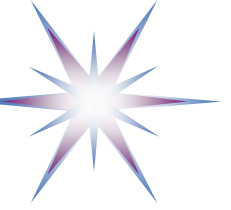




Data Spaces Concept in European Data Policy and Data Spaces Definitions

Yuri Demchenko
BoF Data Spaces Taxonomy
RDA19, 21 June 2022



Outline

- Why do we need taxonomy?
- European Data Spaces and related documents
 - Data Act - Data Governance Act – GDPR - Data Altruism
- Industrial Data Spaces
 - IDSA Reference Architecture Model and Data Sovereignty
 - Open DEI Data Spaces Design principles
- Data Spaces definitions by EU documents and IDSA
- Discussion: What to include into Data Spaces definition?



Goal: Why do we need taxonomy?

- Common language/terms and data models
- Interoperability in professional communication and data sharing
 - Interoperability between Research and Industry
- Basis for interoperable and sustainable infrastructure services
 - Based on the specified Data Spaces properties
- Standardisation: Liaison between RDA and Industry/IEEE/ISO

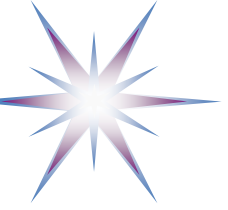
EOSC/European Interoperability Framework – 4 layers

- Technical/infrastructure
- Semantic/metadata
- Organisational
- Legal



European Data Space(s)

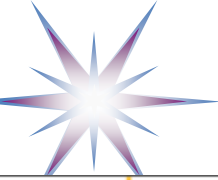
- **European Data Governance Act**
 - Support for set-up and development of common European data spaces in strategic domains, involving both private and public players
- **European rules and values to ensure free flow of data between EU countries and outside**
 - Sharing and re-using data, data monetisation
 - Enabling regulated environment for data processing
- **Industrial Data Space: Sharing data while ensuring data Sovereignty**
 - Data Sovereignty achieved by attaching policy and usage conditions to data
 - Enabling regulated environment for data processing
- **International Data Spaces Association and GAIA-X**



Regulation on European Data Governance (Data Governance Act, Nov 2020)

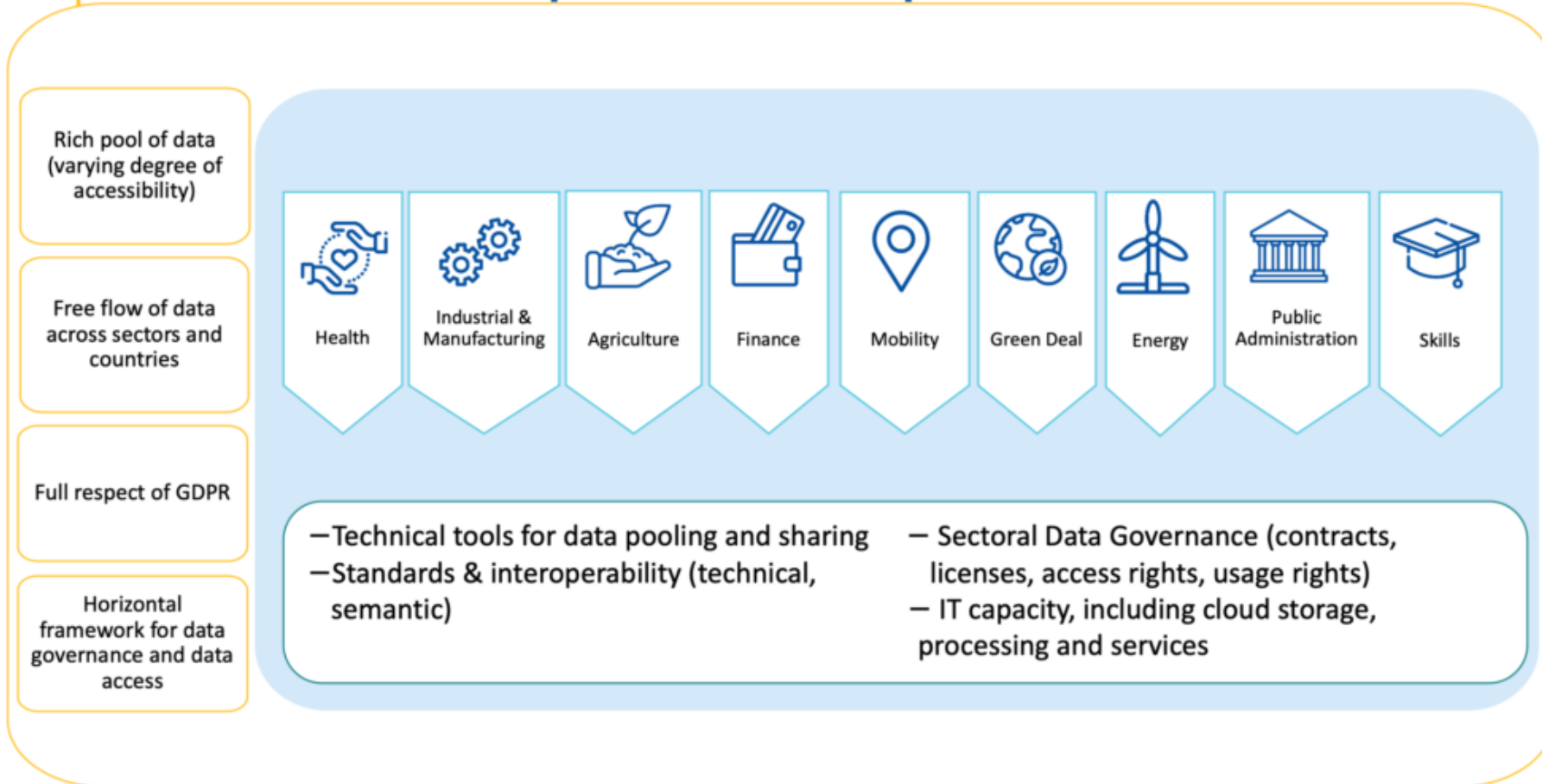
- EU/Parliament Regulation on European data governance (Data Governance Act) SEC(2020) 405 final, Nov 2020, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767>
 - ‘data’ means any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording;
- Facilitate Sharing and Re-use of Data
 - New way of data governance will increase trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the reuse of data.
- Making public sector data available for re-use, in situations where such data is subject to rights of others.
 - “Data the use of which is dependent on the rights of others” or “data subject to the rights of others” covers data that might be *subject to data protection legislation, intellectual property, or contain trade secrets or other commercially sensitive information*.
- Allowing personal data to be used with the help of a ‘personal data-sharing intermediary’, designed to help individuals exercise their rights under the General Data Protection Regulation (GDPR).
 - Allowing data use on *altruistic* grounds.

Questions: What is data? How is it related to Data Spaces? What operations on data?



Data Governance Act: Common European Data Spaces

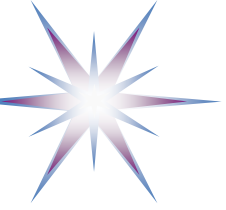
Common European data spaces



- Health
- **Industry and Manufacturing**
- Agriculture
- Finance
- Mobility
- Green Deal
- Energy
- Public Administration
- Skills

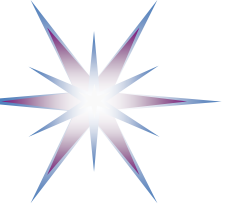
Questions: Are any developments in other domains, except Industry and Manufacturing?

[ref] Common European Data Spaces <http://dataspaces.info/common-european-data-spaces/>



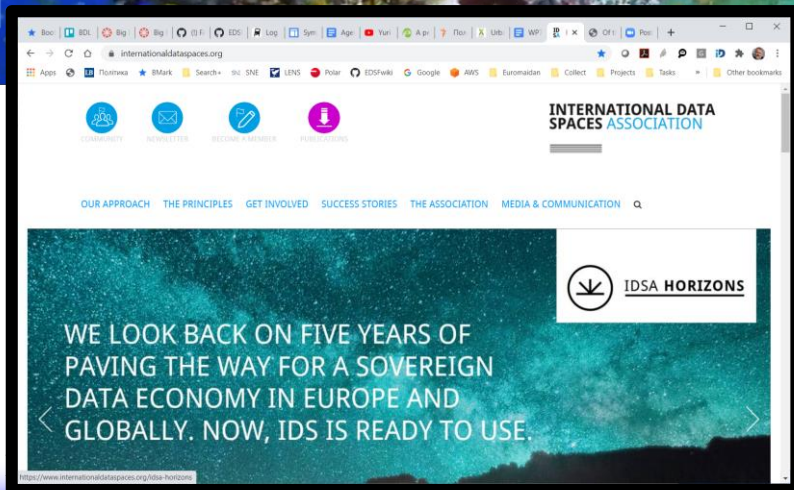
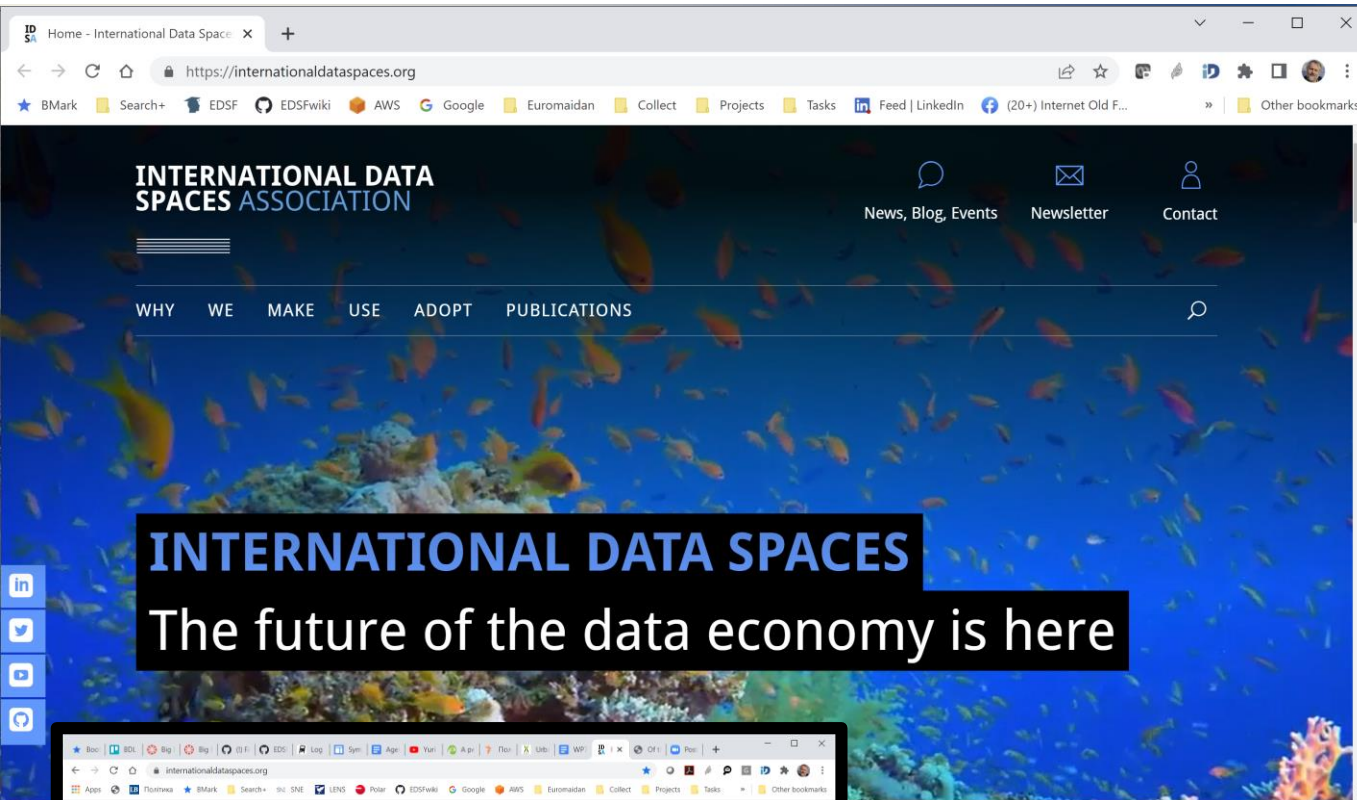
Industrial Data Spaces: Sharing Data

- Data sharing along the whole data handling flow in industrial or business processes – To allow for the 3rd party services
 - Sensors - IoT – Ingest/Load – Ingest/Transform – Process – Act/Operate
- Sharing and Sovereignty are key to make value of data
 - Network of trusted data
- Data Sovereignty achieved by attaching policy and usage conditions to data
 - In non-fungible/blockchain way or protected environment
- Enabling regulated environment for data processing
 - Enclave computing in modern cloud technology

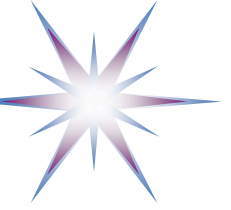


International Data Spaces Association (IDSA)

<https://www.internationaldataspaces.org/>



- More than 120 members
- Started 2016 as Industrial Data Space initiative (supported by German project)
- Re-defined as International Data Space Association (IDSA)
- Published IDSA Reference Architecture Model Version 3.0 (RAM3.0, 2018)
 - Whitepaper and use cases
- Focused on Industrial data sharing
- Key Issue is Data **Sovereignty**
- Number of associated H2020 and HE projects
- Community activities
 - Weekly architecture design meetings
 - Governance task force
 - Winter/Summer schools



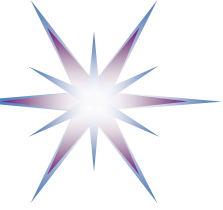
Data Spaces definitions: Review

- European documents
- IDSA and OPEN DEI



Data Spaces in the European Data Governance Policy in Horizon2020 and Horizon Europe

- Term “Data Space” appeared in EU documents related to data regulations and data policy in a form of “**European Data Space**” (European Strategy for Data, 19 Feb 2020) which is defined as
 - “a genuine single market for data, open to data from across the world – where personal as well as non-personal data, including sensitive business data, are **secure and businesses also have easy access** to an almost infinite amount of high-quality industrial data, boosting growth and creating value, while minimising the human carbon and environmental footprint. It should be a space where **EU law can be enforced** effectively, and where all data-driven **products and services comply with the relevant norms** of the EU’s single market.
- Common European rules and efficient enforcement mechanisms should ensure that:
 - *Data can flow within the EU and across sectors;*
 - *European rules and values, in particular personal data protection, consumer protection legislation and competition law, are fully respected;*
 - *Rules for access to and use of data are fair, practical and clear, and there are clear and trustworthy data governance mechanisms in place;*
 - *There is an open, but assertive approach to international data flows, based on European values.*



IDSA Data Spaces – No actual definition, used depending on context

- Reference Architecture Model version 3.0 (referred to as IDSA RAM3.0) the Data Spaces concept
 - *The International Data Spaces (IDS) is **a virtual data space** leveraging existing standards and technologies, as well as governance models well-accepted in the data economy, to facilitate **secure and standardized data exchange and data linkage in a trusted business ecosystem**. It thereby provides a basis for creating smart-service scenarios and facilitating innovative cross-company business processes, while at the same time guaranteeing data sovereignty for data owners.*
- Data Sovereignty is a central aspect of the International Data Spaces. It can be defined as a natural person's or corporate entity's capability of being entirely self-determined with regard to its data.
- Strategic requirements
 - *TRUST, SECURITY AND DATA SOVEREIGNTY, ECOSYSTEM OF DATA*
- **IDS Connector** is introduced as a mandatory element to implement IDS functionality and ensure above properties.



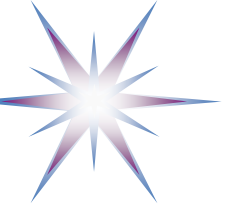
OPEN DEI Data Spaces Design Principles (2021)

- Definition by OPEN DEI is built around infrastructure for data sharing and exchange: **a data space can be defined as a federated data ecosystem within a certain application domain and based on shared policies and rules.** The users of such data spaces are enabled to access data in a secure, transparent, trusted, easy and unified fashion.
 - *Abstracted from this technical definition, a data space can be defined as a federated data ecosystem within a certain application domain and based on shared policies and rules. The users of such data spaces are enabled to access data in a secure, transparent, trusted, easy and unified fashion. These access and usage right can only be granted by those persons or organisations who are entitled to dispose of the data.*
- **A data space is made up of multiple actors, which together form a data ecosystem.** Data will only be shared and exchanged between actors if they decide to do so. Data will only be shared and exchanged between actors if they decide to do so. One reason for doing so may be to create business value. Such business value does not necessarily need to be of monetary nature, but can also manifest itself in a better quality of a product or service. (p.92)
- A data space is defined as a decentralized infrastructure for trustworthy data sharing and exchange in data ecosystem based on commonly agreed principles (p.23). Data spaces require the following elements:
 - *Building blocks such as data platforms, providing support for effective data sharing and exchange as well as for engineering and deployment of data exchange and processing capabilities;*
 - *Building blocks such as data marketplaces, where data providers can offer and data consumers can request data, as well as data processing applications;*
 - *Building blocks ensuring data sovereignty, i.e. the ability for each stakeholder to control their data by making decisions as to how digital processes, infrastructures, and flows of data are structured, built and managed, based on an appropriate governance scheme enabling specification of terms and conditions.*



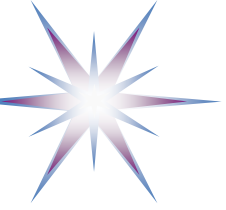
OPEN DEI Data Spaces Design Principles (2021) - 1

- Definition by OPEN DEI is built around infrastructure for data sharing and exchange: **a data space can be defined as a federated data ecosystem within a certain application domain and based on shared policies and rules.** The users of such data spaces are enabled to access data in a secure, transparent, trusted, easy and unified fashion.
 - *Abstracted from this technical definition, a data space can be defined as a federated data ecosystem within a certain application domain and based on shared policies and rules. The users of such data spaces are enabled to access data in a secure, transparent, trusted, easy and unified fashion. These access and usage right can only be granted by those persons or organisations who are entitled to dispose of the data.*
- **A data space is made up of multiple actors, which together form a data ecosystem.** Data will only be shared and exchanged between actors if they decide to do so. Data will only be shared and exchanged between actors if they decide to do so. One reason for doing so may be to create business value. Such business value does not necessarily need to be of monetary nature, but can also manifest itself in a better quality of a product or service. (p.92)



OPEN DEI Data Spaces Design Principles (2021) - 2

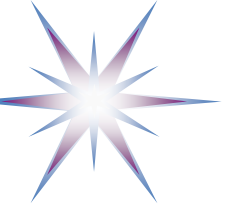
- A data space is defined as a **decentralized infrastructure for trustworthy data sharing and exchange in data ecosystem based on commonly agreed principles** (p.23). Data spaces require the following elements:
 - *Building blocks such as data platforms, providing support for effective data sharing and exchange as well as for engineering and deployment of data exchange and processing capabilities;*
 - *Building blocks such as data marketplaces, where data providers can offer and data consumers can request data, as well as data processing applications;*
 - *Building blocks ensuring data sovereignty, i.e. the ability for each stakeholder to control their data by making decisions as to how digital processes, infrastructures, and flows of data are structured, built and managed, based on an appropriate governance scheme enabling specification of terms and conditions.*



State of Data Spaces, Report by Finland Ministry of Transport and Communications (October 2021)

This document in general share the definitions by OPEN DEI both Data Space and Data Ecosystem.

- **Data space:** *a framework and a medium that creates a secure space for data exchange.*
 - *Decentralized infrastructure for trustworthy data sharing and exchange in data ecosystems based on commonly agreed principles.*
 - *Purpose-or sector-specific or cross-sectoral interoperable frameworks of common standards and practices to share or jointly process data for, inter alia, development of new products and services, scientific research or civil society initiatives.*
- **Data sharing:** conditioned exchange of data aiming to create added value
- **Data ecosystem:** integration of and interaction between stakeholders to access and share data
- **Data sovereignty:** self-determination in a digital world,



Operations on data(sets) and Data/Data Spaces properties

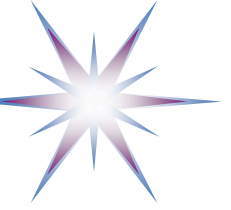
- The traditional POSIX file system object permission model defines three classes of users called owner, group, and other. Each of these classes is associated with a set of permissions: read (r), write (w), and execute (x).
- The most of traditional database define the four basic operations that can be performed on the database records: Create, Read, Update, Delete (typically referred to as CRUD operations).
- Data in general and datasets in particular are complex digital objects that may characterised by multiple parameters, properties and models.
- The following are the common operations on data or datasets can be assumed:
 - Basic
 - Create
 - Store (write)
 - Modify/update
 - Delete
 - Transfer/Copy: Send – Receive - Stream
 - Extended
 - Metadata: Create – Read – Update – Delete
 - DOI: Create – Delete
 - To be extended

Data/dataset properties (generic) – To check RDM recommendations!

- Data model
- Namespace
- Language
- Encoding
- Accuracy/Statistical
- Owner
- Source, Origin
- DOI/PID

Questions to discuss:

- 1) What operations are subject to Data Spaces and what are applications and system internal?
- 2) What infrastructure or system components are required to perform these operations?
- 3) From FAIR data principles to related data properties



Discussion Questions

- How to match all existing definitions that are shared by large community of practitioners?
- Given active development in Industrial Data Space domain, are any developments in other domains, except Industry and Manufacturing?
- What is data? How is it related to Data Spaces? What operations on data?
- What properties to included into Data Spaces definition? – To achieve the main goal in facilitating data sharing.