

# Chapter 14

## IDS as a Foundation for Open Data Ecosystems



Fabian Kirstein and Vincent Bohlen

**Abstract** Open data is a popular and flourishing concept. The availability of open and structured data is the foundation of new business models, citizen engagement, and scientific research. However, open data still faces many issues to unfold its full potential, including usability, quality, legal, privacy, strategic, and technical barriers. In addition, the public sector remains its main provider, while industry stakeholders are still reluctant to participate in open data ecosystems. In this article, we present an architecture to overcome these drawbacks by utilizing the concepts, specifications, and technologies provided by International Data Spaces. We developed a prototype to demonstrate and evaluate the practical adoption of our architecture. Our work shows that IDS can act a vital foundation for open data ecosystems. The presented solution is available as open source software.

### 14.1 Introduction

Open data is a popular and continuously evolving concept, focusing on the open provision and reuse of structured datasets. Established publishers are public bodies, such as public administrations, research institutes, and nonprofit organizations. In addition, the relevance for private companies and industries as publishers of open data is increasing. Typical domains are traffic, weather, geographical, statistical data, and research data [1]. Open data creates transparency, supports innovations, and contributes to participation processes. The data should meet a set of key characteristics; in particular it should be complete, as raw as possible, up-to-date, accessible, free of charge, and machine-readable. Open data is mostly made available via web

---

F. Kirstein (✉)  
Fraunhofer FOKUS, Berlin, Germany

Weizenbaum Institute for the Networked Society, Berlin, Germany  
e-mail: [fabian.kirstein@fokus.fraunhofer.de](mailto:fabian.kirstein@fokus.fraunhofer.de)

V. Bohlen  
Fraunhofer FOKUS, Berlin, Germany  
e-mail: [vincent.bohlen@fokus.fraunhofer.de](mailto:vincent.bohlen@fokus.fraunhofer.de)

portals, provided by data publishers [2]. Currently more than 2600 individual portals exist worldwide [3]. An integral element of open data is the aggregation of available data into centralized platforms to create single access points and harmonized views. An example for open government data is the European Data Portal (EDP), aiming to make available public sector information from across Europe [4]. A similar platform is operated by the European project OpenAIRE, gathering research data and publications from a plethora of distributed repositories [5]. However, open data is still facing numerous barriers on technical, legal, organizational, strategic, and usability levels [6]. In this article, we argue that the technologies, specifications, and methodologies of the International Data Spaces (IDS) can act as a powerful foundation to build and maintain open data ecosystems and improve and remedy existing challenges and issues. The core objective of our work is to lift open data into a unified and common data space. Therefore, we developed an IDS-based architecture, allowing the publication and dissemination of open data in a decentralized and timely fashion. We evaluated our work through a working prototype, thus demonstrating its practical feasibility. Finally, we discuss our findings and indicate directions for future developments. The main contributions of our work are:

1. An IDS-based software architecture and toolkit to build and operate open data ecosystems and to foster the adoption of open data in industry-driven scenarios.
2. A new paradigm for disseminating open data according to its often decentralized origins and provenance, facilitating more timely and sovereign access.

## 14.2 Barriers of Open Data

Barriers and issues in open data are well researched and constitute the principal motivation of our work. Beno et al. [6] collected extensive data about these barriers based on existing literature and an online survey. The authors distinguish between the users and the publishers of open data. The most relevant issues for users are the limited availability and functionality of machine-readable representations and interfaces for the data, outdated and obsolete data, and incomplete metadata. In addition, users reported challenges in searching and browsing, poor performance of the portals, and lacking information on data quality. Publishers are more concerned with legal aspects, ownership of the data, and potential loss of control. In addition, they seek resource-efficient methods to publish their data. In general, low data quality, the lack of a definite standard for describing metadata, and poor interoperability between data providers impede open data to meet its full potential [7].

A crucial reason for the current state of open data is the strong fragmentation and decentralization of data providers, although a single and central point of access and a unified data space are desired. This leads to a plethora of individual solutions, various implementations, and differing publication schemes. The current mitigation strategy is to set up centralized platforms to aggregate and unify open data, such as through Google Dataset Search [8]. However, these services may lead to undesired vendor lock-in and loss of data sovereignty.

### 14.3 Related Work

The technical implementation of open data is highly related to standardized data formats, open-source and propriety platform solutions, Semantic Web standards, and decentralization approaches.

An important foundation is the Data Catalog Vocabulary (DCAT). DCAT aims to describe meta-datasets meaningfully and seeks to increase interoperability between data catalogs. It is based on the Resource Description Framework (RDF) and is agnostic with regard to the actual data. Its core comprises three main classes: catalog, dataset, and distribution. One catalog consists of multiple datasets, and a dataset is referencing to one or more distributions, constituting a representation of the actual data [9]. The DCAT Application Profile for data portals in Europe (DCAT-AP) is an extension of DCAT, designed to describe public sector datasets. It is designed as a unifying standard for the publication of open data in Europe and extends DCAT with additional properties and mandatory controlled vocabularies. For instance, properties like language, spatial information, or file format can be aligned with pre-defined and centrally managed vocabularies that are formalized with the Simple Knowledge Organization System (SKOS) [10]. DCAT-AP is widely adopted in the European Open Data landscape with country-specific extensions emerging as well.

Many solutions exist for publishing open data on the Web and for making it accessible for both humans and machines. The open-source solution CKAN has grown to be the de facto standard in the public sector for creating open data catalogs, but is increasingly adopted by private companies too. CKAN offers many functionalities for managing the entire publication process and offers a wide range of plugins. Its API is well documented and provides an extensive way to access the data programmatically [11]. Another open source solution is uData, which puts a focus on social interaction and customization. It offers specific features for data reuse and community contributions, distinguishing it from other solutions [12]. A proprietary and closed source solution for creating open data platforms is OpenDataSoft. It is widely adopted and emphasizes interactive feature, for example, visualizations and the generation of extensive APIs for structured data [13].

With the increasing relevance and dissemination of open data throughout the Web, aggregation and federation services are becoming more crucial. Therefore, there is an intrinsic interest of the open data movement to implement methods and systems to deliver a unified and central point of access. Besides national and pan-national portals, like the EDP, the Google Dataset Search aims to make all open data, published worldwide on the Web, available in a central place. It employs structured metadata, embedded on web sites that are publishing open data. The Google search index is used to extract this metadata, process it, and aggregate it into a convenient, harmonized view [8]. Yet, it remains proprietary and does neither offer a machine-readable interface nor transparency for data publishers.

Besides straightforward harvesting approaches of the datasets, approaches that are more elaborate have evolved. One popular interface standard applied in open data is SPARQL, offering the possibility for federated queries. This allows the

retrieval of data from multiple endpoints simultaneously. Many other implementations and methods are available, but the solutions are not yet ready for production use. Rakhmawati et al. [14] compiled a comprehensive overview of existing research in this field. More recent research was conducted in the combination of peer-to-peer mechanisms and open data, especially regarding the rise of blockchain and distributed ledger technologies. Truong et al. [15] proposed a solution based on Hyperledger Fabric and the InterPlanetary FileSystem (IPFS) to improve the provision and integrity of open data. The tool Regerator follows the methodology of CKAN and effectively enables the creation of decentralized registries for open data, by employing smart contracts on the Ethereum blockchain [16]. García-Barriocanal et al. [17] sketched an architecture based on Ethereum, IPFS, and the decentralized database BigchainDB to facilitate a secure and trustworthy metadata repository.

Finally, the Dat project constitutes a vivid initiative, a range of open source tools, and a custom peer-to-peer protocol to handle decentralized publication and archiving of data across multiple organizations. Its core objective is to shift the provision of data from central and commercial services to self-controlled consortium networks. Currently it primarily targets the research community but can be applied for any kind of data [18].

However, these partly experimental decentralization approaches are highly complex. In addition, many of them do not take the existing specifications for open data into account. To the best of our knowledge, no work pursues the middle ground between centralized aggregation and decentralization to publish open data, like the methodology and architecture of IDS can offer. Additionally, existing approaches are not considering any connection to industry data processing and exchange.

## 14.4 International Data Spaces and Open Data

In principle, open data and International Data Spaces address different data sharing problems. While open data aims to create transparency, the central objective of IDS is to establish trust between companies and to ensure the secure handling of data. However, these two approaches are not mutually exclusive. For instance, companies utilizing the IDS to share confidential data could also consume open data to support data analyses or even publish non-confidential data under an open license. While the IDS offers various approaches to ensure data sovereignty of data providers, the underlying objective is to enable participants to exchange data. From a purely technical perspective, exchanging open data is no different from exchanging industry data. Therefore, IDS can be used to distribute both closed and open data.

In the following, we will present our IDS-based open data architecture and illustrate the mutual benefits, arising from this combination.

### ***14.4.1 IDS as an Open Data Technology***

The underlying concepts and technologies of open data and IDS are very similar. Both initiatives rely on metadata repositories to share information about the availability and accessibility of data. Such repositories store knowledge about participating data publishers, available data offers, and possible access restrictions, without the need to transfer actual data into central data hubs. Therefore, both concepts follow principles of decentralization and transfer metadata to and from central information access points, i.e., metadata brokers and open data portals follow the same basic conception. The actual data remains under control of the data publisher's infrastructure until a potential data user issues a request for it. Both open data portals and IDS metadata brokers act as gateways to their respective ecosystems, providing (comprehensive) information about the available data.

As introduced in our related work, open data uses Semantic Web specifications, like DCAT. Likewise, the IDS information model is based on Linked Data principles and DCAT. This makes the two systems easily compatible and ensures straightforward interoperability.

### ***14.4.2 IDS Components in an Open Data Environment***

The architectural similarities of open data and International Data Spaces allow a straightforward matching of artifacts. To build open data ecosystems based on the IDS specifications, the corresponding components and actors must be practically aligned. IDS connectors provide the means to publish and expose data, matching the role of a data provider in the domain of open data. IDS metadata brokers collect and provide the metadata about available data offers, taking the role of open data portals. An IDS open data ecosystem therefore consists of at least one IDS connector and at least one IDS metadata broker:

- **Open Data Connector**

The IDS connector takes the role of an open data provider and becomes an open data connector. Each data publishing entity in an open data ecosystem applies an instance of the connector to announce availability and grant access to data resources. Data consumers request said data from the connector, which then responds by serving the actual data from internal data management systems. In contrast to other IDS connectors, usage policies or access restrictions are not necessary requirements for exchanging data under open licenses. Therefore, the open data connector's usage control features can be reduced to a minimum, allowing for easier configuration and handling.

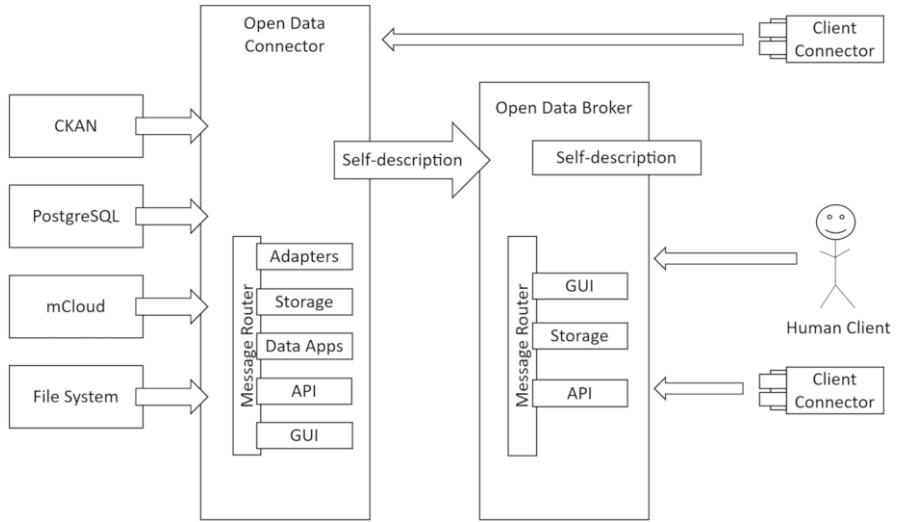


Fig. 14.1 Component overview and (meta)data flow in the IDS open data ecosystem

• **Open Data Broker**

The IDS metadata broker fulfills a very similar role as an open data portal and becomes an open data broker. It represents a central entity, distributing information about what kind of data is available from which participant and which conditions apply for using the data. Data consumers use the metadata acquired from the broker to find and select their desired data and request the data directly from the publishing connector. As in the domain of open data, a broad variety of open data brokers can exist. Municipalities, e.g., on regional, country, or continent level, or private companies may operate their own open data broker. Different metadata brokers can also assemble data offers from specific domains, such as transport or energy.

Connectors and brokers communicate via well-defined IDS communication means, such as IDS messages, defined by the IDS information model, or the IDS Communication Protocol (IDSCP). This mechanism permits the IDS connector to pick the specific IDS metadata brokers to publish the metadata. It also provides the connector with the possibility to revoke the publishing of data while instantly informing the metadata broker about the changes made to the data offer.

An architecture overview of the described components is provided in Fig. 14.1. The figure depicts the (meta)data flow in the open data ecosystem. Additionally, an identity provider can be applied to ensure that the offered data is served from the expected publisher, tackling issues concerning the authenticity and integrity of data.

### 14.4.3 *Benefits*

Open data faces issues with data availability and metadata quality. Especially in open data portals that are harvesting large numbers of data sources, accessibility and overall usability depend on the metadata supplied by the original data providers. Standards like DCAT provide a common ground, but it remains within the responsibility of the data provider to follow these standards. IDS provides stricter specifications, not only covering metadata, but the entire communication process. This allows a much more harmonized communication and improved interoperability. Hence, the number of datasets published with insufficient metadata can be reduced significantly, and unified data exchanges between publishers and portals can be established.

In traditional open data environments, availability of datasets can be a problem if data publishers cannot ensure the accessibility of said data. As a result, open data portals must deal with unavailable data, dead links, or server limitations on the provider's side. To counteract these problems, portals periodically harvest the publisher's data catalogs and perform availability checks to confirm the data is still reachable. In the IDS open data ecosystem, the responsibility to keep the information at the portal/broker up to date is reversed. The open data connector informs the open data broker about currently available or updated datasets. Whenever changes occur to the actual data or the metadata, the broker is informed about the changes and adapts the information published accordingly. The "pull" approach of responsibility in traditional open data environments is inverted to a "push" approach in the IDS open data ecosystem. This approach does not only shift the responsibility of housekeeping the data offers from the portal/broker to the data publisher but presents the data publisher with new possibilities to control the initial placement of said data offers. Traditionally, publishers make their datasets available for download on their infrastructure, and the data is harvested by different open data portals. With IDS, the publisher chooses which metadata broker to register to, giving the data publisher additional data sovereignty. Obviously, this control ends when the data is downloaded—as there are no usage controls applied in the open data domain. For the data consumers, this approach can lead to improvements regarding the findability of timely data and reduces fragmentation. A fully developed IDS open data ecosystem will, thus, act as one virtual data space and removes the need to search for data in different places.

An exemplary comparison of the two dissemination approaches is illustrated in Fig. 14.2, modelling the differences between data flows. While open data and industry data are conceptually distinct in various aspects, the core concept is equally relevant for both domains. Consequently, an alignment of the two types of data on a technological level is highly beneficial. New possibilities for integration and reuse emerge if open data can be acquired, handled, and processed with the same tools and applications that are already applied in industry. The entire process of using, analyzing, and creating value from open data can be made more efficient and lower the entry barriers to using and publishing open data. On the other hand, it

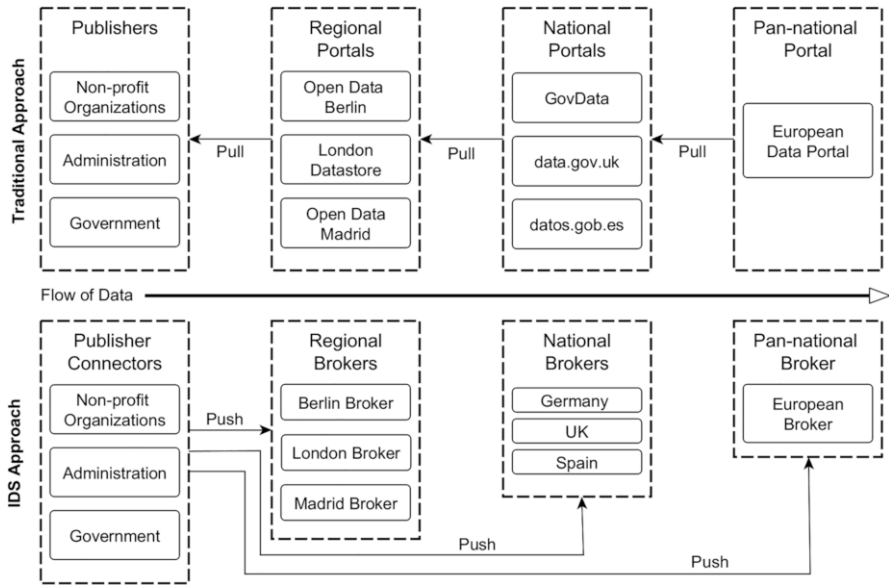


Fig. 14.2 Comparison of different data flows

can lower the barrier for industries to enter the IDS in the first place and act as a catalyst for adoption. Since legal hurdles and complex usage control are negligible when dealing with open data, it can also encourage interested parties to start sharing data via IDS. This may be followed by industry use cases. The following table provides a summary of how IDS can mitigate typical open data barriers:

Open data barrier	IDS solution
Limited availability of machine-readable representations and interfaces	Metadata exchange between data publishers and open data portals via the highly standardized IDS information model
Challenges in searching and finding open data	The virtual data space can lower fragmentation and supports the creation of a single point of truth, thus simplifying the search for suitable data
Potential loss of control	Data sovereignty through the ability to freely choose the brokers for publication and revocation of data offers
Poor interoperability between data providers	Data sources (open or closed) operating under the same standardized IDS conditions. Support for direct industry contributions
Central and timely point of access	The IDS “push” approach allows more control and immediate updates to central access points



## 14.5 The Public Data Space

The Public Data Space is a practical implementation and prototype of the aforementioned IDS open data ecosystem. It consists of an open data connector used to publish offers of open data to the ecosystem and an open data broker to collect and present metadata about available data offers. The Public Data Space successfully establishes a link between IDS and the Web by publicly exposing the open data offers. In addition, it targets human users through a full-fledged open data frontend for convenient browsing, searching, and querying. The components are published under the open source Apache 2.0 license, and the source code is available on GitHub.<sup>1</sup>

The following section will introduce the components in more detail and present an open government data use case as an example.

### 14.5.1 *The Open Data Connector*

The Open Data Connector is an implementation of the IDS connector specification with the aim of providing data owners an easy to use, out-of-the-box solution for publishing their data as open data. The connector is built to seamlessly interact with the Open Data Broker and offers functionality to communicate within the realm of International Data Spaces while simultaneously offering technology-agnostic ways to retrieve the published data.

The Open Data Connector is extensible and able to connect to any standard or proprietary data storage solution via data source adapters. Currently, paradigmatic implementations of adapters for SQL queries to PostgreSQL databases, the CKAN API metadata schema, and the mCLOUD<sup>2</sup> API metadata schema are available. Data is not replicated inside the connector but is retrieved on demand from the registered data sources. API guidelines for the development of data source adapters are provided as an OpenAPI<sup>3</sup> specification and can be easily applied to provide access to the desired data source solution.

An overview of the Connector's internal component architecture is presented in Fig. 14.3. The Open Data Connector offers two ways of communication: (1) it accepts IDS multipart messages and is able to respond to incoming requests for artifacts and self-descriptions and (2) additionally, the connector offers HTTP GET endpoints to retrieve the offered datasets without the technical barrier of having to set up a consuming IDS connector. This design decision was made to reflect the open nature of the offered resources and avoid imposing additional technological barriers on the retrieval of data.

---

<sup>1</sup><https://github.com/public-data-space>

<sup>2</sup><https://www.mcloud.de/>

<sup>3</sup><https://www.openapis.org/>

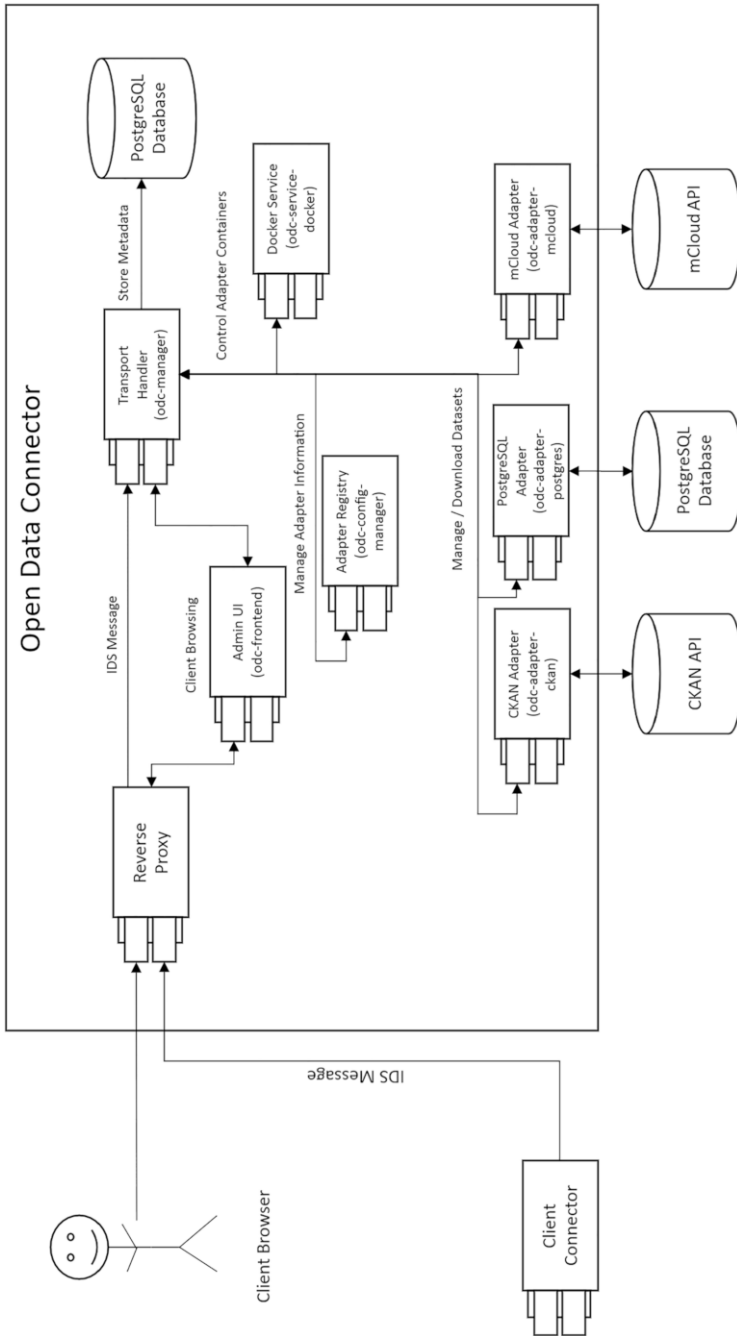


Fig. 14.3 Component architecture of the Open Data Connector

There exist no particular technical differences in consuming open data compared to any other type of data published via IDS. Therefore, the Open Data Connector is not intended to consume data and does not contain such functionalities. Data from Open Data Connectors should be consumed, using the same connectors already set up to retrieve commercial data for a given use case.

### ***14.5.2 The Open Data Broker***

The Open Data Broker is an implementation of the IDS metadata broker specification. In contrast to existing IDS metadata broker implementations, the Open Data Broker is not only offering endpoints for IDS components to consume and navigate the data offerings, but is equipped with the functionality of Open Data portals. Human users can search and browse the registered data using a modern web frontend without the need to deal directly with the IDS interfaces, used in the backend of the portal. Therefore, the HTTP GET endpoints of the Open Data Connector are used. This allows users of the Broker's portal to directly download data from the connectors through the user interface. The Broker's open data portal functionality is based on the open data management platform piveau [19]. Piveau offers comprehensive functionalities to store and search metadata based on Linked Data, as well as a modern portal web frontend.

In addition, the Open Data Broker offers the IDS endpoints specified in the IDS metadata broker specification. Via the */infrastructure* endpoint, the Broker accepts IDS multipart messages and is able to respond to incoming requests to register or unregister a connector or dataset and handles self-description requests. Via the */data* endpoint, the Broker accepts IDS query messages containing SPARQL payloads. Figure 14.4 presents an overview of the Broker's internal component architecture.

The Open Data Broker differs from other IDS metadata brokers in that it provides a web frontend specific to the needs of open data portals and by offering optimized interoperability with the Open Data Connector. Furthermore, the IDS information model was extended with additional metadata fields from the DCAT specification. Both components make use of this additional metadata and allow an improved user experience and more fine-grained information displayed at the Broker.

### ***14.5.3 Use Case: Publishing Open Government Data***

The Public Data Space is employed as a demo showcase and publicly available since December 2020.

The domain of open data is largely defined by an innumerable amount of different data publishers and data portals. The showcase has been set up to replicate this environment and to demonstrate how the Public Data Space solution runs in a production environment. The demo showcase consists of ten Open Data Connectors

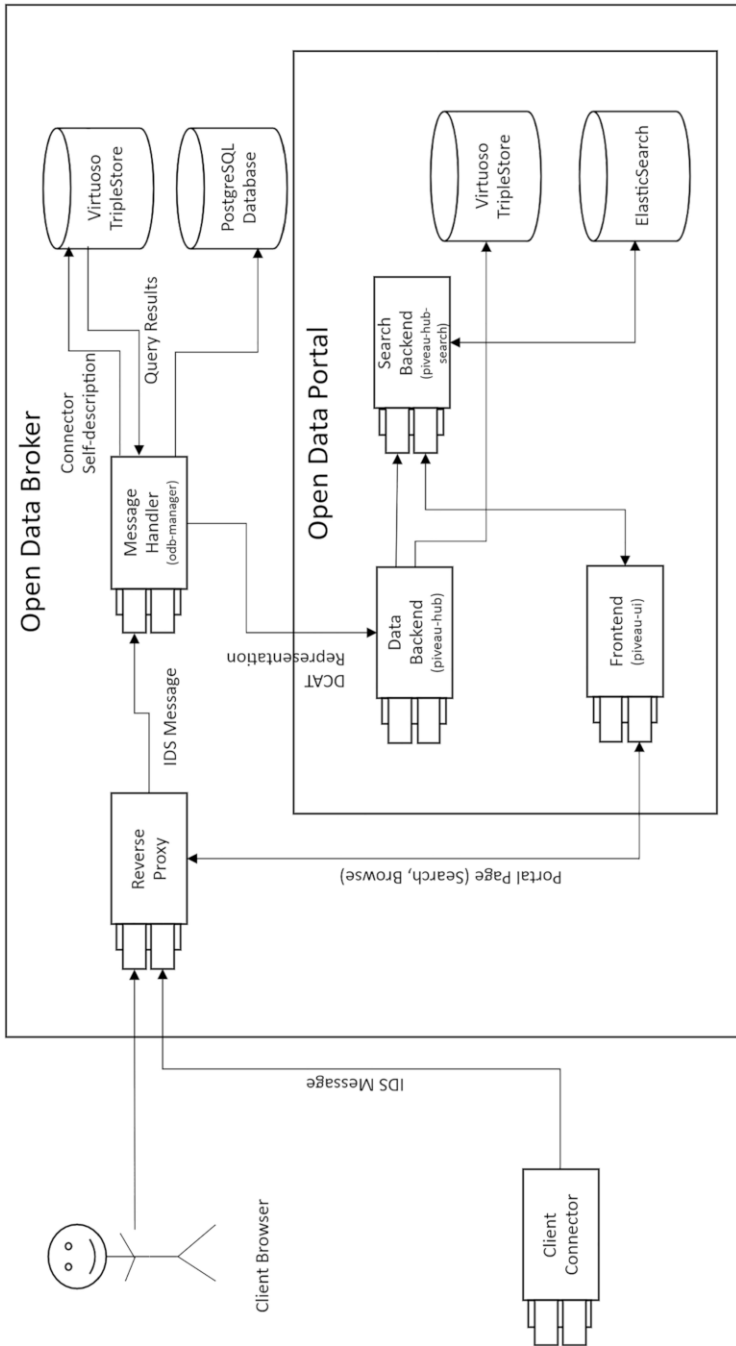


Fig. 14.4 Open Data Broker component architecture

and one Open Data Broker. The connectors have been set up to simulate one specific municipal data owner publishing their data. The connectors expose existing open data offers, by connecting to established open data portals, such as Berlin Open Data, Deutsche Bahn Open Data, and Open.NRW Köln. They are registered to an instance of the Open Data Broker representing a central open data portal for open government data. As such, the portal created in this showcase resembles existing portals such as GovData or the European Data Portal in the way they are combining data offers from a variety of different publishers and smaller portals. Each Open Data Connector publishes 30–50 exemplary datasets totaling 318 datasets registered at the Open Data Broker. Figure 14.5 shows the web frontend of the Open Data Broker in the open government data demo showcase. With this use case, we demonstrate that the Public Data Space solution can replicate the existing open data domain of data publishers and open data portals. Furthermore, the use case shows that communication between open data publishers and open data portals can be achieved by utilizing IDS. The scenario confirms the Open Data Connector’s and Open Data Broker’s ability to exchange IDS messages, allowing the Connector to exert full control over the metadata registered at the Broker and enabling the benefits presented in Sect. 14.4. In addition, the Open Data Connector showed to be fully functional in providing data

The screenshot displays the web interface of the Open Data Broker. At the top, there is a navigation bar with 'International Data Spaces', 'Datasets', 'Connectors', and 'English'. A search bar on the right contains the text 'Search datasets...' and a magnifying glass icon. Below the search bar, it indicates '318 datasets found' and provides options for 'Last Modified' and 'Relevance'. On the left side, there is a sidebar with a map of Europe, 'Settings' (Operator: AND/OR), 'Countries' (Germany: 318), and a list of 'Connectors' with their respective dataset counts. The main content area shows three dataset cards:

- Parkmöglichkeiten**: Created: 18.09.2020 12:45, Updated: 16.12.2020 03:23. Description: 'Dieser Datensatz umfasst die Standorte der Parkmöglichkeiten in der Hanse- und Universitätsstadt Rostock mit Informationen zu Adresse, Art, Standort, Betreiber, Pkw-Stellplätzen, Wohnmobil-Stellplätzen, Bus-Stellplätzen und Gebühren. Die Ressourcen werden in der Regel im folgenden Zeitabstand aktual...' Available formats: WFS, WMS, GeoJSON, CSV, GML, Excel XL, KML.
- Parkscheinautomaten**: Created: 18.09.2020 12:44, Updated: 16.12.2020 03:23. Description: 'Dieser Datensatz umfasst die Standorte der Parkscheinautomaten der Hanse- und Universitätsstadt Rostock mit Informationen zu Nummer, Bezeichnung, Lage, Zone, Handyparkzone, Tarif, Bewirtschaftungszeiten, Mindestparkdauer zum Normaltarif (Zone W zur Hauptsaison), Maximalparkdauer zum Normaltarif (Zon...' Available formats: WMS, Excel XL, GML, KML, WFS, GeoJSON, CSV.
- Reisebusparkplätze und -terminals**: Created: 18.09.2020 12:43, Updated: 16.12.2020 03:23. Description: 'Dieser Datensatz umfasst die Standorte der Reisebusparkplätze und -terminals in der Hanse- und Universitätsstadt Rostock mit Informationen zu Adresse, Bezeichnung, Art, Stellplätzen, Gebühren und Einschränkungen. Die Ressourcen werden in der Regel im folgenden Zeitabstand aktualisiert: 365 Tage' Available formats: GML, CSV, GeoJSON, WFS, KML, Excel XL, WMS.

Fig. 14.5 Open Data Broker with published open government data

to a consuming instance of the Enterprise Integration Connector<sup>4</sup> and an instance of the DataSpace Connector<sup>5</sup> during the IDSA Plugfest activities in August 2020.

## 14.6 Discussion and Conclusion

In this paper, we have investigated and demonstrated how the specifications and artifacts of the Industrial Data Spaces can act as a powerful foundation for building and fostering open data ecosystems. We have designed a comprehensive IDS-based architecture and a working prototype to create a Public Data Space, where both public and private organizations can disseminate open data in decentralized, timely, and self-determined fashion.

Open data is currently mostly provided by actors from the public domain and rarely by private companies. Yet, it holds great economic and social potential. A variety of open specifications and software solutions exists for publishing open data. However, it still faces many barriers, mostly regarding data quality and availability. The IDS architecture can be a powerful building block to mitigate these barriers and enable the industry to become an open data user and provider. Many IDS use cases can benefit from using open data, e.g., using traffic data in a digital supply chain. In addition, the engagement with open data in the context of IDS can create awareness about it, leading to the creation of industrial open data.

IDS offer a selection of specifications and artifacts for trustworthy data sharing. Although many features relate to confidential data sharing, the very core mechanisms are suitable matches for creating open data ecosystems. This includes the well-defined information model, the principle of decentralization, and the standardized communication protocols. In addition, the IDS Connector-Broker architecture allows implementing a novel approach for making open data accessible. The data providers actively push their data offering to a central broker platform while maintaining sovereignty. This increases availability and timeliness in comparison to pull or harvesting mechanisms, applied by central services, like the Google Dataset Search or the European Data Portal. Hence, it maps the decentralized nature of open data to a technical solution while avoiding the complexity of fully decentralized approaches, like peer-to-peer methods. In addition, both IDS and open data foster the same metadata standards (DCAT).

Our Public Data Space prototype applies and consolidates the IDS artifacts information model, connector, broker, and data source adapters to an out-of-the-box solution for IDS-conform data publishing. It currently can integrate and publish data from four sources: CKAN, PostgreSQL, file system, and the open data platform mCloud. In addition, the entire solution is available as open source and acts as a

---

<sup>4</sup>[https://www.dataspace.fraunhofer.de/en/software/connector/ei\\_connector.html](https://www.dataspace.fraunhofer.de/en/software/connector/ei_connector.html)

<sup>5</sup><https://www.dataspace-connector.io/>

transparent implementation of IDS artifacts, representing a starting point for potential IDS open data providers or other IDS use cases.

A successful broader adoption of our solution will require an adoption by actual providers of open data from the public and private domain. Our solution can be integrated on top of existing platforms, in principle allowing a smooth migration. Yet, this process requires substantial efforts and the support of the current actors of open data. We believe that with the rising popularity of International Data Spaces, the demand for integration and adoption in domains beyond industry data sharing will increase.

In future work, we aspire a Base Security Profile certification of our connector, adding an additional layer of trust to open data. In addition, we plan to extend the application domain towards the publication and sharing of open research Data. In addition, the Open Data Broker's interoperability with consuming IDS connectors will be demonstrated further in upcoming IDSA Plugfest activities. Finally, we will provide an extended data model to offer a richer, semantic description of the available data.

## References

1. Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258–268. <https://doi.org/10.1080/10580530.2012.716740>
2. Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., et al. (2018). The open data landscape. In Y. Charalabidis, A. Zuiderwijk, C. Alexopoulos, et al. (Eds.), *The world of open data: Concepts, methods, tools and experiences* (pp. 1–9). Springer International Publishing.
3. Opendatasoft. (2021a). Open data inception - A comprehensive list of 2600+ open data portals in the world. In *Open Data Inception*. Accessed January 25, 2021, from <https://opendatainception.io/>
4. Kirstein, F., Dittwald, B., Dutkowski, S., et al. (2019). Linked data in the European data portal: A comprehensive platform for applying DCAT-AP. In I. Lindgren, M. Janssen, H. Lee, et al. (Eds.), *Electronic government* (pp. 192–204). Springer International Publishing.
5. Manghi, P., Manola, N., Horstmann, W., & Peters, D. (2010). An infrastructure for managing EC funded research output: The OpenAIRE project. *The Grey Journal (TGJ): An International Journal on Grey Literature*, 6(1), 31–40.
6. Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017). Perception of key barriers in using and publishing open data. *JeDEM - eJournal of eDemocracy and Open Government*, 9, 134–165. <https://doi.org/10.29379/jedem.v9i2.465>
7. Neumaier, S., Thurnay, L., Lampoltshammer, T. J., & Knap, T. (2018). *Search, filter, fork, and link open data: The ADEQUATE platform: Data- and community-driven quality improvements*. In Companion Proceedings of the Web Conference 2018 (pp 1523–1526). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland.
8. Noy, N., Burgess, M., & Brickley, D. (2019). *Google dataset search: Building a search engine for datasets in an open Web ecosystem*. In 28th Web Conference (Web Conf 2019).
9. W3C. (2020). *Data Catalog Vocabulary (DCAT) - Version 2*. Accessed March 10, 2019, from <https://www.w3.org/TR/vocab-dcat-2/>

10. European Commission. (2021). About DCAT application profile for data portals in Europe. *Joinup*. Accessed February 10, 2019, from <https://joinup.ec.europa.eu/solution/dcat-application-profile-data-portals-europe/about>
11. CKAN Association. (2021) *CKAN*. Accessed March 9, 2019, from <https://ckan.org/>
12. Etalab. (2021). *uData Documentation*. Accessed February 5, 2021, from <https://udata.readthedocs.io/en/latest/>
13. OpenDataSoft. (2021b) OpenDataSoft - Turn your data into actions. In *OpenDataSoft*. Accessed March 17, 2019, from <https://www.opendatasoft.com/>
14. Rakhmawati, N. A., Umbrich, J., & Karnstedt, M., et al., (2013). Querying over Federated SPARQL Endpoints: A state of the art survey. *arXiv*, 13061723 [cs].
15. Truong, D.-D., Nguyen-Van, T., Nguyen, Q.-B., et al. (2019). Blockchain-based open data: An approach for resolving data integrity and transparency. In T. K. Dang, J. Küng, M. Takizawa, & S. H. Bui (Eds.), *Future data and security engineering* (pp. 526–541). Springer International Publishing.
16. Tran, A. B., Xu, X., Weber, I., et al. (2017). Reperator: A registry generator for Blockchain. *CEUR Workshop Proceedings*, 1848, 81–88.
17. García-Barriocanal, E., Sánchez-Alonso, S., & Sicilia, M.-A. (2017). Deploying metadata on Blockchain technologies. In E. Garoufallou, S. Virkus, R. Siatiri, & D. Koutsomiha (Eds.), *Metadata and semantic research* (pp. 38–49). Springer International Publishing.
18. Robinson, D. C., Hand, J. A., Madsen, M. B., & McKelvey, K. R. (2018). The Dat project, an open and decentralized research data tool. *Scientific Data*, 5, 180221. <https://doi.org/10.1038/sdata.2018.221>
19. Kirstein, F., Stefanidis, K., Dittwald, B., et al. (2020). Piveau: A large-scale open data management platform based on semantic web technologies. In A. Harth, S. Kirrane, A.-C. Ngonga Ngomo, et al. (Eds.), *The semantic web* (pp. 648–664). Springer International Publishing.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

