUIs and APIs to facilitate metadata publication services.

To carry out its metadata management responsibility, a BSP must provide an interface for data owners to publish their metadata, including descriptions of their *data connectors* and the *data catalogs* accessible through those data connectors. The metadata can be stored in the BSP's internal repository and made available for structured queries submitted by the data user. In addition to supporting the data users in retrieving the metadata of participating IDS connectors or cataloged data resources, the IDS metadata broker should also help the data owners register, activate, update, passivating, or remove metadata entries [15]. Additionally, it should provide an interface describing additional information about its functionalities and indexing services, such as supported query languages, available add-on services, and their data endpoints.

By hosting the IDS metadata broker, the BSP offers its service to the data space to support data users in finding and discovering IDS connectors and data sources provided by the data owners. Two processes must take place for this service to deliver its full potential. First, before enacting any data or metadata exchange, the data owners and users should already be in control of a certified IDS connector. Acquiring the so-called IDS-ready labels for software components is one of the requirements for business actors to participate in an IDS ecosystem after being approved on the organizational level [6, 5]. Secondly, the data owners could submit the self-description and the metadata describing the data used by their data connectors to the IDS metadata broker via the exposed interface based on a standardized protocol (e.g., REST API, OpenAPI 3.0, etc.) [8, 14, 15]. This process occurs after they create the data and define their data usage policies. Next, the data users could discover these catalogs by browsing the IDS metadata broker's metadata based on contextual information (e.g., keywords, language, usage policies, maintainer, etc.). Finally, the data users could receive the information required to access the data owner's IDS connector to request the desired data.

3 Semantic Discovery and Selection of IDS Connectors

The metadata broker specification document states that different metadata broker implementations may be developed and made available by various providers in International Data Spaces [14]. By extending previous research [13], this paper reports on the development of a Data Connector Store, which will operate as an extension of the IDS metadata broker by providing additional functionality to support the semantic discovery and selection of IDS connectors [13]. It aims, therefore, to help data owners and users discover and select the data connectors most suitable for their needs and capabilities based on
AQ3 information about the context in which the connectors could operate (Table 1).

The contextual information describing the IDS connectors derives from a conceptual ontology model proposed in earlier work [13]. The ontology grounded the development of the Data Connector Store proposed in this paper by providing a taxonomy of data connectors and properties that characterize their operational context. Those properties are

Table 1. Data connector's Ontology Requirements Specification Document [13]

Purpose
To describe IDS data connectors for potential participants of an IDS ecosystem
Scope
Contextual information about the business ecosystem where the data connector will operate, e.g., business domain, pricing model, and enforced data access policy
Implementation Language
The ontology is represented in OntoUML, with further translation into OWL
Intended End-Users
User 1.Business representatives of potential and existing IDS participants User 2.IT representatives of potential and current IDS participants User 3.Software and service providers who develop and supply IDS Connectors User 4.Scholars exploring the ontology's knowledge representation capabilities
Intended Uses
Use 1.Software and service providers publish their offered data connectors' metadata on the IDS Connector Store to make their data connectors discoverable Use 2.Business representatives search for IDS-compliant partners operating in the same business domain, complying with common standards, etc Use 3.IT representatives search for data connectors that match their needs and capabilities Use 4.Scholars search and import the ontology into their IDS proof-of-concept tools
Ontology Requirements
Non-Functional Requirements
NFR 1.The ontology must at least use English NFR 2.The ontology must comply, reuse and integrate with the existing IDS Ontology specified under the IDS Information Model
Functional Requirements: Competency Questions
CQ 1.*What software provider offers data connectors?* CQ 2.*Which data connectors are developed for a specific business domain?* CQ 3.*Which data connectors are complying with a particular standard?* CQ 4.*Which data connectors are offered in this pricing model?* CQ 5.*Which data connectors support these data usage agreements?* CQ 6.*Which data connectors were developed in which development framework?* CQ 7.*Which data connectors are offered in this deployment context?* CQ 8.*Which IDS actors use a particular data connector from a specific software provider?* CQ 9.*Which IDS participants operate in a particular business domain?* CQ 10.*Which IDS participants comply with a particular standard?*
Terms from Competency Questions & Frequency
Business Domain, Data usage agreement, deployment, IDS connector, participant, pricing mode, software provider, standards, technology
Objects and Terms for Answers

(*continued*)

Table 1. (*continued*)

Purpose
-Gatewise IDS Connector, Supplydrive IDS Connect-or, TradeCloud IDS Connector; -Transport Logistics, Glass Manufacturing, Steel Manufacturing; -Delete After Interval, Connector-restricted Agreement, Logging Agreement; -Vandaglas B.V., Van Egmond Groep, Meijer Metal; -ECI Software Solutions, Tradecloud, OTM, GS1, EDI4STEEL; -Flat Rate, Freemium, Pay per User, Pay per Feature; -Java, Spring Boot, JavaScript, NodeJS, VueJS, Python, On-Premise, cloud SaaS

defined to answer a list of *ontology competency questions* (CQs) related to the discovery and selection of data connectors and their respective software providers, data owners, and users. By complying with the Ontology Requirements Specification Document (ORSD) detailed in Table 1, the Data Connector Store aims to recommend data connectors that are: (1) developed by a specific service or software provider; (2) developed for a specific business domain; (3) offered in a specific pricing model; or (4) developed using a particular technology.

Figure 2 depicts an alternative ArchiMate viewpoint detailing the internal infrastructure of the Data Connector Store. It exposes the IDS connector *provisioning metadata publication service* through its *provisioning interface*, which extends the *metadata publication service* defined primarily as a standard to implement a metadata broker. This service enables software and service providers to register, update, passivate, and delete their metadata entries and the data connectors they offer.

The Data Connector Store combines Linked Data principles and Semantic Web technologies to store and provide metadata to describe data connectors and their providers. It also aims to facilitate the integration of disparate open data sources in a standardized way [18]. This design decision also relies on the IDSA technical specifications, which indicate that an IDS metadata broker should allow the discovery of data and other resources based on Linked Data principles [6, 14]. Therefore, the Data Connector Store uses the Resource Description Framework (RDF) format to describe the metadata of the data connectors and data sources by annotating them with a layer of semantics to form subject-predicate-object triples [19]. Examples of relevant triples could include: *"Company A uses data connector X"; "software provider B develops data connector Y"; "data connector X is specialized in the Transport Logistics sector"*; or *"data connector Y is offered in a flat-rate pricing model"*. These knowledge representation triples could therefore support the IDS actors in discovering resources of interest based on semi-automated machine reasoning.

To store these metadata represented in RDF, the IDSA suggests the use of a triple store database (e.g., Apache Jena Fuseki, TriplyDB, etc.) or any other storage back end that fits the purpose [2, 11]. The Data Connector Store also needs to provide a SPARQL endpoint to allow data owners and users to accept and send messages that comply with the IDS Information Model (IM), as well as execute the metadata operation queries [9, 12, 13]. These IDS IM compliant messages refer to the publish message that pushes metadata into the repository and the query message that pulls metadata from it [2].

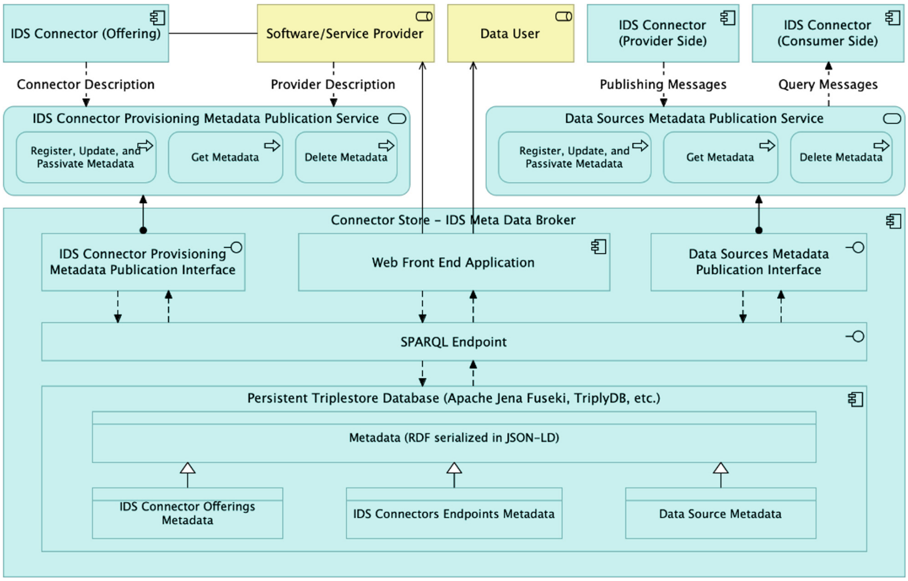

Fig. 2. Enterprise architecture infrastructure viewpoint of the Data Connector Store

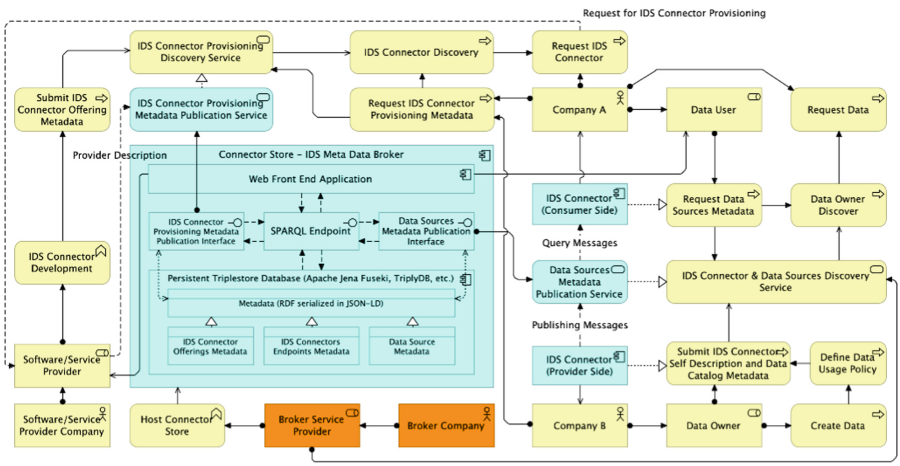

Fig. 3. Connector store ecosystem interaction viewpoint

The Data Connector Store supports the ecosystem operation illustrated in Fig. 3, according to the architecture previously depicted in Fig. 2. It helps software and service providers submit the metadata describing the data connectors they develop and offer in this ecosystem. Accordingly, the Data Connector Store allows data owners and users to find and acquire the best fit data connectors based on their contextual information and through its IDS connector provisioning service (in addition to the *data sources metadata publication service* for the essential IDS metadata broker). Its mechanism enhances the process of discovering and selecting data connectors for the participants, enabling them to quickly onboard to the IDS ecosystem.

4 IDS Data Connector Store: A Proof-of-Concept Application

This section elaborates on the prototype of the Data Connector Store, which demonstrates the technical feasibility of the architecture introduced in Sects. 2 and 3. The Data Connector Store combines multiple application components and application interfaces. It comprises a front-end web application, APIs, and a triple store database, as illustrated in Figs. 2 and 3. The front-end web application operates as a GUI for the participants to browse for the metadata of the required resources. Meanwhile, the APIs are responsible for exposing the IDS connector *provisioning metadata publication service* and the *data sources metadata publication service*. The triple store database is responsible for persisting the metadata represented in RDF triples. A SPARQL endpoint runs on top of the triple store database to allow the front-end application and the APIs to execute the queries for retrieving, inserting, and updating metadata entries.

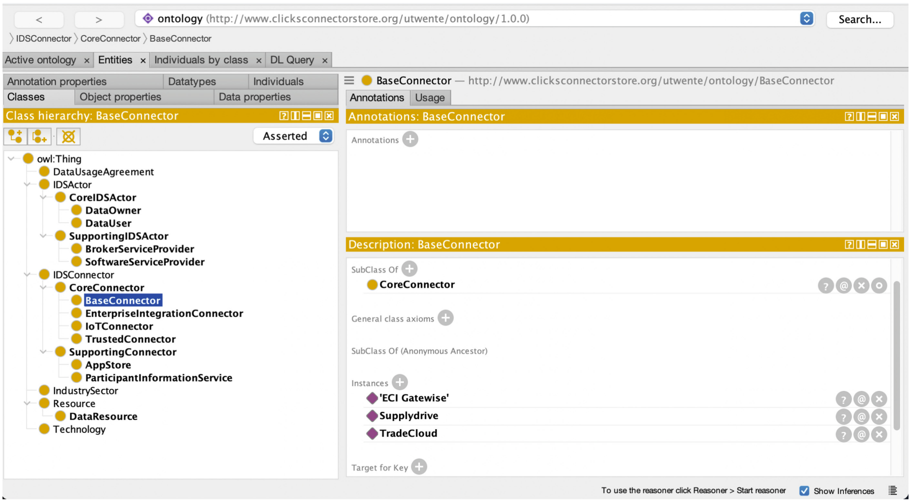

Fig. 4. Visualization of the axioms of the IDS Connector Ontology in the Protégé tool

The first phase of the prototyping process comprised translating the conceptual ontology model proposed in earlier work [13] into an OWL serialization[6], according to the NEON Ontology Engineering methodology [20]. This ontology serves as a knowledge base for the triple store database to structure and store the metadata instances. The Protégé tool [21] supported the design of the OWL ontology and the verification of its axioms. Accordingly, ontology individuals were created manually by referring to the

6 https://raw.githubusercontent.com/danniarreza/connectorstoreontology/main/connectorstorev7.owl.

objects and terms listed in the earlier work's Ontology Requirements Specification Document (ORSD) to reduce the development process complexity [13]. Figure 4 depicts part of the operational ontology design with the Protégé tool. Figure 5 renders a graphical visualization of the ontology with the OntoGraf Protégé plugin.

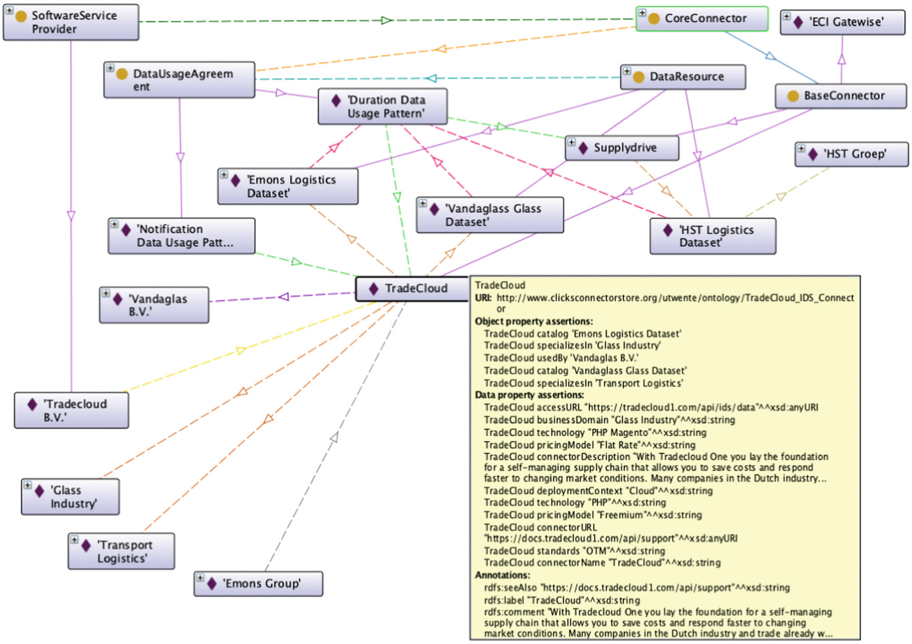

Fig. 5. Visualization of individuals of the IDS Connector Ontology in the OntoGraf plugin

In the second phase of the prototyping process, the OWL model of the IDS connector ontology was uploaded to a triple store database and made available for further querying. This research project uses a triple store associated with Platform Linked Data Netherlands (PLDN), an instance of TriplyDB [22, 23]. The triple store database accepts an ontology graph represented in the Turtle format or its equivalent (i.e., N-triples, JSON-LD, or CSV – except the default OWL or RDF/XML formats), which allows exporting the output ontology to the target format. After that, queries were formulated to retrieve metadata describing the data connectors offered by software and service providers and metadata relating to data sources provided by the data owners. Figure 6 illustrates two SPARQL queries formulated to obtain a list of data resources and descriptions of a data connector. Other SPARQL queries formulated to provide metadata for the prototype of the Data Connector Store are publicly available for external scrutiny[7].

7 https://data.pldn.nl/danniar/-/queries.

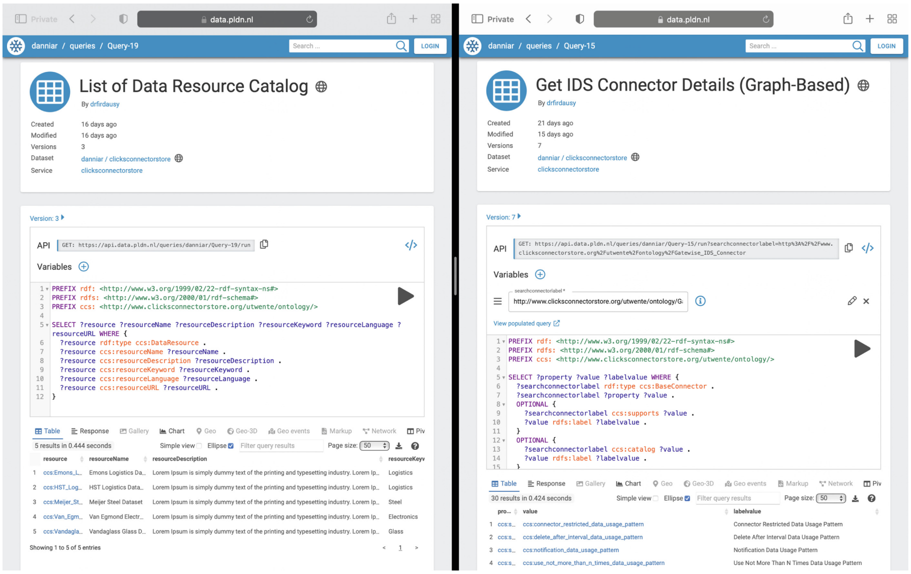

Fig. 6. SPARQL endpoint executing queries to retrieve a list of data resources and connectors

The third phase of the prototyping process focused on developing the front-end part of the Data Connector Store for interaction with the participants of an IDS ecosystem. In this research, the Web application was developed using Mendix. This low-code application development platform allows rapid development and provides an extensive system integration capability, such as via REST API [24]. Figure 7 indicates two distinct processes prescribed for the participants, as indicated in Fig. 3. The upper part of the figure shows how the Data Connector Store provides the participants with a list of data connectors and their respective descriptions. The conceptual ontology proposed in earlier work characterizes the data connectors according to their industry sector, pricing model, data catalog, supported standards, underlying technology, deployment context, and data usage policies [13]. These metadata serve as the filtering attributes to request the data connectors. The bottom part of Fig. 7 indicates a list of data sources and the details of a particular data source owned and offered by a data owner through a respective data connector. Additional metadata are also made available to further describe data sources, for instance, the usage policies constraining the data usage, data representation language, and keywords related to the data content.

Additionally, some properties describing the data connectors and sources were made available as hypertext reference (href) links. This decision aimed to maximize the advantage of following the Linked Data principles, which identifies subjects and objects with HTTP Uniform Resource Identifiers (URIs) and enables associating with one another through their URIs to leverage resource discoverability at the users' side [19, 25]. In the example illustrated in Fig. 8, when the users click on the "offered by" and "industry sector" properties from Fig. 7, they will be supplied with the details of the corresponding software and service provider and the industry sector.

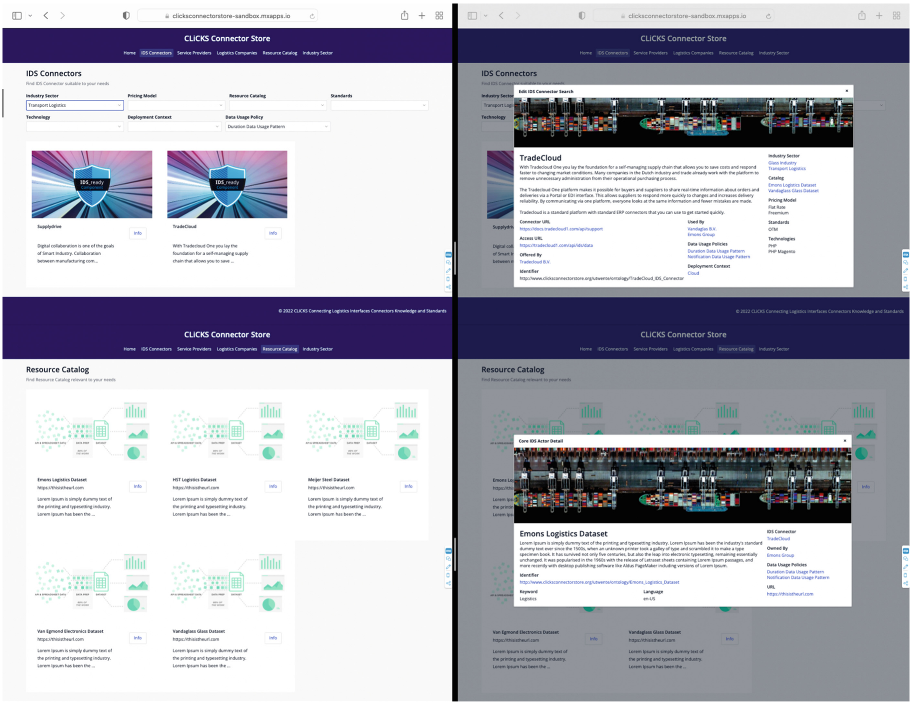

Fig. 7. Data connector store web application – request for a data connector's provisioning metadata and data sources metadata

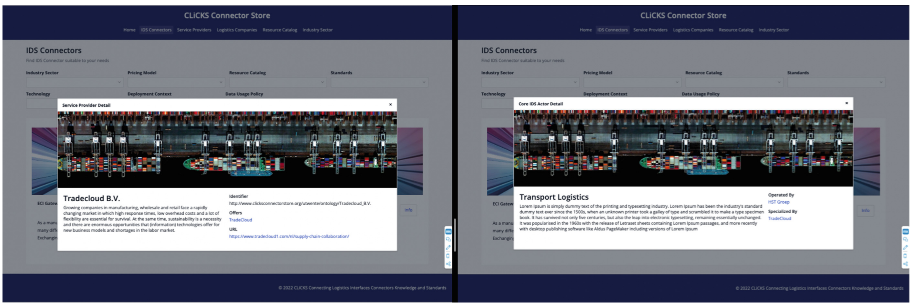

Fig. 8. Connector store web application – software and service provider and Industry sector metadata

Finally, the APIs to facilitate metadata publication services are also implemented and made available. Figure 9 depicts the documentation of the Data Connector Store's preliminary implementation of its REST APIs. Currently, three endpoints are exposed. The first one reveals the metadata broker's self-description, in which its fields refer to the attributes used by the IDS metadata broker reference implementation [8, 15]. Meanwhile, the second and the third endpoints publish the metadata that describes a particular data connector and data source, respectively. The front-end user interfacing

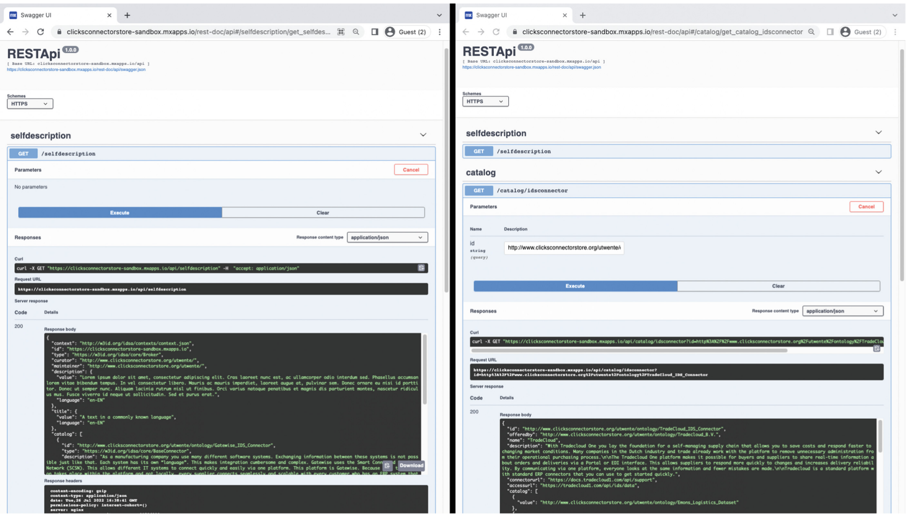

Fig. 9. Connector store web application – metadata broker self-description, IDS connector provisioning, and data sources metadata publication interface

part of the Data Connector Store's prototype and REST API documentation is publicly available for evaluation and testing[8,9,10].

5 Related Work

There are at least three companion approaches closely related to the research reported in this paper: (1) the Mobility Data Space (MobiDS) introduced by Drees et al. [26]; (2) the Smart Connected Supplier Network (SCSN) introduced by Stolwijk, Berkers [27]; and (3) the Maritime Data Space (MDS) introduced by Rødseth, Berre [28].

Drees et al. [26] proposed the MobiDS as an initiative to realize a trustworthy data-sharing ecosystem dedicated to the mobility sector, one of the nine critical sectors referred by the European Commission with a high demand for IDS technology in Europe. To foster discoverability, accessibility, and trustworthiness traits in IDS, the authors describe the implementation of a metadata broker that publishes and allows data users to search for data sources and services. The authors also propose an architecture in which they specify the interactions of the metadata broker with the other intermediary roles in an IDS ecosystem (i.e., clearinghouse, vocabulary provider, identity provider, and data apps provider). Despite the extensive description of the demonstrator (including prototypes of the intermediary services), the architecture is limited to a high-level organizational viewpoint. It does not elaborate on what kind of technologies could support the technical

8 https://clicksconnectorstore-sandbox.mxapps.io.
9 https://clicksconnectorstore-sandbox.mxapps.io/rest-doc/api.
10 https://github.com/danniarreza/connectorstoreontology.

implementation of its components or how to use them, which may limit its acceptance by business practitioners, developers, and scholars in exploring particular business cases. However, the architecture presented in this paper characterizes the technological components individually and how they interact with one another in alignment with the upper layer's business processes and roles. Still, the authors of MobiDS report the evaluation of its usability by highlighting use cases on a granular level (i.e., intermodal mobility, on-demand rural transport, first/last mile, and traffic management). The evaluation of the prototype's usability introduced in this paper is part of future work.

Stolwijk, Berkers [27] introduced the concept of a Smart Connected Supply Network (SCSN) as a data-sharing ecosystem for the manufacturing domain where multiple service providers are fully interconnected using IDS technology. It allows a manufacturing company to interoperate with other companies reachable in the network once they connect with one of the service providers. In an SCSN ecosystem, the service providers become the central components enacting a data space by providing manufacturing companies with IDS connectors whereby companies could exchange data. Although the authors present a high-level architecture specifying the relationships between manufacturing companies, software components, and service providers, they provide no architectural viewpoints to clarify how to realize the SCSN from a technical perspective. Their work focuses on bringing forward the projected benefits and savings learned from implementing the SCSN standard to facilitate interoperability between manufacturing companies through a simulation of three different ICT service providers.

Lastly, the Maritime Data Space (MDS) proposed by Rødseth, Berre [28] comprehends an IDS-based Maritime ecosystem aiming to tackle data ownership, access rights, and interoperability issues originating from the existing decentralized physical systems architecture and data storage. The authors propose a conceptual model for MDS, which explains its active business roles, business services, datasets, and metadata. Yet, similar to the two related approaches, their conceptual model falls short of elaborating the technical aspects of developing and implementing the proposed data space. Such a limitation may hinder technology adoption by companies interested in customizing it in their business cases.

6 Conclusion and Future Research

This paper addressed the research problem of designing an application to support the broker service provider's role in an International Data Space. Focusing on developing technology to help this role is a crucial enabler of the simplest form of an IDS-based business ecosystem, which still includes data owners' and users' roles. This work also accomplishes part of a Design Science cycle [11] started in recent work [12, 13]. The problem investigation phase combined requirements from the IDSA technical specifications [6–8] with Enterprise Interoperability and data sovereignty requirements elicited from representatives of the Dutch Logistics sector and Enterprise Integration software companies interested in joining IDS-based ecosystems shortly [12]. The primary outcome of the treatment design phase was a reference architecture to guide the implementation of IDS-based applications initially evaluated in [12] and extended in this paper with a white-box view of the components of the broker service provider. Moreover, the treatment design combined IDSA technical specifications with Semantic Web and Linked

Data technologies. The Data Connector Store represents a proof-of-concept application that partially validates the utility of the Enterprise Architecture model also described in this paper.

The main contributions of this work are threefold. First, it demonstrates the *technical feasibility* of implementing a broker service provider. As the state-of-the-art technology evolves with companies providing their data connectors and further advances in IDS infrastructure, companies, especially SMEs, are expected to become more motivated to join IDS-based ecosystems and invest in their underlying technology. Second, although not yet evaluated by end-users, the prototype of the Data Connector Store could *facilitate* the discovery and selection of data connectors for companies interested in exploring IDS ecosystems. Moreover, the architecture is customizable, including other ontologies to describe new types of data connectors and catalogs. Last, the Data Connector Store innovates the IDSA technical guidelines by combining Semantic Web and Linked Data technologies to integrate IDS technology already available for testing.

There are also some limitations threatening the validity of this work. The first one concerns the correctness of the architecture. Even though the development of this work is based on the reference architecture model and technical guidelines provided by the IDSA, a proper evaluation by an Evaluation Facility appointed by the IDSA will further guarantee its compliance with IDS specifications. The second one concerns the lack of evaluation of the prototype with end-users. Although it had a positive preliminary assessment of some of the representatives of the Dutch Logistics sector involved in this research, a direct examination by companies interested in joining an IDS-based ecosystem is necessary. That leads to the third limitation, which calls for a business case to demonstrate the prototype's utility and usability. The IDSA call for business cases is urgent [9], but companies are still skeptical about the practical benefits of investing in IDS technology. On top of this, business cases will test the prototype's performance and scalability.

There are three immediate research directions to explore from this work. The first one will comprehend the extension of the Enterprise Architecture proposed in this work with the clearinghouse role, which is essential to ensure consistency of the data transactions in an IDS-based ecosystem. Besides, it is critical to show its technical feasibility with a prototype. According to the IDSA Reference Architecture Model, a broker service provider could accumulate the clearinghouse function in a federated IDS-based ecosystem. There are also indications that the Enterprise Integration software companies involved in this research may explore their private control towers as a mechanism to realize the functionalities of a clearinghouse in IDS. The second direction is extending the Semantic Web and Linked Data functionalities used in this work to leverage the discoverability of the IDS data connectors' data applications. This work assumed that these applications were preconfigured and ensembled in a data connector for a while. However, it should be possible for companies to reuse their data transformation applications to build customizable data connectors to operate in diverse business domains. Finally, a business case involving the exchange of sensitive data among Dutch Transport Logistics partners could provide more robust evidence for the acceptance and adoption of IDS technology, at least for regional ecosystems.

Acknowledgments. This research is financially supported by the Dutch Ministry of Economic Affairs and co-financed via TKI DINALOG and NWO. The CLICKS project has granted funding for this work (grant no. 439.19.633). CLICKS is the acronym for *Connecting Logistics Interfaces, Converters, Knowledge, and Standards*. The authors thank the involved consortium partners for their support and the anonymous reviewers for their constructive feedback.

References

1. Braud, A., Fromentoux, G., Radier, B., Le Grand, O.: The road to European digital sovereignty with Gaia-X and IDSA. IEEE Netw. **35**(2), 4–5 (2021)
2. Autolitano, S., Pawlowska, A.: Europe's quest for digital sovereignty: GAIA-X as a case study. IAI Papers **21**(14), 1–22 (2021)
3. Bader, S., et al.: The international data spaces information model–an ontology for sovereign exchange of digital content. In: Bader, S., et al. (eds.) The Semantic Web – ISWC 2020. ISWC 2020. Lecture Notes in Computer Science, vol. 12507, pp. 176–192. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62466-8_12
4. Otto, B., Jarke, M.: Designing a multi-sided data platform: findings from the International Data Spaces case. Electron. Mark. **29**(4), 561–580 (2019). https://doi.org/10.1007/s12525-019-00362-x
5. IDSA: IDSA rule book version 1.0. In: Steinbuss, S. (ed.) International Data Spaces Association, Berlin (2020)
6. IDSA: IDS reference architecture model version 3.0. In: Steinbuss, S. (ed.) International Data Spaces Association, Berlin (2019)
7. IDSA: IDS Metadata Broker (2022). https://github.com/International-Data-Spaces-Association/metadata-broker-open-core. Accessed 11 June 2022
8. IDSA: IDS Metadata Broker API Reference Implementation (2020). https://app.swaggerhub.com/apis/idsa/IDS-Broker/1.3.1. Accessed 10 June 2022
9. Otto, B.: Interview with Reinhold Achatz on "data sovereignty and data ecosystems." Bus. Inf. Syst. Eng. **61**(5), 635–636 (2019)
10. Gort, E.: Developing a maturity model-based approach supporting the decision to adopt International Data Spaces. Master's thesis, University of Twente (2021)
11. Wieringa, R.J.: Design Science Methodology for Information Systems and Software Engineering. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43839-8
12. Firdausy, D.R., Silva, P.D.A., van Sinderen, M., Iacob, M.-E.: Towards a reference enterprise architecture to enforce digital sovereignty in international data spaces. Paper Presented at the IEEE 24th Conference on Business Informatics (CBI), Amsterdam, The Netherlands (2022)
13. Firdausy, D., Silva, P.D.A., van Sinderen, M.J., Iacob, M.E.: Semantic discovery and selection of data connectors in international data spaces. In: Interoperability for Enterprise Systems and Applications, I-ESA 2022 (2022)
14. Bader, S., et al.: Specification: IDS meta data broker. In: Steinbuss, S. (ed.) IDSA (2020)
15. IDSA Certification Working Group: Component certification criteria catalog V2.1.2. In: International Data Spaces Association (2021)
16. Josey, A.: ArchiMate® 3.0. 1-A pocket guide. Van Haren (2017)
17. Iacob, M.E., Meertens, L.O., Jonkers, H., Quartel, D.A.C., Nieuwenhuis, L.J.M., van Sinderen, M.J.: From enterprise architecture to business models and back. Softw. Syst. Model. **13**(3), 1059–1083 (2012). https://doi.org/10.1007/s10270-012-0304-6
18. Soylu, A., et al.: Enhancing public procurement in the european union through constructing and exploiting an integrated knowledge graph. In: Pan, J.Z., et al. (eds.) ISWC 2020. LNCS, vol. 12507, pp. 430–446. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62466-8_27

19. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Sci. Am. **284**(5), 34–43 (2001)
20. Suárez-Figueroa, M.C., Gómez-Pérez, A., Fernández-López, M.: The NeOn methodology for ontology engineering. In: Suárez-Figueroa, M., Gómez-Pérez, A., Motta, E., Gangemi, A. (eds.) Ontology Engineering in a Networked World, pp. 9–34. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-24794-1_2
21. Sivakumar, R., Arivoli, P.: Ontology visualization PROTÉGÉ tools–a review. Int. J. Adv. Inf. Technol. (IJAIT) **1** (2011)
22. Kadaster B.V.: Platform Linked Data Nederland (PLDN). https://data.pldn.nl/
23. TriplyDB (2022). https://triply.cc/. Accessed 14 Apr 2022
24. Henkel, M., Stirna, J.: Pondering on the key functionality of model-driven development tools: the case of mendix. In: Forbrig, P., Günther, H. (eds.) Perspectives in Business Informatics Research. BIR 2010. Lecture Notes in Business Information Processing. Springer, Heidelberg, vol. 64, pp. 146–160 (2010). https://doi.org/10.1007/978-3-642-16101-8_12
25. Berners-Lee, T.: Linked Data (2006). https://www.w3.org/DesignIssues/LinkedData.html. Accessed 15 June 2022
26. Drees, H., Kubitza, D.O., Lipp, J., Pretzsch, S., Langdon, C.S.: Mobility data space–first implementation and business opportunities. In: ITS World Congress (2021)
27. Stolwijk, C., Berkers, F.: Scalability and agility of the smart connected supplier network approach. In: TNO (2020)
28. Rødseth, Ø.J., Berre, A.J.: From digital twin to maritime data space: transparent ownership and use of ship information. In: Proceedings of 13th International Symposium on Integrated Ship's Information Systems & Marine Traffic Engineering Conference (ISIS-MTE), pp. 1–8 (2018)

Author Queries

Chapter 14

Query Refs.	**Details Required**	**Author's response**
AQ1	Please check and confirm if the authors given and family names have been correctly identified.	
AQ2	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ3	Please check and confirm if the inserted citation of Table 1 is correct. If not, please suggest an alternate citation.	