WEBINAR

Data spaces: Discovering the building blocks

data.europa academy

6 October 2023

10.00 — 11.30 CET

Rules of the game

The webinar will be recorded

For questions, please use the ClickMeeting chat.

Please reserve 3 min after the webinar to help us improve by filling in our feedback form

data.europa academy

Introduction

Flora Kopelou
Data.europa.eu,
Publications Office of the
EU

Clara Pezuela
FIWARE Foundation

Edward Curry
University of Galway

Alexandra Balahur
European Commission
DIGIT

data.
europa
academy

Agenda

10.00 – 10.10	Opening and introduction of the series
10.10 – 10.30	The building blocks of data spaces – *Clara Pezuela*
10.30 – 10.50	Exploring the Research Challenges with Data Spaces – *Edward Curry*
10.50 – 11.10	SEMIC Specifications, services and trainings in support of Data Spaces – *Alexandra Balahur*
11.10 – 11.25	Questions and answers
11.25 – 11.30	Closing remarks

The building blocks of data spaces

Clara Pezuela

DATA SPACES
SUPPORT CENTRE

Discovering the building blocks

Clara Pezuela (FIWARE Foundation)

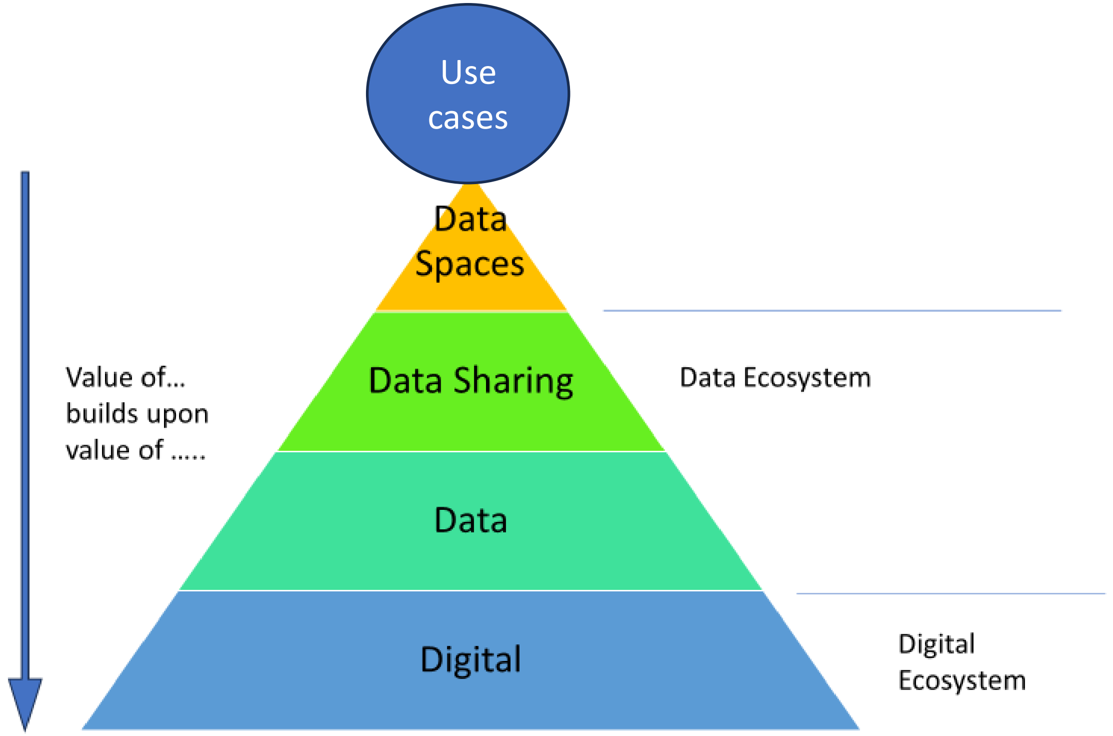
Data Europe webinar - 6 October 2023

Funded by
the European Union

Context of data spaces

DATA SPACES SUPPORT CENTRE

- The value of data spaces relies on the value of data sharing and therefore on the value of data
- Data Spaces operate in digital ecosystems and inherit some of their characteristics
- Data Spaces were introduced in the European Data Strategy to support the idea of “single market of data”
- Being concrete through use cases



Funded by the European Union

Data Space concept

DATA SPACES
SUPPORT CENTRE

A distributed system defined by a governance framework, that enables trustworthy data transactions between participants while supporting trust and data sovereignty (DSSC Glossary)

What is?

- Several infrastructures supporting one or more use cases:
  - Distributed structure
  - Governance framework
  - Enabling trusted data transactions
  - Enabling data sovereignty

What is not?

- A data lake
- Only a data platform
- Only a digital ecosystem

Funded by
the European Union

DSSC mission

DATA SPACES SUPPORT CENTRE

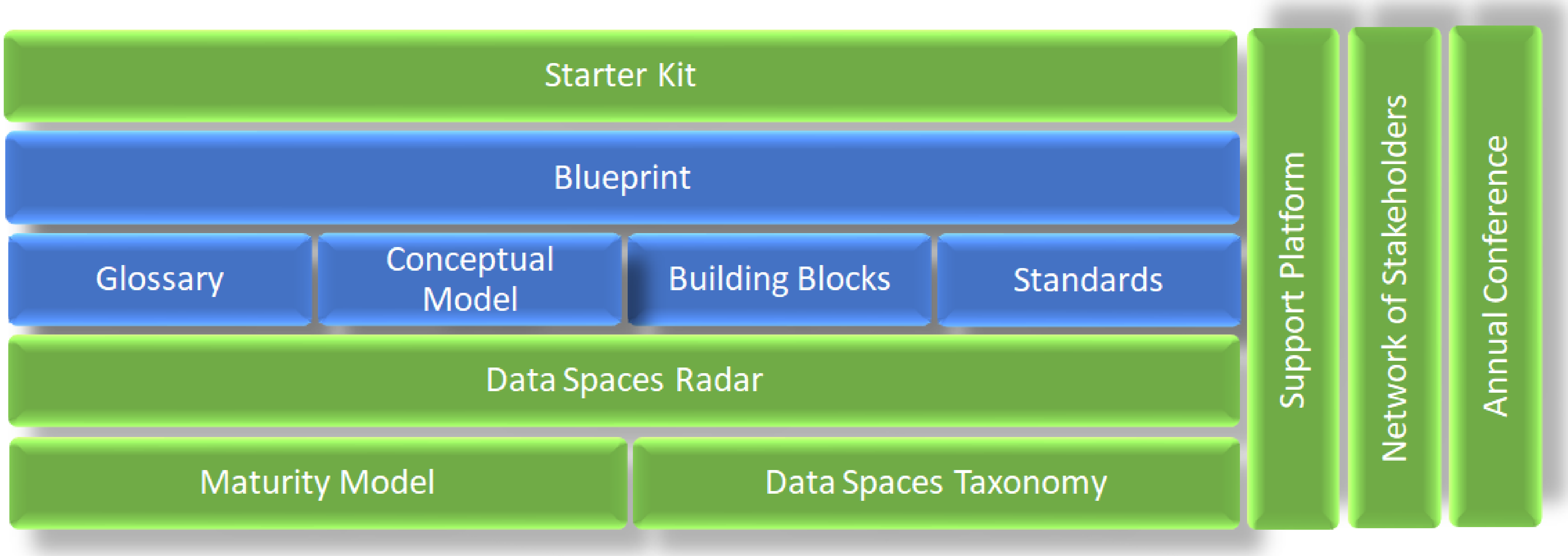
- Enable data spaces to reach a higher flight level faster: a **quick start** and an **accelerated scale-up**
- Provide the tools to address the **basic organizational and technical matters**, required by every data space
- This includes a **blueprint**, best practices, common standards and reference implementations which will be developed according to a **co-creation process**
- Enable dataspaces to focus on their domain-specific business challenges and **provide business benefits to their participants**

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

DSSC assets

DATA SPACES
SUPPORT CENTRE



Starter Kit

Blueprint

Glossary

Conceptual Model

Building Blocks

Standards

Data Spaces Radar

Maturity Model

Data Spaces Taxonomy

Support Platform

Network of Stakeholders

Annual Conference

DELIVERY PLAN

Funded by
the European Union

Blueprint v0.5 just released

DATA SPACES SUPPORT CENTRE

https://dssc.eu/page/knowledge-base

Starter Kit

Blueprint

Building Blocks

Conceptual model

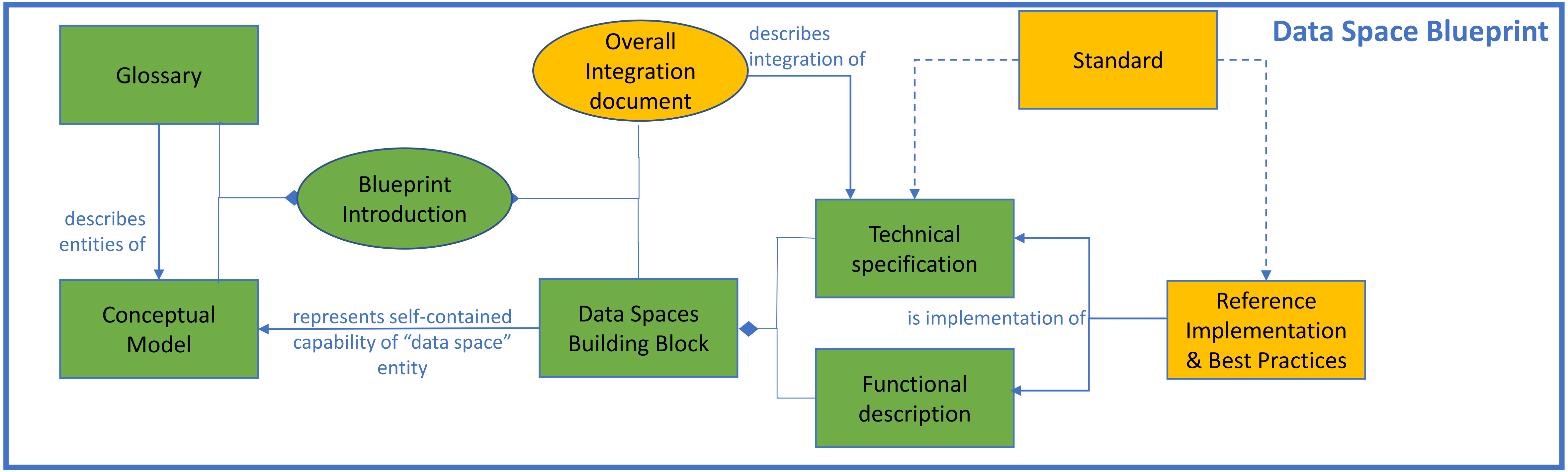
Glossary

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Blueprint

Data Spaces Blueprint Structure



🟩 V0.5 – Publication: September 2023

🟨 V1.0 – Publication: April 2024

Funded by the European Union

Blueprint v0.5 elements

DATA SPACES SUPPORT CENTRE

Glossary

- Curated set of terms and definitions
- V2.0

Conceptual Model

- Model which represents concepts and relationships among them
- V0.5

Building Blocks

- Taxonomy of basic units or components that can be implemented and combined with other building blocks to achieve the functionality of a data space
- V0.5

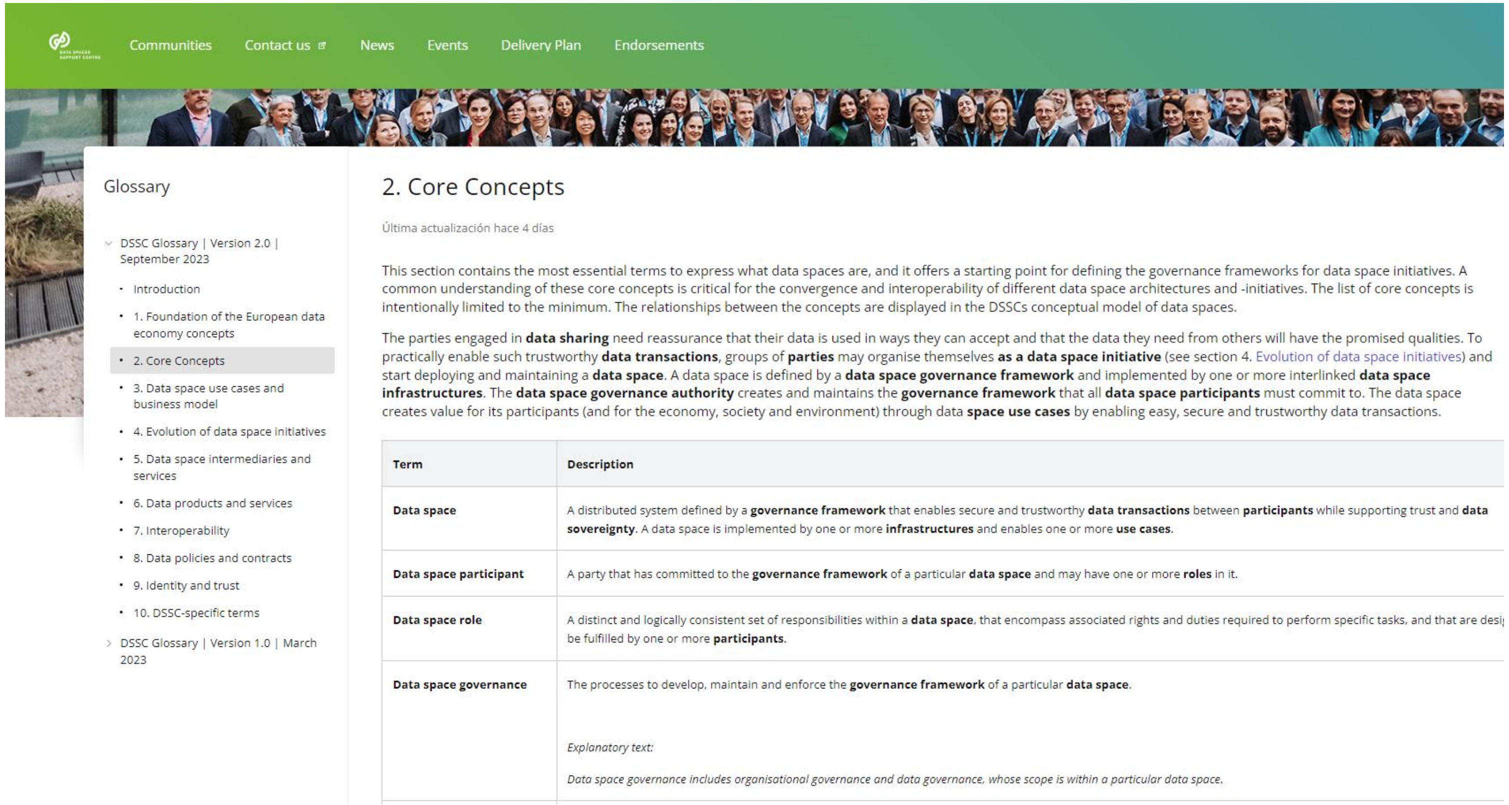
The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Glossary

DATA SPACES SUPPORT CENTRE

- In co-development with data spaces
- Common vocabulary for DSSC publications and communications
- Terms are provided with a criterion and definition
  - Same term may have different meanings in different contexts
- Others may keep their own terminology
- Adopting definitions from EU legislation and international standards, when possible


Communities
Contact us
News
Events
Delivery Plan
Endorsements
Glossary
DSSC Glossary | Version 2.0 | September 2023
Introduction
1. Foundation of the European data economy concepts
2. Core Concepts
3. Data space use cases and business model
4. Evolution of data space initiatives
5. Data space intermediaries and services
6. Data products and services
7. Interoperability
8. Data policies and contracts
9. Identity and trust
10. DSSC-specific terms
DSSC Glossary | Version 1.0 | March 2023

2. Core Concepts

Última actualización hace 4 días

This section contains the most essential terms to express what data spaces are, and it offers a starting point for defining the governance frameworks for data space initiatives. A common understanding of these core concepts is critical for the convergence and interoperability of different data space architectures and -initiatives. The list of core concepts is intentionally limited to the minimum. The relationships between the concepts are displayed in the DSSCs conceptual model of data spaces.

The parties engaged in **data sharing** need reassurance that their data is used in ways they can accept and that the data they need from others will have the promised qualities. To practically enable such trustworthy **data transactions**, groups of **parties** may organise themselves **as a data space initiative** (see section 4. Evolution of data space initiatives) and start deploying and maintaining a **data space**. A data space is defined by a **data space governance framework** and implemented by one or more interlinked **data space infrastructures**. The **data space governance authority** creates and maintains the **governance framework** that all **data space participants** must commit to. The data space creates value for its participants (and for the economy, society and environment) through data **space use cases** by enabling easy, secure and trustworthy data transactions.

Term	Description
Data space	A distributed system defined by a **governance framework** that enables secure and trustworthy **data transactions** between **participants** while supporting trust and **data sovereignty**. A data space is implemented by one or more **infrastructures** and enables one or more **use cases**.
Data space participant	A party that has committed to the **governance framework** of a particular **data space** and may have one or more **roles** in it.
Data space role	A distinct and logically consistent set of responsibilities within a **data space**, that encompass associated rights and duties required to perform specific tasks, and that are desi be fulfilled by one or more **participants**.
Data space governance	The processes to develop, maintain and enforce the **governance framework** of a particular **data space**. *Explanatory text:* *Data space governance includes organisational governance and data governance, whose scope is within a particular data space.*

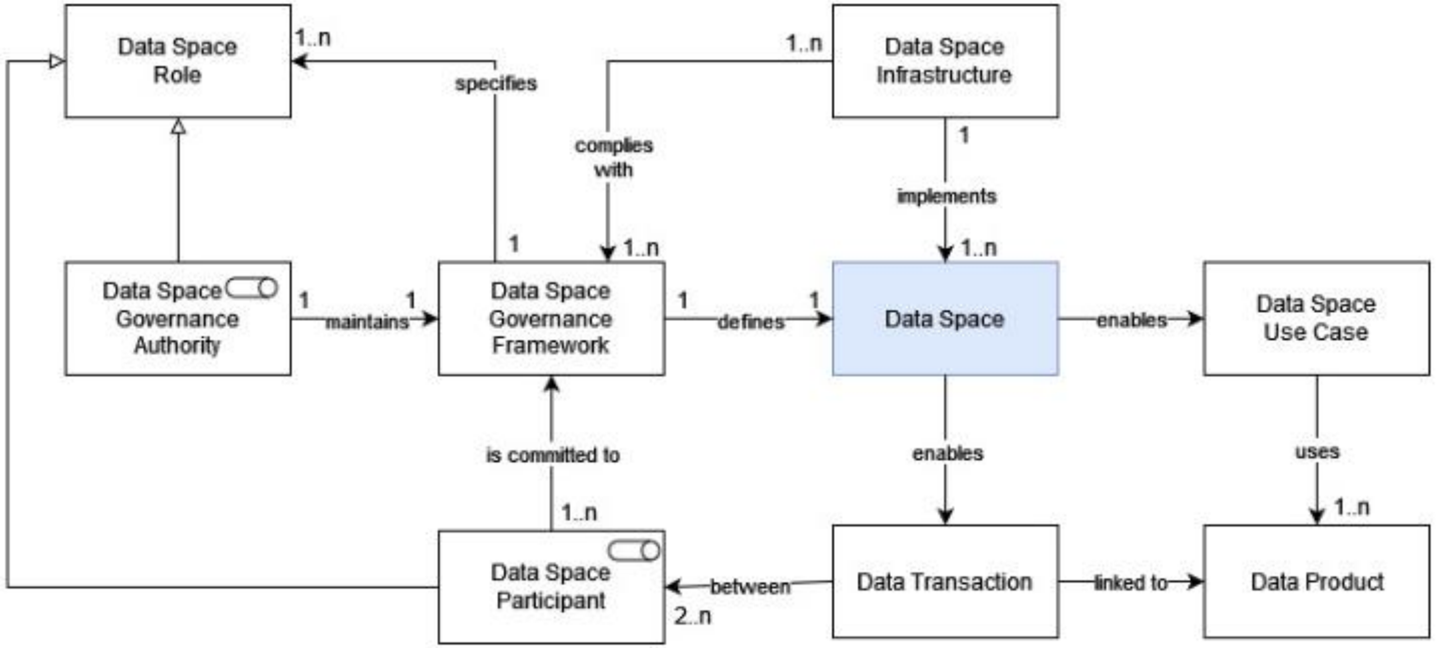
The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Conceptual Model

- Provide a well-defined language
- Provide a high-level view of what a data space is
- Future structure will provide:
  - An overarching view of the data space environment (level 0)
  - Detailed view on individual concepts (level 2)

Level 1: basic terminology and key concepts



The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Building Blocks (BBs)

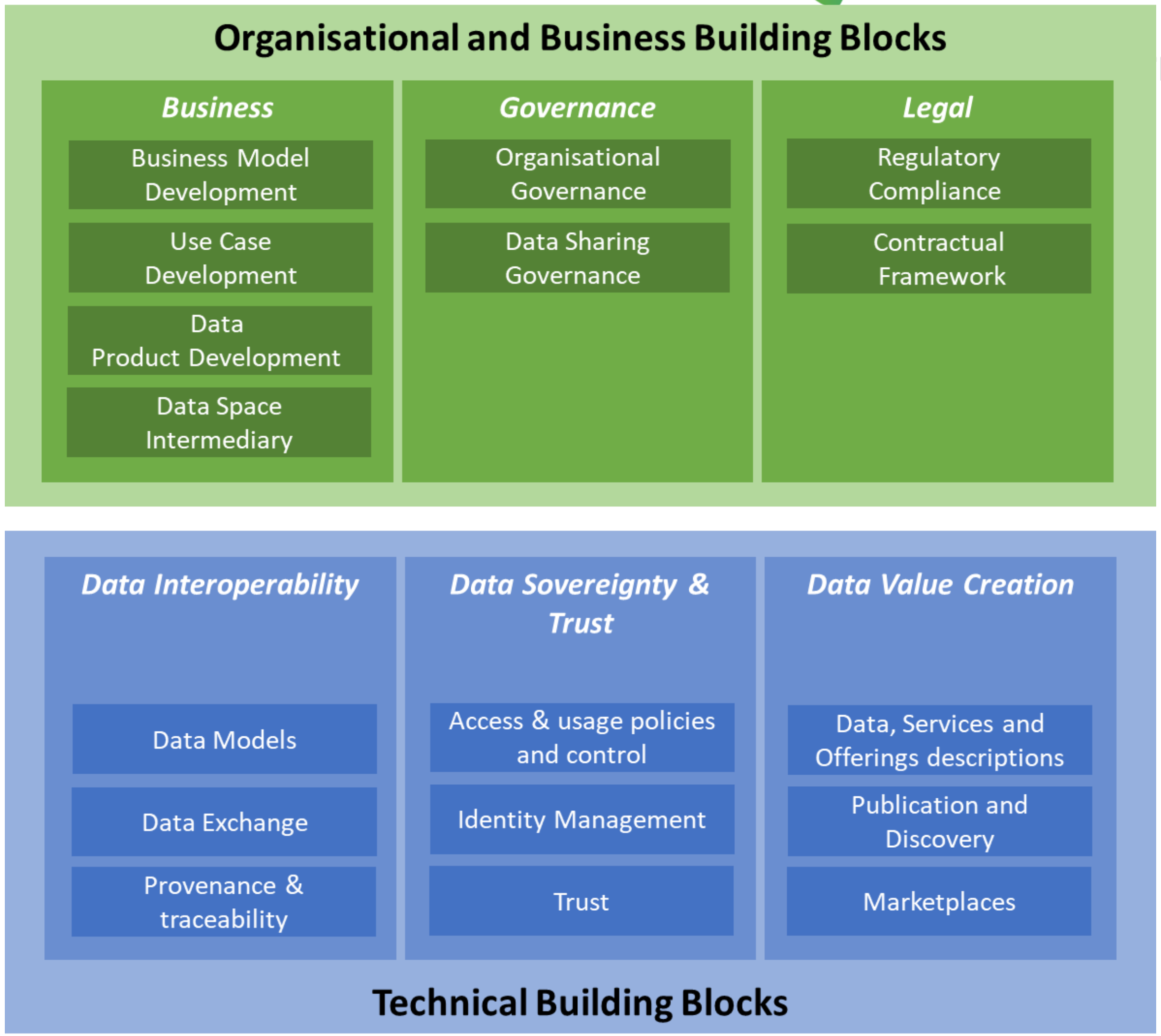
Organisational & business building blocks

Describe capabilities on a governance, legal and business level.

Technical Building Blocks

Describe capabilities on a technical level about:

Data Interoperability
Data Sovereignty & Trust
Data Value Creation



Organisational and Business Building Blocks

Business
- Business Model Development
- Use Case Development
- Data Product Development
- Data Space Intermediary

Governance
- Organisational Governance
- Data Sharing Governance

Legal
- Regulatory Compliance
- Contractual Framework

Technical Building Blocks

Data Interoperability
- Data Models
- Data Exchange
- Provenance & traceability

Data Sovereignty & Trust
- Access & usage policies and control
- Identity Management
- Trust

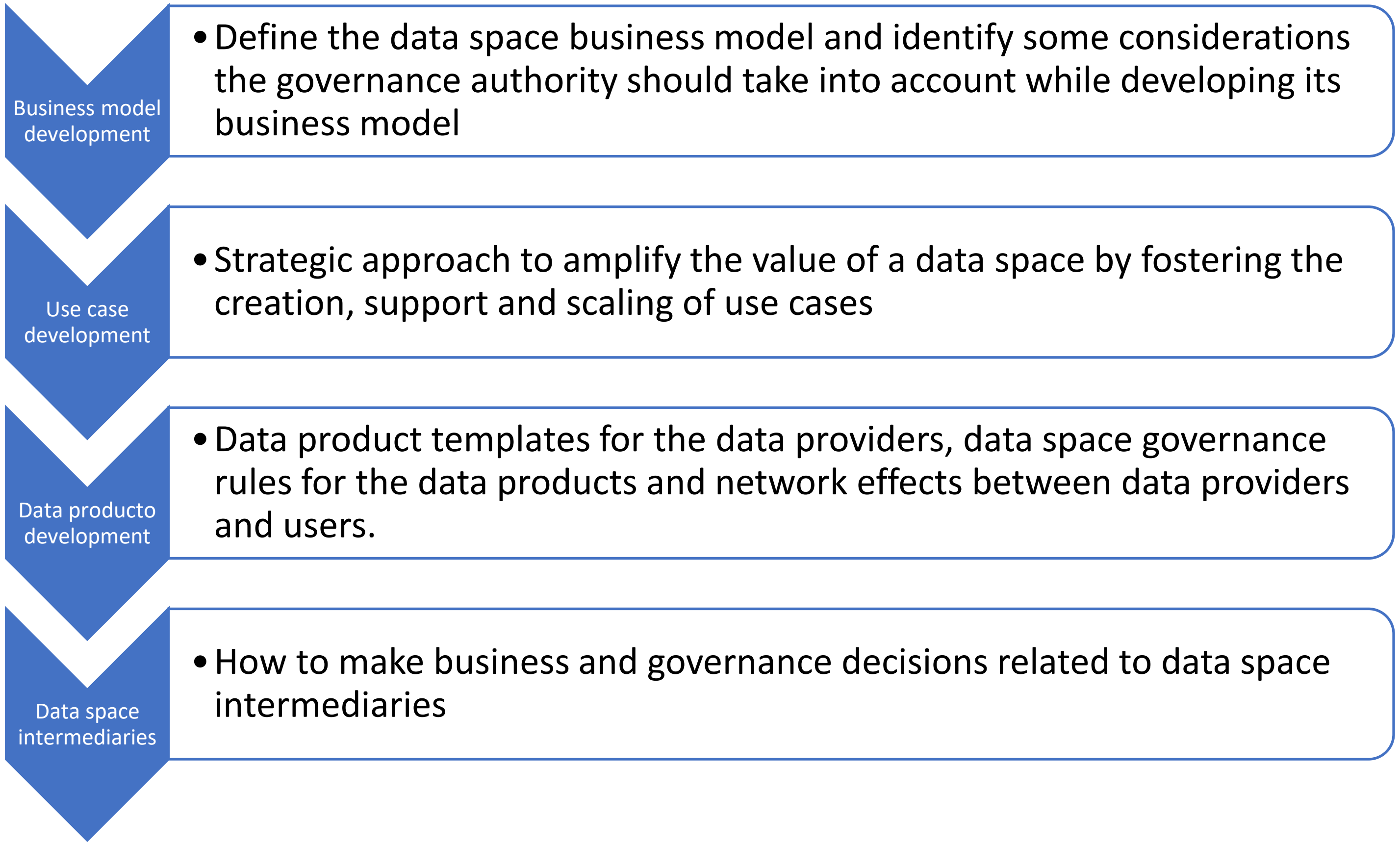
Data Value Creation
- Data, Services and Offerings descriptions
- Publication and Discovery
- Marketplaces

Funded by
the European Union

Organisational and Business BBs - Business

DATA SPACES SUPPORT CENTRE

- Provides the essential concepts necessary in the business modelling of a data space
- Data space business model vs data space participant business model



Business model development

- Define the data space business model and identify some considerations the governance authority should take into account while developing its business model

Use case development

- Strategic approach to amplify the value of a data space by fostering the creation, support and scaling of use cases

Data producto development

- Data product templates for the data providers, data space governance rules for the data products and network effects between data providers and users.

Data space intermediaries

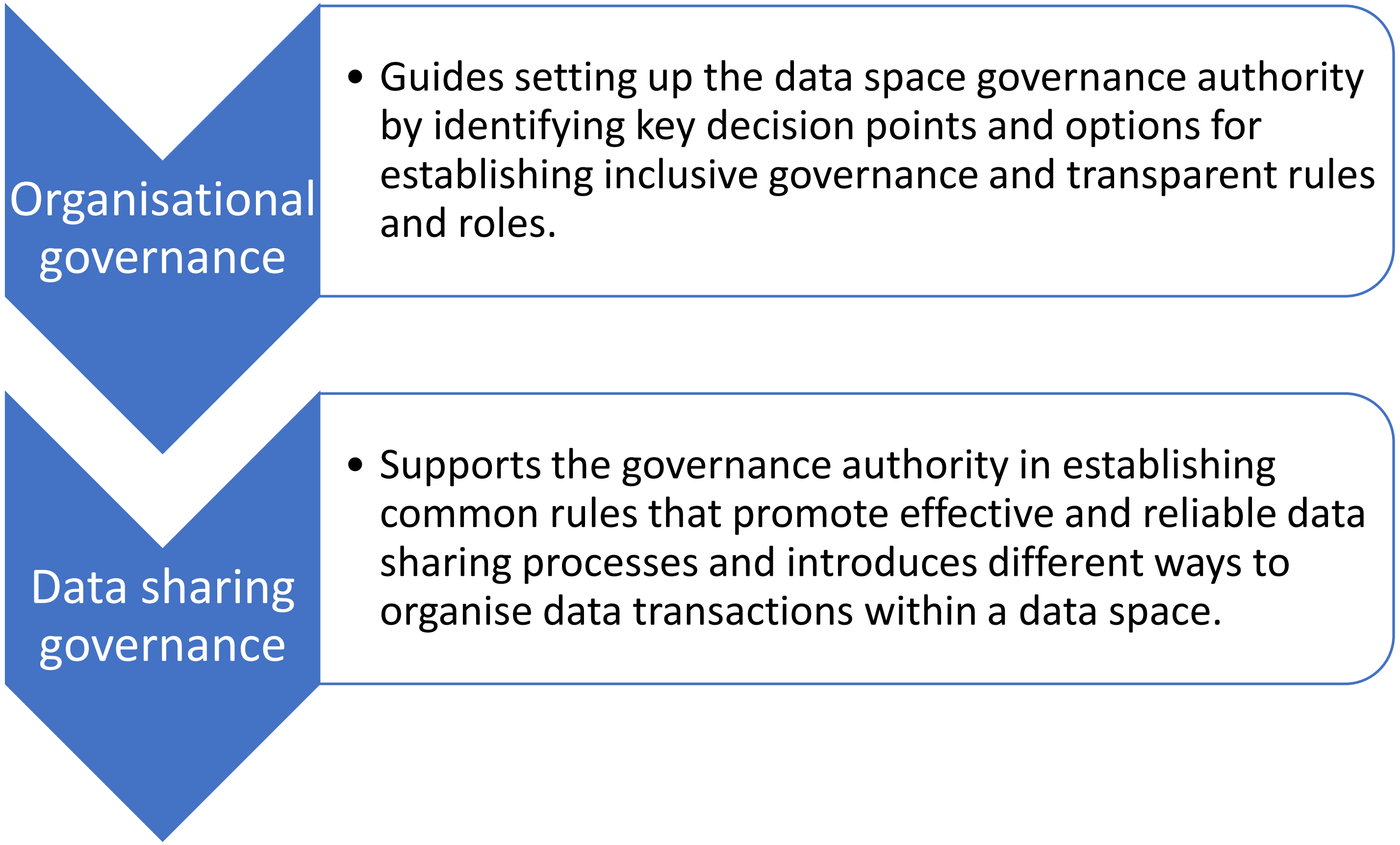
- How to make business and governance decisions related to data space intermediaries

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Organisational and Business BBs - Governance

- Focused on data space-level governance, emphasizing its dynamic nature.
- Governance in data spaces needs to adapt as the data space evolves.
- This requires data space participants in the data space to work together strategically for effective governance.



Organisational governance

- Guides setting up the data space governance authority by identifying key decision points and options for establishing inclusive governance and transparent rules and roles.

Data sharing governance

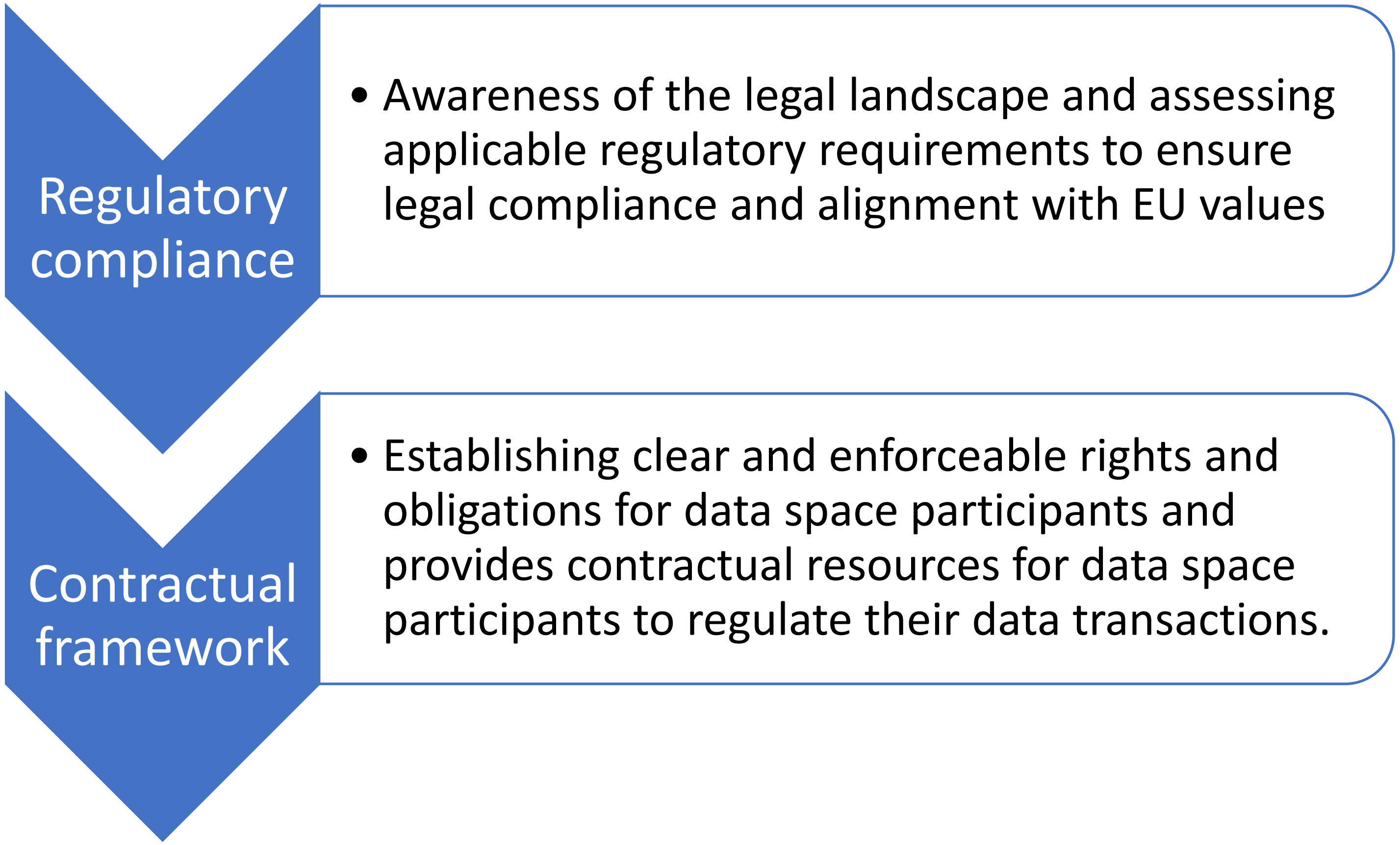
- Supports the governance authority in establishing common rules that promote effective and reliable data sharing processes and introduces different ways to organise data transactions within a data space.

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by
the European Union

Organisational and Business BBs - Legal

- Provide guidance and resources for the data space initiatives to ensure compliance with legislation and establish a robust contractual framework.



Regulatory compliance

- Awareness of the legal landscape and assessing applicable regulatory requirements to ensure legal compliance and alignment with EU values

Contractual framework

- Establishing clear and enforceable rights and obligations for data space participants and provides contractual resources for data space participants to regulate their data transactions.

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by
the European Union

Technical BBs – Data Interoperability

	Description	Key capabilities	Link to specifications
Data Models	Capabilities to define and use shared semantics in a data space	• Vocabularies • Vocabulary management process • Vocabulary hub	Data Models
Data Exchange	Capabilities relating to the actual exchange and sharing of data	• Meta specifications and best practices for the adoption of existing data exchange APIs • Generic purpose data exchange API, including methods to query, update and delete data • Tooling to use and maintain data exchange APIs	Data Exchange
Provenance & Traceability	Capabilities for tracking the process of data sharing, so it becomes traceable and compliant	• Framework for requirements for observability • Third parties to provision or use evidence • Mechanisms to provide and use evidence of the activities of a transaction	Provenance & Traceability

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by
the European Union

Technical BBs – Data Sovereignty & Trust

	Description	Key capabilities	Link to specifications
Access and usage policies and control	Ability to specify and policies within a given data space, by the data space authority and the individual participants	- Access policies specify the conditions to access services and data - Usage Policies specify rights and obligations for the usage of the data, this includes future usages of data - To enable the decision-making process in evaluation policies connection to other building blocks is required for Identification, Claim Management, Authentication and Authorization	Access and usage policies and control
Identity Management	Management of identities within a data space	- Onboarding (and offboarding) participants into a data space, by providing them with an identity of verifying their identity - Issuing, holding and verifying identities	Identity Management
Trust	Being able to verify that a participant of a data space adheres to certain rules	- Semantic to describe and verify the description of service, data products and policies, notably for access control, usage purpose, consent, authorisation and rights delegation - Level of assessment activity - either a declaration or certification - for making specific claim - List of accredited parties eligible to issue identifiers, certifications and cryptographic material such as certificates - Procedure for parties to request and issue identifiers for themselves, services or data products	Trust

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Technical BBs – Data Value Creation

DATA SPACES SUPPORT CENTRE

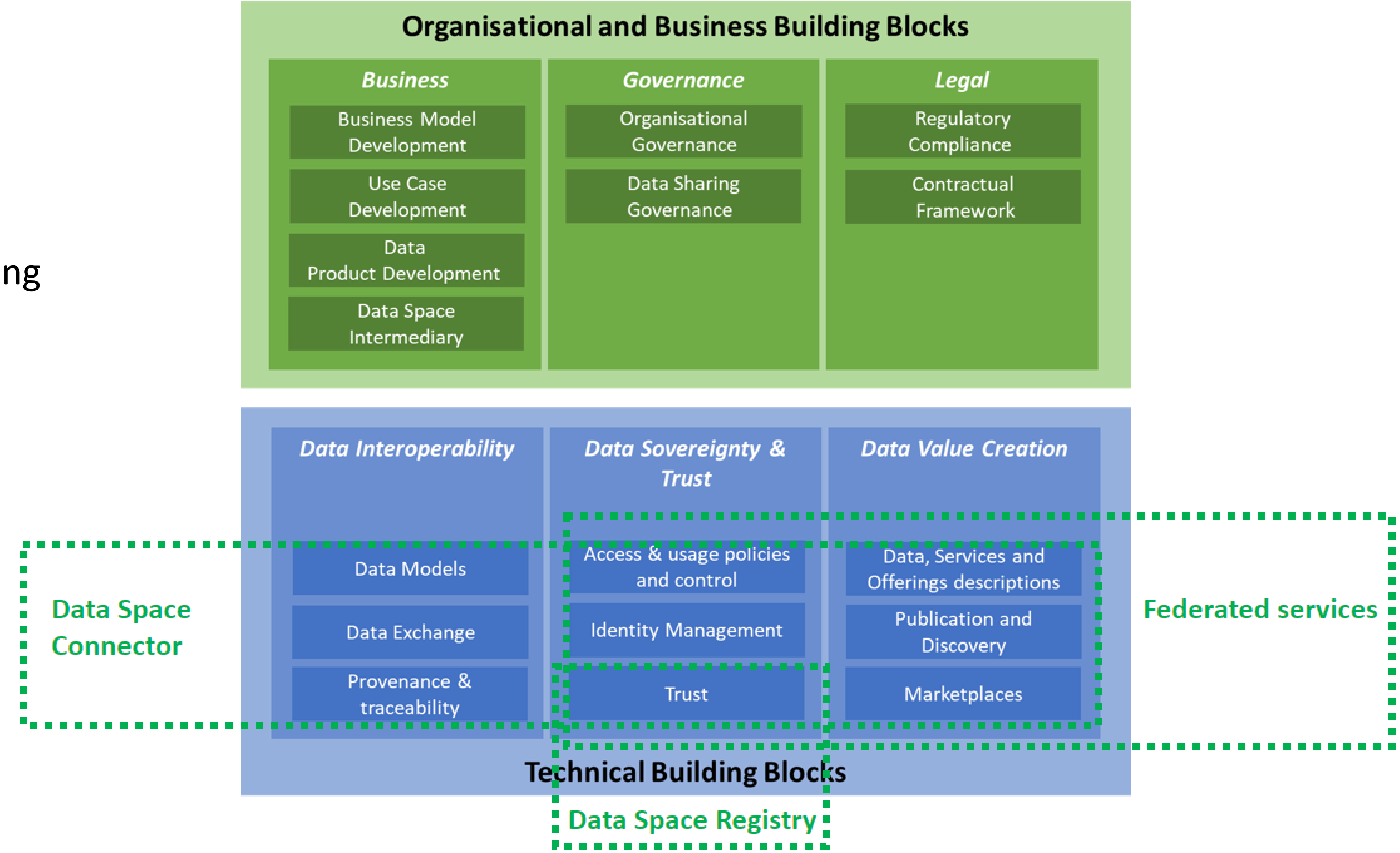
	Description	Key capabilities	Link to specifications
Data, Services and Offerings descriptions	provides to data providers the tools to describe appropriately, and in a complete way, a data product, in a manner that will be understandable by any participant in the data space.	• detailed explanation assists users in understanding the data product content and structure • rich description allows users to get an impression of the data product before conducting any analysis or processing • complete data description is required for evaluating the dataset's quality and dependability • well-documented data description contributes to data governance activities such as data privacy, security, and regulatory compliance	Data, Services and Offerings descriptions
Publication and Discovery	Allows data providers to publish the description of their data, services and offerings, in a way that they can be discovered by users, following the FAIR principles	• Management of self-descriptions • Discovery of self-descriptions • Enable dynamic transactions • Mange access to self-descriptions	Publication and Discovery
Marketplaces	Marketplace capabilities, in such a way that provider and user can enter into a relationship about the access, provision and use of a data product	• Catalogue management • Contract and producto order management • Aftersales assistance and customer service	Marketplaces

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Systems architecture view

- Technology is needed to implement the capabilities of Technical BBs
- Certain services are implementing the BBs
- To facilitate the onboarding of data space participants



Funded by the European Union

How to use the Blueprint

DATA SPACES SUPPORT CENTRE

- When posible follow the proposed structure in your data space
- Do the mapping between your building blocks and DSSC ones
- Identify missing or incomplete functionalities
  - Specific for your domain – refer to the DSSC building block that you are extending in your domain
  - Relevant for other domains – propose the addition to the DSSC to be incorporated to the DSSC taxonomy
- Keep updated about future releases
- Propose implementations (upcoming)

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Summary

- Demand on data sharing to address business and societal challenges
- Data Spaces as instruments to facilitate the data sharing
- Bringing an extra value to the existing ways of data sharing
- Guidelines and tolos are needed for data spaces initiatives
- DSSC is here to support the understanding and development of data spaces
- Providing a set of assets, including a blueprint
- Blueprint v0.5 is out! Go and take it!

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

How to keep posted

DATA SPACES SUPPORT CENTRE

- News and publications at dssc.eu
  - Get support
- Join the Thematic Groups
  - Knowledge base and periodic meetings
  - Get involved
- Monthly DSSC Insights Series
  - Next one about the Blueprint on 12th October at 16:00
- Data Spaces Symposium in March 2024

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

DATA SPACES SUPPORT CENTRE

Thanks!

More info at contact@dssc.eu

The Data Spaces Support Centre receives funding from the European Union Digital Europe Programme under grant agreement n° 101083412.

Funded by the European Union

Exploring the Research Challenges with Data Spaces

Edward Curry

A World Leading SFI Research Centre

Science Foundation Ireland SFI For what's next

Exploring the Research Challenges with Data Spaces

Prof. Edward Curry

University of Galway

Insight SFI research Centre for Data Analytics

Insight

SFI RESEARCH CENTRE FOR DATA ANALYTICS

Data spaces: Discovering the building blocks

Webinar on 6 October 2023

data.europa.eu academy

HOST INSTITUTIONS

DCU · Ollscoil na Gaillimhe University of Galway · UCC University College Cork, Ireland · UCD Dublin

PARTNER INSTITUTIONS

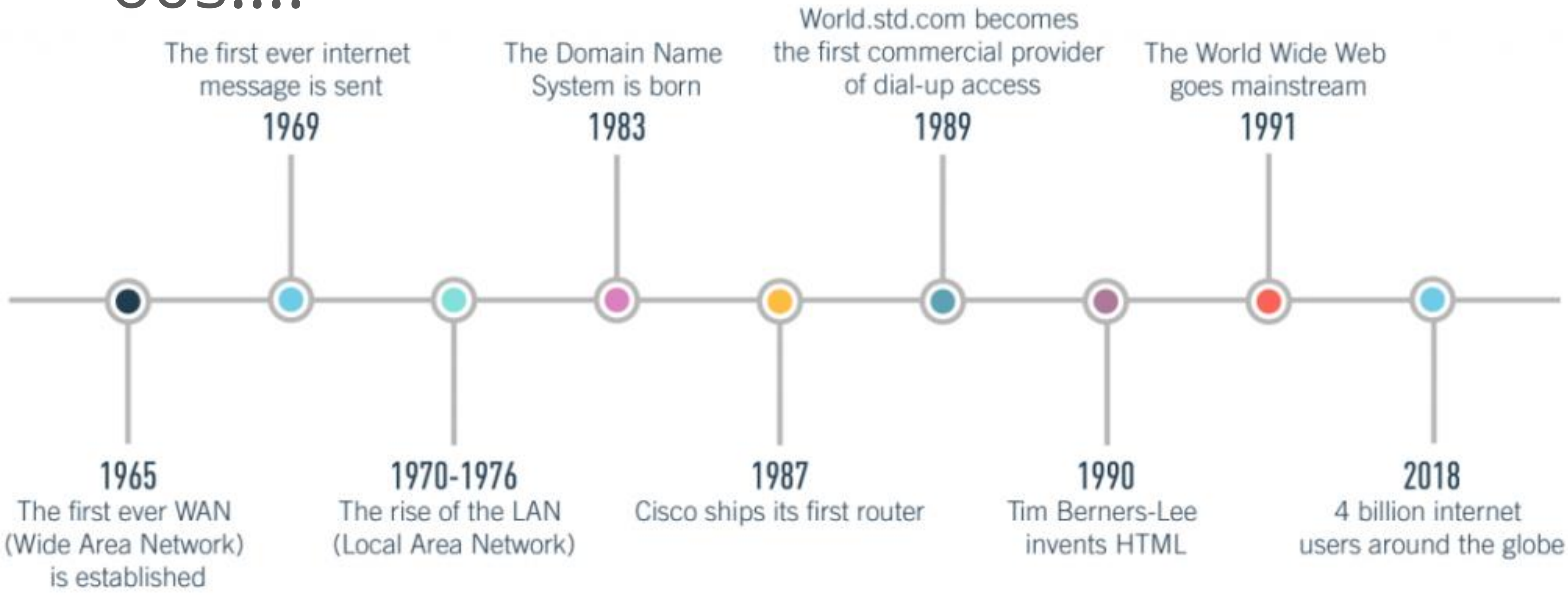
Maynooth University · Tyndall National Institute · Trinity College Dublin · University of Limerick

FUNDED BY:

European Union European Regional Development Fund · Science Foundation Ireland SFI For what's next

Common European Data Spaces

Foundations of the Web go back to the 60s….



1965
The first ever WAN (Wide Area Network) is established

1969
The first ever internet message is sent

1970-1976
The rise of the LAN (Local Area Network)

1983
The Domain Name System is born

1987
Cisco ships its first router

1989
World.std.com becomes the first commercial provider of dial-up access

1990
Tim Berners-Lee invents HTML

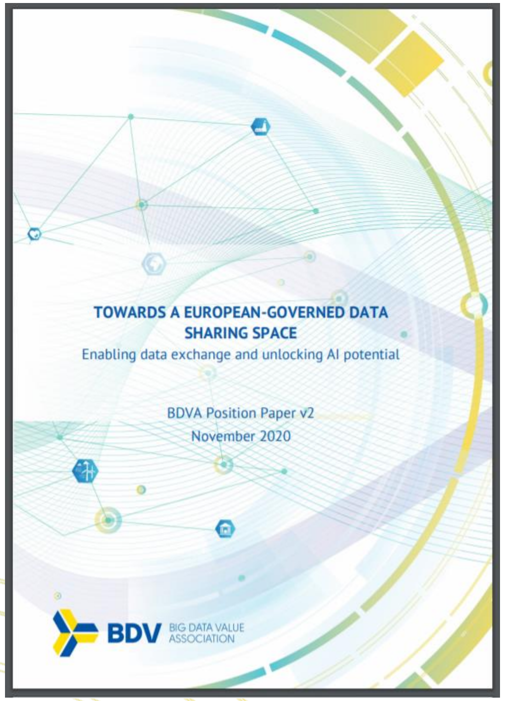
1991
The World Wide Web goes mainstream

2018
4 billion internet users around the globe

Image credit: https://ventcube.com/history-of-the-internet-timeline/

Common European Data Spaces
……..A Ten Year Journey

BDV BIG DATA VALUE ASSOCIATION



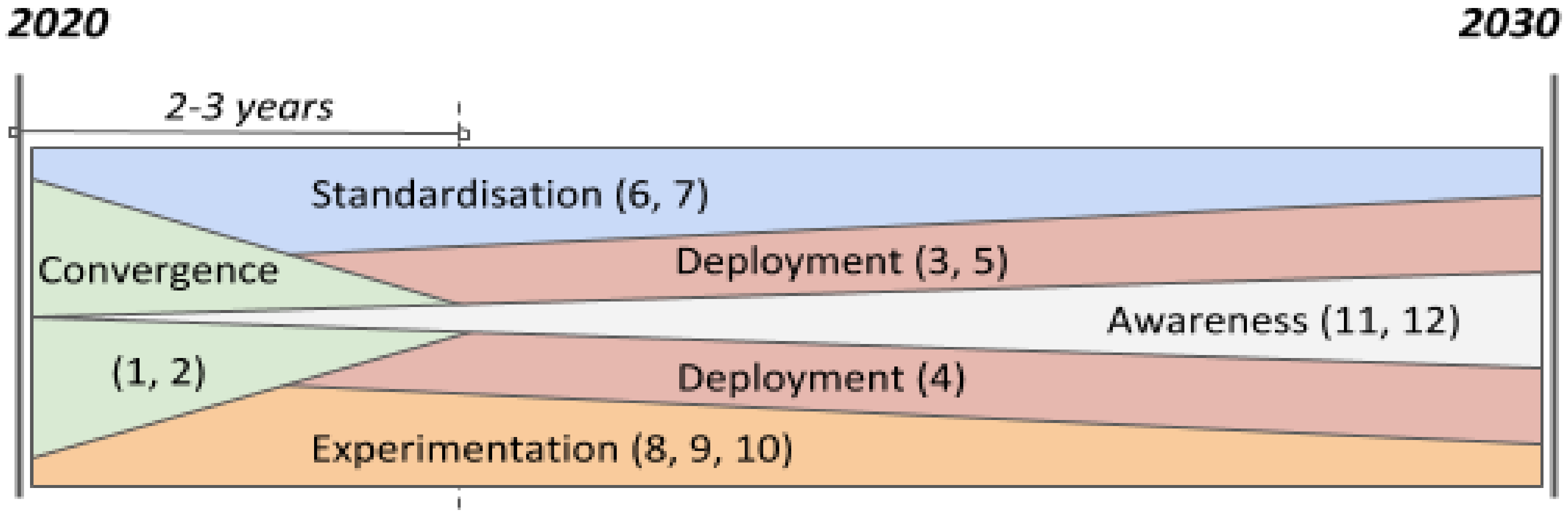
TOWARDS A EUROPEAN-GOVERNED DATA SHARING SPACE

Enabling data exchange and unlocking AI potential

BDVA Position Paper v2

November 2020

BDV BIG DATA VALUE ASSOCIATION



Data Spaces Business Alliance

Unleashing the Data Economy

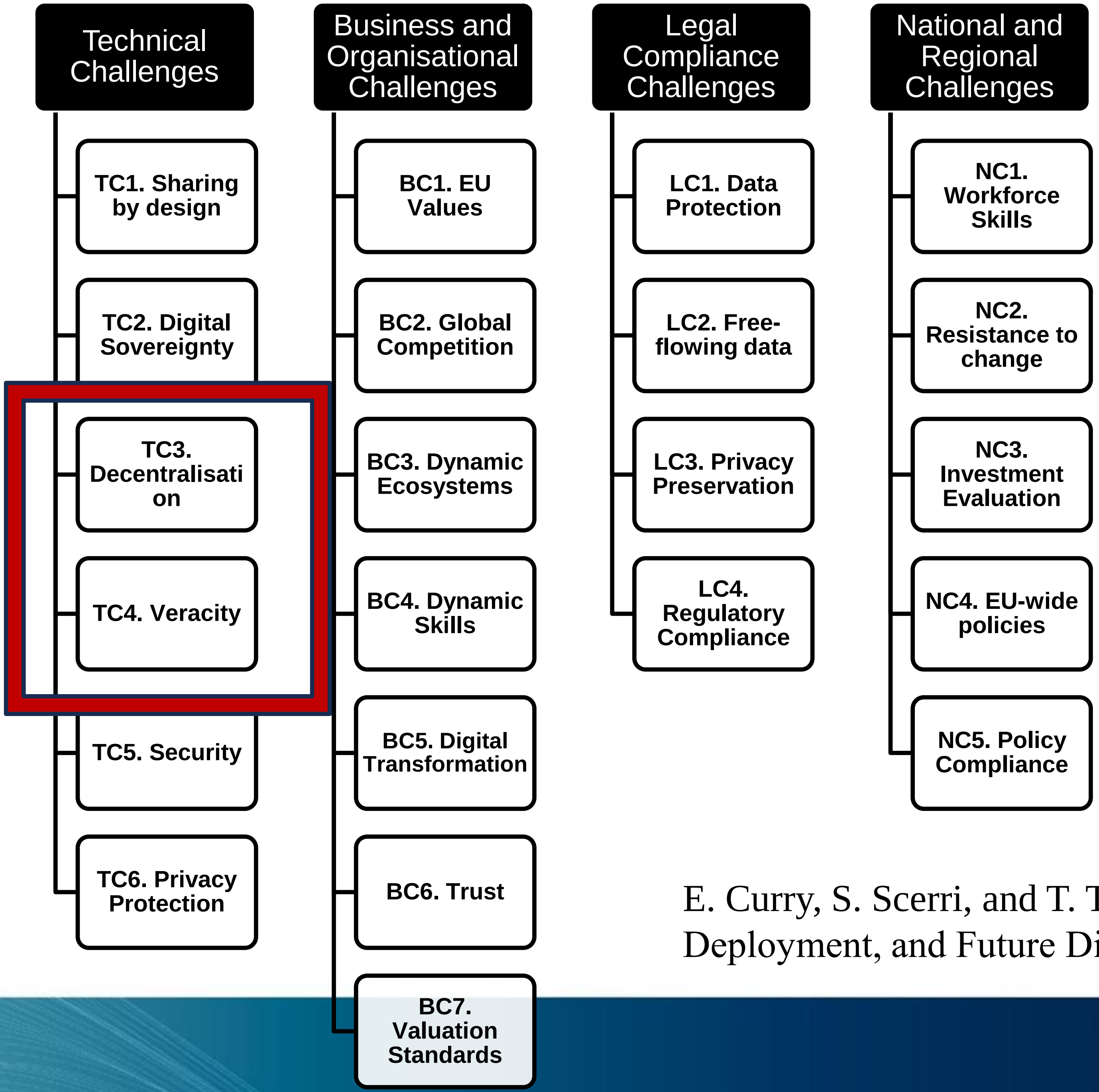
BDV BIG DATA VALUE ASSOCIATION FIWARE FOUNDATION gaia-x INTERNATIONAL DATA SPACES ASSOCIATION

DATA SPACES SUPPORT CENTRE

Next Generation Data Spaces

interoperable europe

Research Challenges



Technical Challenges

- TC1. Sharing by design
- TC2. Digital Sovereignty
- TC3. Decentralisati on
- TC4. Veracity
- TC5. Security
- TC6. Privacy Protection

Business and Organisational Challenges

- BC1. EU Values
- BC2. Global Competition
- BC3. Dynamic Ecosystems
- BC4. Dynamic Skills
- BC5. Digital Transformation
- BC6. Trust
- BC7. Valuation Standards

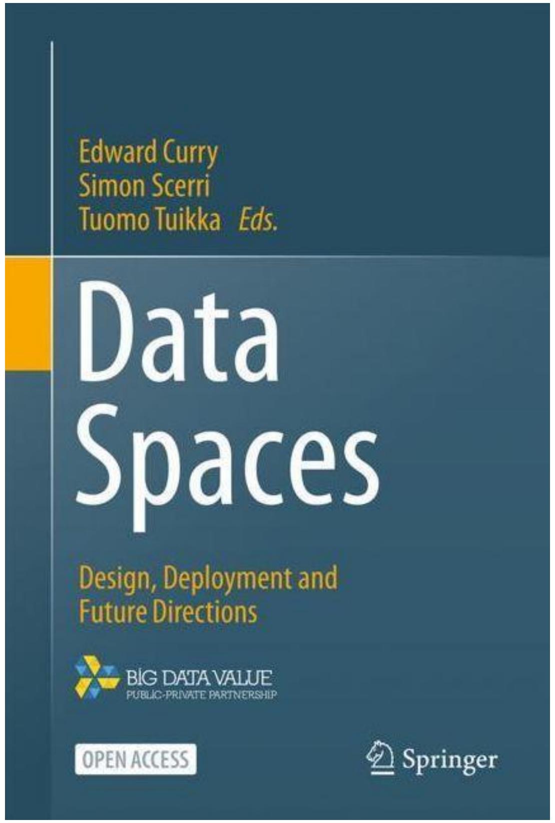
Legal Compliance Challenges

- LC1. Data Protection
- LC2. Free-flowing data
- LC3. Privacy Preservation
- LC4. Regulatory Compliance

National and Regional Challenges

- NC1. Workforce Skills
- NC2. Resistance to change
- NC3. Investment Evaluation
- NC4. EU-wide policies
- NC5. Policy Compliance

OPEN ACCESS



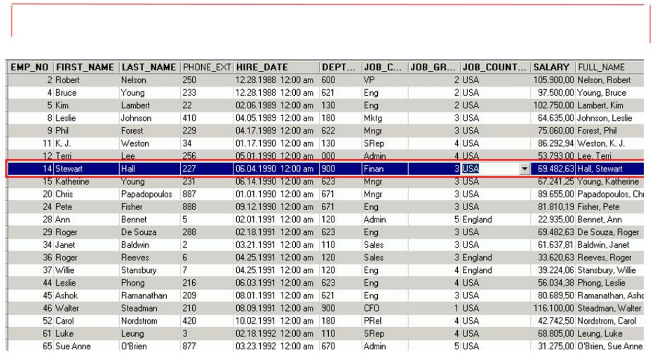
E. Curry, S. Scerri, and T. Tuikka, “Data Spaces: Design, Deployment, and Future Directions,” in *Data Spaces*, 2022.

Expanding Content Space...

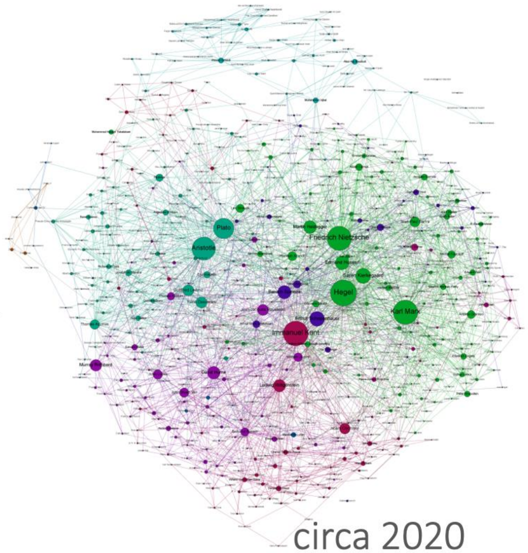
- Heterogeneous, complex and large-scale data
- Very-large and dynamic “schemas”
- Open Environments: distributed, decentralised decoupled data sources, anonymous users, multi-domain, lack of global order of information flow

circa 2000

10s-100s attributes



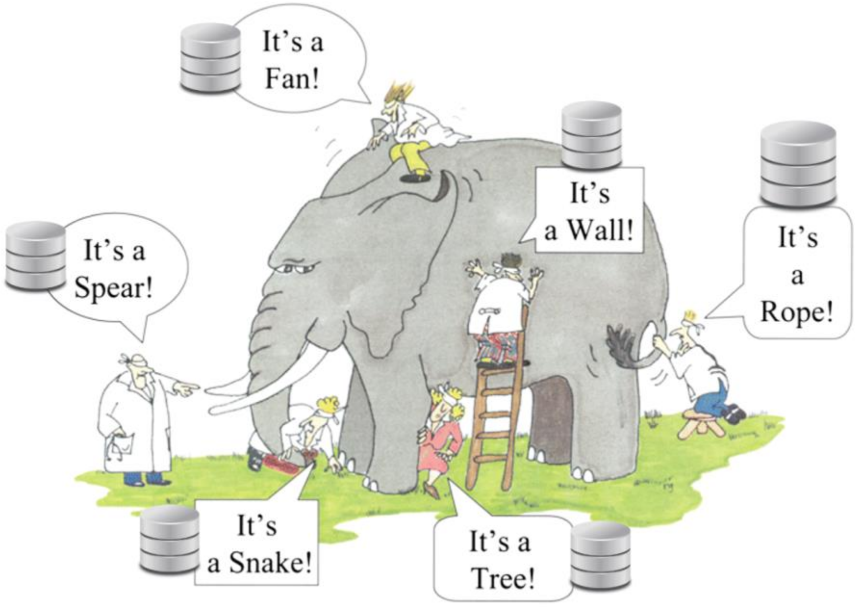
EMP_NO	FIRST_NAME	LAST_NAME	PHONE_EXT	HIRE_DATE	DEPT...	JOB_C...	JOB_GR...	JOB_COUNT...	SALARY	FULL_NAME
2	Robert	Nelson	250	12.28.1988 12:00 am	600	VP	2	USA	105.900,00	Nelson, Robert
4	Bruce	Young	233	12.28.1988 12:00 am	621	Eng	2	USA	97.500,00	Young, Bruce
5	Kim	Lambert	22	02.06.1989 12:00 am	130	Eng	2	USA	102.750,00	Lambert, Kim
8	Leslie	Johnson	410	04.05.1989 12:00 am	180	Mktg	3	USA	64.635,00	Johnson, Leslie
9	Phil	Forest	229	04.17.1989 12:00 am	622	Mngr	3	USA	75.060,00	Forest, Phil
11	K. J.	Weston	34	01.17.1990 12:00 am	130	SRep	4	USA	86.292,94	Weston, K. J.
12	Terri	Lee	256	05.01.1990 12:00 am	000	Admin	4	USA	53.793,00	Lee, Terri
14	Stewart	Hall	227	06.04.1990 12:00 am	900	Finan	3	USA	69.482,63	Hall, Stewart
15	Katherine	Young	231	06.14.1990 12:00 am	623	Mngr	3	USA	67.241,25	Young, Katherine
20	Chris	Papadopoulos	887	01.01.1990 12:00 am	671	Mngr	3	USA	89.655,00	Papadopoulos, Chr
24	Pete	Fisher	888	09.12.1990 12:00 am	671	Eng	3	USA	81.810,19	Fisher, Pete
28	Ann	Bennet	5	02.01.1991 12:00 am	120	Admin	5	England	22.935,00	Bennet, Ann
29	Roger	De Souza	288	02.18.1991 12:00 am	623	Eng	3	USA	69.482,63	De Souza, Roger
34	Janet	Baldwin	2	03.21.1991 12:00 am	110	Sales	3	USA	61.637,81	Baldwin, Janet
36	Roger	Reeves	6	04.25.1991 12:00 am	120	Sales	3	England	33.620,63	Reeves, Roger
37	Willie	Stansbury	7	04.25.1991 12:00 am	120	Eng	4	England	39.224,06	Stansbury, Willie
44	Leslie	Phong	216	06.03.1991 12:00 am	623	Eng	4	USA	56.034,38	Phong, Leslie
45	Ashok	Ramanathan	209	08.01.1991 12:00 am	621	Eng	3	USA	80.689,50	Ramanathan, Ashc
46	Walter	Steadman	210	08.09.1991 12:00 am	900	CFO	1	USA	116.100,00	Steadman, Walter
52	Carol	Nordstrom	420	10.02.1991 12:00 am	180	PRel	4	USA	42.742,50	Nordstrom, Carol
61	Luke	Leung	3	02.18.1992 12:00 am	110	SRep	4	USA	68.805,00	Leung, Luke
65	Sue Anne	O'Brien	877	03.23.1992 12:00 am	670	Admin	5	USA	31.275,00	O'Brien, Sue Anne



circa 2020

1,000s-1,000,000s attributes

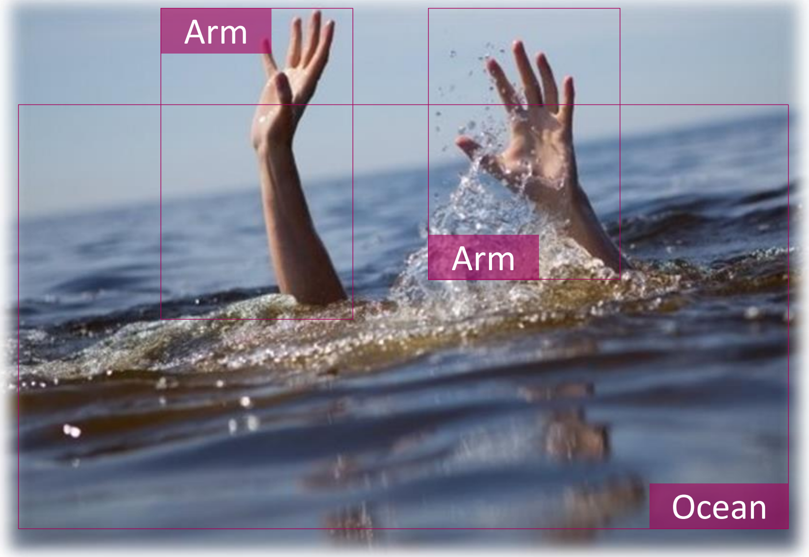
- Multiple perspectives (conceptualisations) of the reality.
- Ambiguity, vagueness, inconsistency.



... Fundamental Decentralisation

Increasing Amounts of Multimodal Data….

(More than 220k people in Ireland are blind or visually impaired)



(An average of 110 drownings occur in Ireland every year)

(Every 3 minutes in Ireland someone gets a cancer diagnosis; every hour someone dies from cancer)

.J. Khan, J.G. Breslin, E. Curry, “Expressive Scene Graph Generation using Commonsense Knowledge Infusion for Visual Understanding and Reasoning,” Extended Semantic Web Conference (ESWC) 2022, Crete, Greece, May-June 2022.

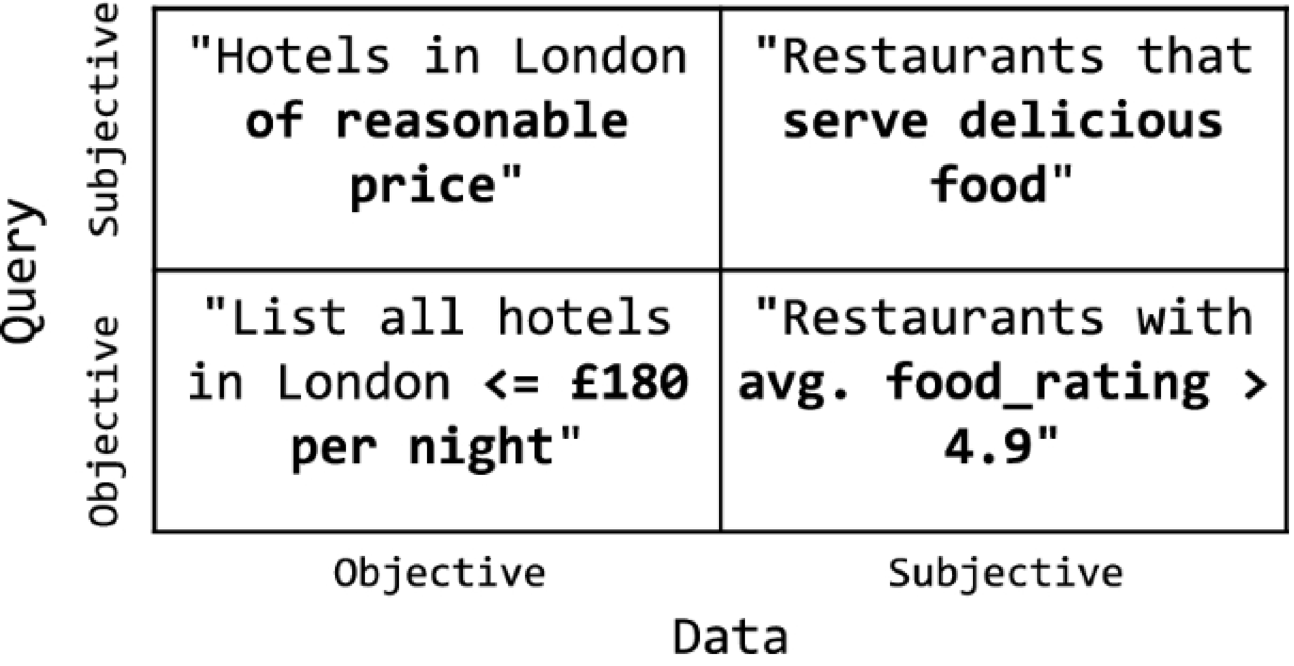
Increasing amounts of Subjective Data…

$
OBJECTIVE REALITY
SUBJECTIVE INTERPRETATION

> “The very concept of objective truth is fading out of the world.”
>
> *George Orwell*

Image Credit: https://www.yesvedanta.com/objectivity-vs-subjectivity-in-light-of-vedanta/

Subjective and Objective Attributes and Query



Query	Subjective	"Hotels in London **of reasonable price"**	"Restaurants that **serve delicious food"**
	Objective	"List all hotels in London **<= £180 per night"**	"Restaurants with **avg. food_rating > 4.9"**
		Objective	Subjective
		Data	

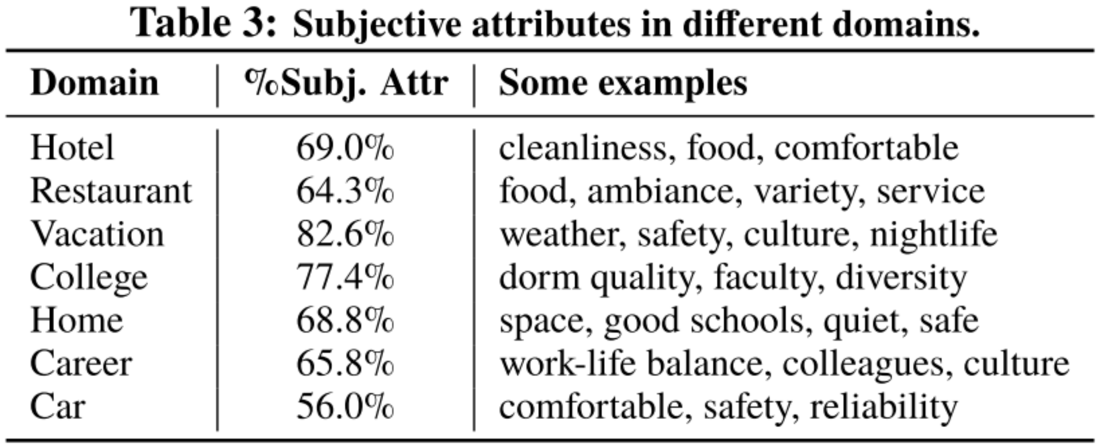


Table 3: **Subjective attributes in different domains.**

Domain	**%Subj. Attr**	**Some examples**
Hotel	69.0%	cleanliness, food, comfortable
Restaurant	64.3%	food, ambiance, variety, service
Vacation	82.6%	weather, safety, culture, nightlife
College	77.4%	dorm quality, faculty, diversity
Home	68.8%	space, good schools, quiet, safe
Career	65.8%	work-life balance, colleagues, culture
Car	56.0%	comfortable, safety, reliability

Li, Y., Feng, A., Li, J. *et al.* Querying subjective data. *The VLDB Journal* **30,** 115–140 (2021). https://doi.org/10.1007/s00778-020-00634-5

The Red Queen Hypothesis

"It takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"

Lewis Carroll's Through the Looking-Glass

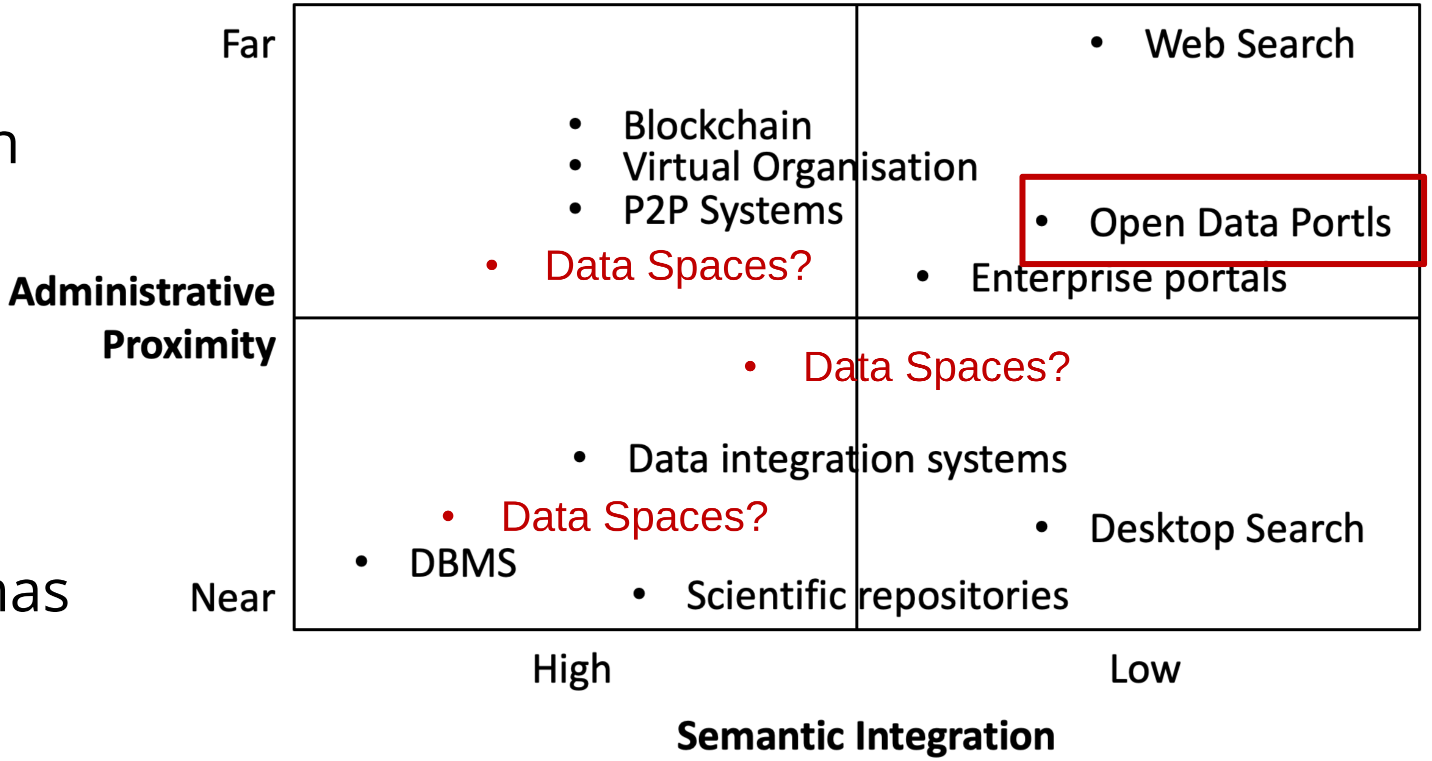
Control and Coordination

Administrative Proximity

- Close vs. Loose Coordination
- Assumptions concerning guarantees such as data, access, quality, and consistency,

Semantic Integration

- Degree to which data schemas are matched up (types, attributes, and names).



Far

Administrative Proximity

Near

- Web Search
- Blockchain
- Virtual Organisation
- P2P Systems
- Data Spaces?
- Open Data Portls
- Enterprise portals
- Data Spaces?
- Data integration systems
- Data Spaces?
- DBMS
- Desktop Search
- Scientific repositories

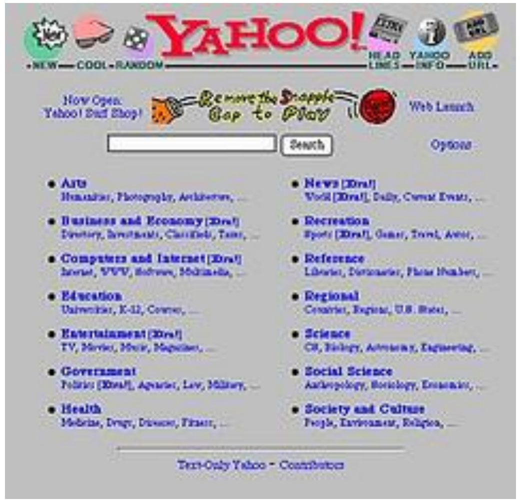
High

Low

Semantic Integration

Halevy, A., Franklin, M. and Maier, D. 2006. Principles of dataspace systems. *25th ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems - PODS ’06* (New York, New York, USA, 2006), 1–9.

Evolving Human Interactivity…

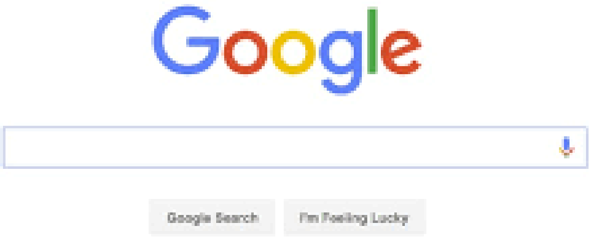


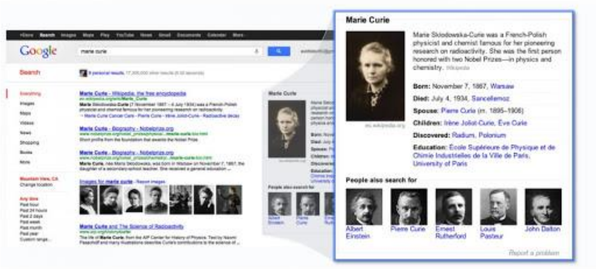
Now Open: Yahoo! Surf Shop!

Web Launch

- **Arts**
  Humanities, Photography, Architecture, ...
- **Business and Economy** [Xtra!]
  Directory, Investments, Classifieds, Taxes, ...
- **Computers and Internet** [Xtra!]
  Internet, WWW, Software, Multimedia, ...
- **Education**
  Universities, K-12, Courses, ...
- **Entertainment** [Xtra!]
  TV, Movies, Music, Magazines, ...
- **Government**
  Politics [Xtra!], Agencies, Law, Military, ...
- **Health**
  Medicine, Drugs, Diseases, Fitness, ...
- **News** [Xtra!]
  World [Xtra!], Daily, Current Events, ...
- **Recreation**
  Sports [Xtra!], Games, Travel, Autos, ...
- **Reference**
  Libraries, Dictionaries, Phone Numbers, ...
- **Regional**
  Countries, Regions, U.S. States, ...
- **Science**
  CS, Biology, Astronomy, Engineering, ...
- **Social Science**
  Anthropology, Sociology, Economics, ...
- **Society and Culture**
  People, Environment, Religion, ...

Text-Only Yahoo - Contributors





Marie Curie

Marie Skłodowska-Curie was a French-Polish physicist and chemist famous for her pioneering research on radioactivity. She was the first person honored with two Nobel Prizes—in physics and chemistry. Wikipedia

Born: November 7, 1867, Warsaw

Died: July 4, 1934, Sancellemoz

Spouse: Pierre Curie (m. 1895–1906)

Children: Irène Joliot-Curie, Ève Curie

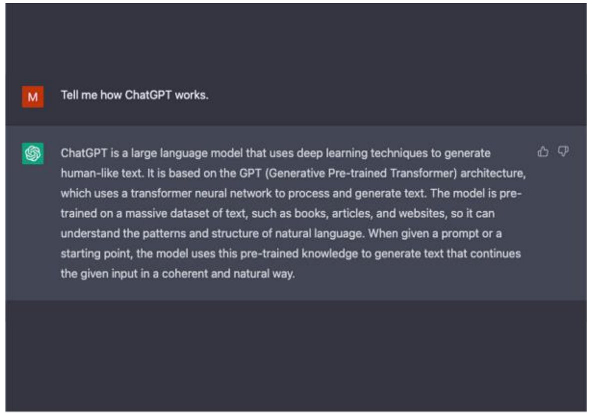
Discovered: Radium, Polonium

Education: École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris, University of Paris

People also search for

Albert Einstein, Pierre Curie, Ernest Rutherford, Louis Pasteur, John Dalton

Report a problem



M Tell me how ChatGPT works.

ChatGPT is a large language model that uses deep learning techniques to generate human-like text. It is based on the GPT (Generative Pre-trained Transformer) architecture, which uses a transformer neural network to process and generate text. The model is pre-trained on a massive dataset of text, such as books, articles, and websites, so it can understand the patterns and structure of natural language. When given a prompt or a starting point, the model uses this pre-trained knowledge to generate text that continues the given input in a coherent and natural way.

From Structure → **to Search** → **to Knowledge Graph** → **to Conversations ?**

From Structure

~1995
~100K Websites
Exact Results
Human Curated

to Search

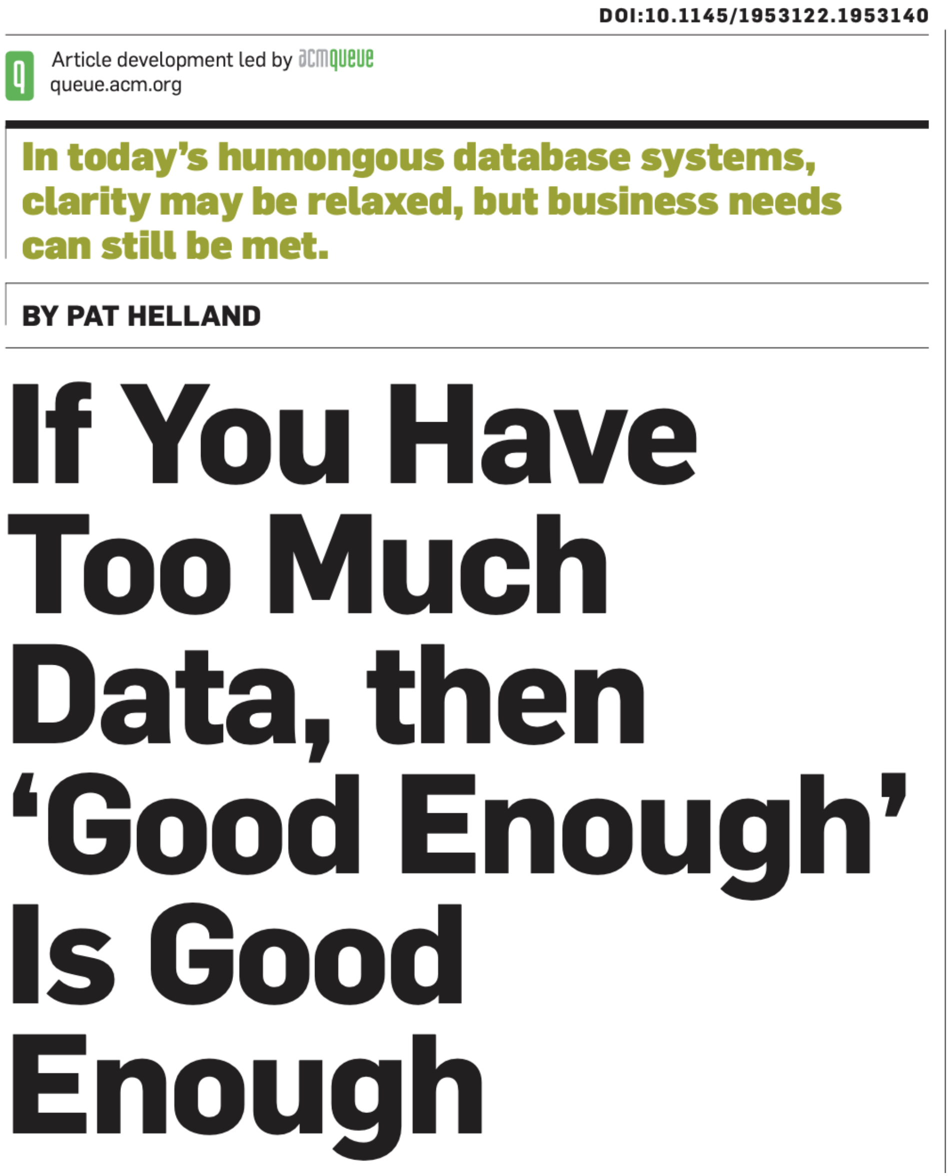
~1998
~2.4M Websites
Approximate Results
Computed

to Knowledge Graph

~2012
~700M
Approximate Results + Exact
Computed + Crowd

to Conversations ?

~2023
~2B
Approximate Results ?
?



DOI:10.1145/1953122.1953140

Article development led by acmqueue
queue.acm.org

In today's humongous database systems, clarity may be relaxed, but business needs can still be met.

BY PAT HELLAND

If You Have Too Much Data, then 'Good Enough' Is Good Enough

"We can no longer pretend to live in a clean world. SQL and its Data Definition Language (DDL) assume a crisp and clear definition of the data, but that is a subset of the business examples we see in the world around us. It's OK if we have lossy answers—that's frequently what business needs."

What is a Data Space?

“Dataspaces are not a data integration approach; rather, they are more of a **data co-existence approach**. The goal of dataspace support is to provide base functionality over all data sources, regardless of how integrated they are.” (Halevy, A., Franklin, M. and Maier, D. 2006.)

Incrementalism, Approximate, Interactive

Data Space Enablers

Edward Curry

Real-time Linked Dataspaces

Enabling Data Ecosystems for Intelligent Systems

SpringerOpen

OPEN ACCESS

Data Space Support Platform (Halevy et al.)

- Must deal with **many different formats** of data.
- **Does not subsume** existing systems; they still provide individual access via their native interfaces.
- Queries in are provided on a **best-effort and approximate basis**.
- Must provide **pathways to improve the integration** among the data sources, including streams and events, in a pay-as-you-go fashion.

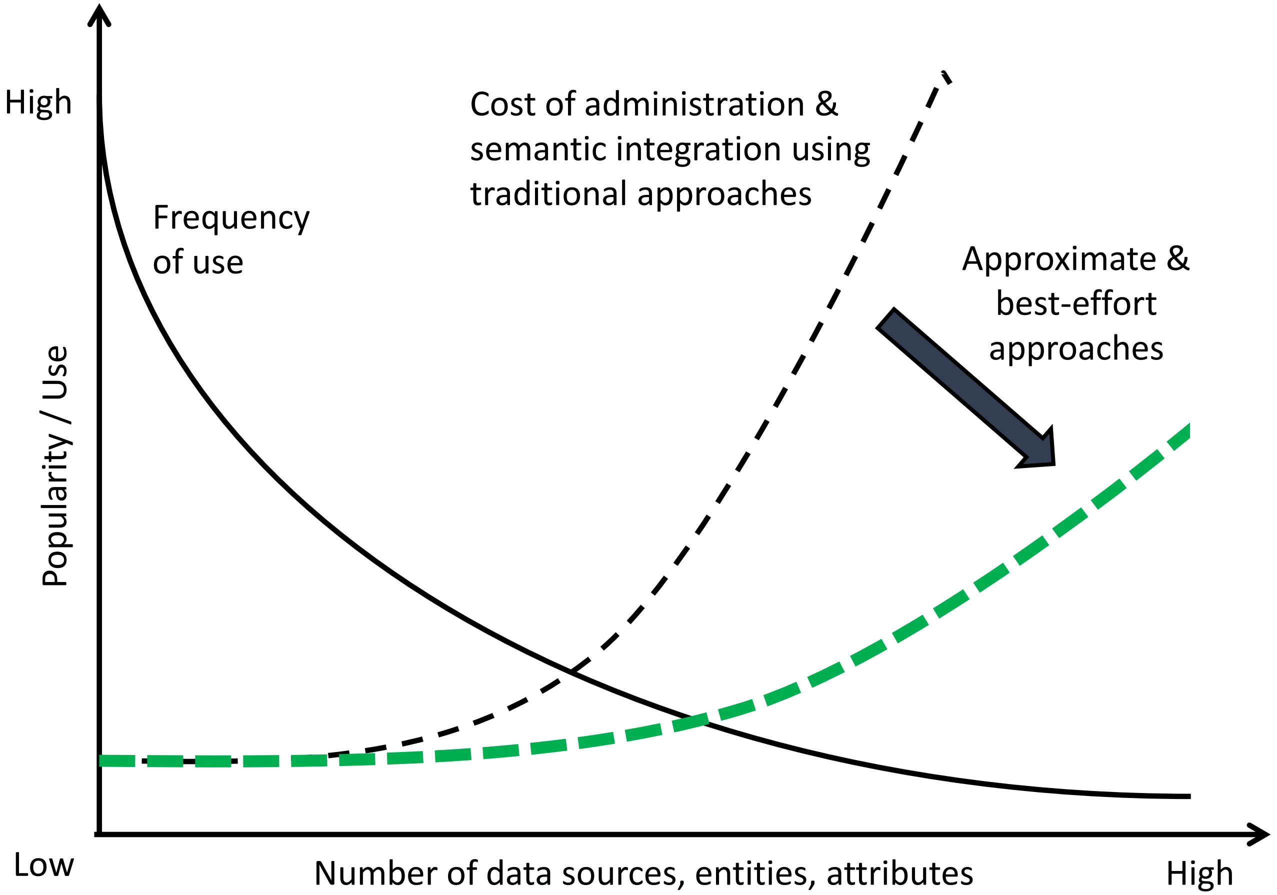
Data Space Support Services

- Catalog and Browse
- Search and Query
- Local Store and Index
- Discovery
- Enhancement
- Administration

Reusing Human Attention

- Learn from users' activities
    - Create meaningful relationships between data sources
    - Enhance data sources

Approximate and Best Effort Approaches



The Long Tail of Data

Creating Approximate Services to support incremental data management is a key challenge…

Edward Curry

Real-time Linked Dataspaces

Enabling Data Ecosystems for Intelligent Systems

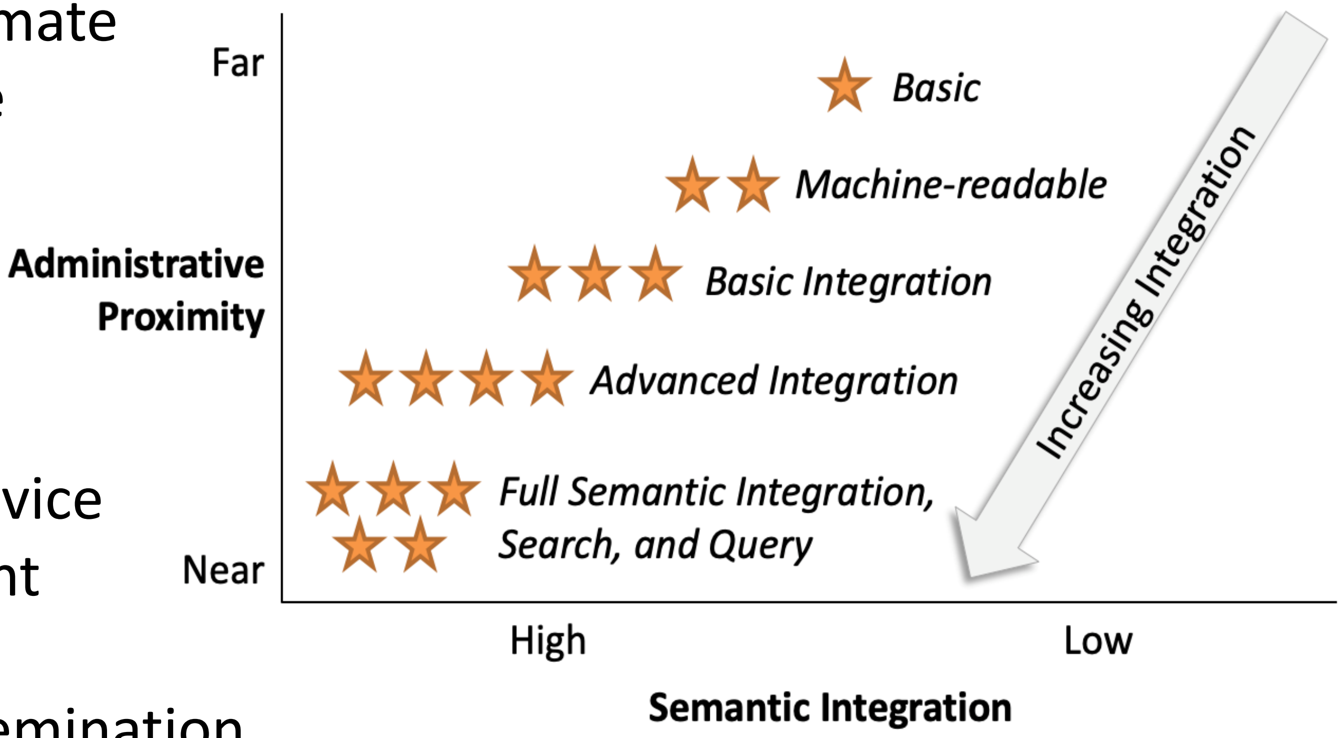
Springer Open

OPEN ACCESS

Investigate techniques to enable approximate and best-effort support services for loose administrative proximity and semantic integration

Incremental support services

- Catalog
- entity management
- query and search
- data discovery
- human tasks
- quality of service
- complex event processing
- streams dissemination
- approximate semantic event matching



Curry, E. 2020. Real-time Linked Dataspaces. Springer International Publishing.

http://dataspaces.info

Research Trends

Hybrid Neuro-Symbolic Approaches

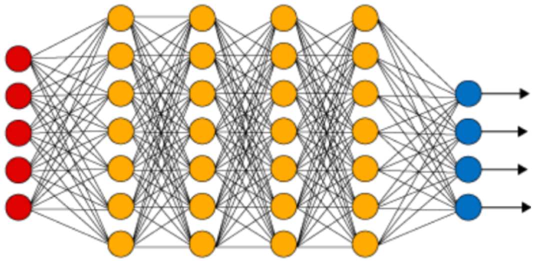
- We need Semantics, not just statistics!

- Combing rules-based AI approaches (Knowledge Graphs) with statistical learning techniques (deep learning)

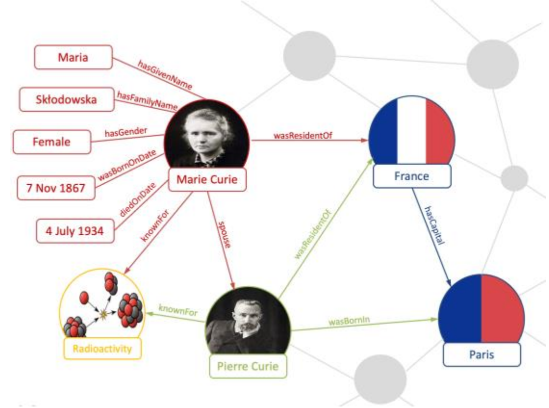
Neuro-Symbolic AI

Neural Networks
(Statistical Learning + Reasoning)

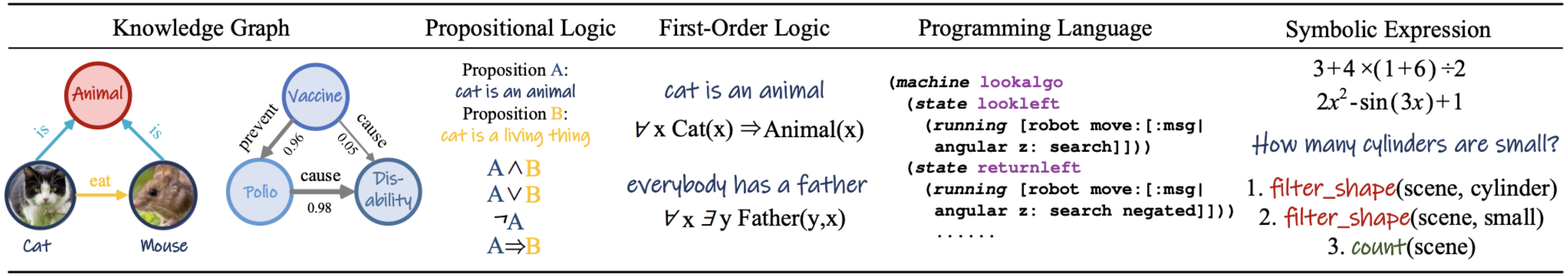
Knowledge Graphs
(Expert Knowledge + Reasoning)



+



Examples of Symbolic Knowledge



```
(machine lookalgo
  (state lookleft
    (running [robot move:[:msg|
     angular z: search]]))
  (state returnleft
    (running [robot move:[:msg|
     angular z: search negated]]))
     ......
```

Fig. 8. Illustrative overview of symbolic knowledge representations in NeSy.

- Neural- Symbolic Integration
- Knowledge Representation
- Knowledge Embedding

Wang, Wenguan, and Yi Yang. "Towards data-and knowledge-driven artificial intelligence: A survey on neuro-symbolic computing." *arXiv preprint arXiv:2210.15889* (2022)

Examples of NeSy Systems

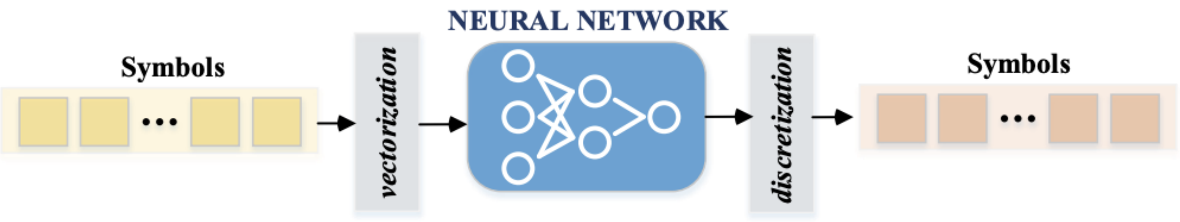


Fig. 2. Type 1: Symbolic Neuro Symbolic.

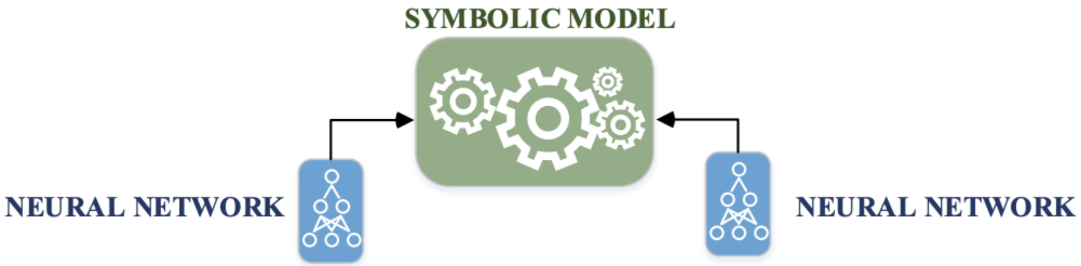


Fig. 3. Type 2: Symbolic[Neuro].

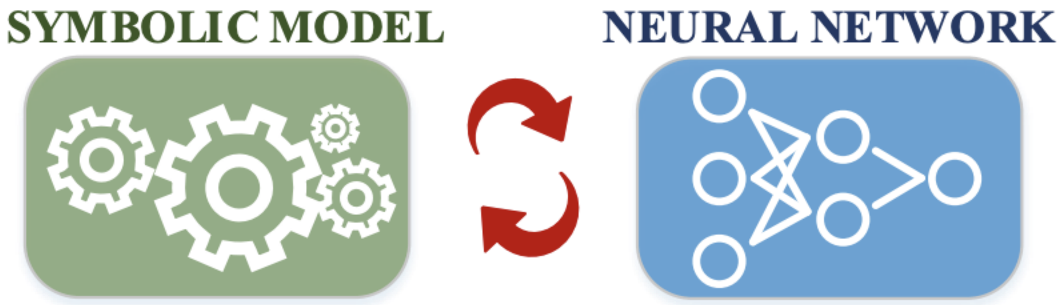


Fig. 4. Type 3: Neuro|Symbolic.

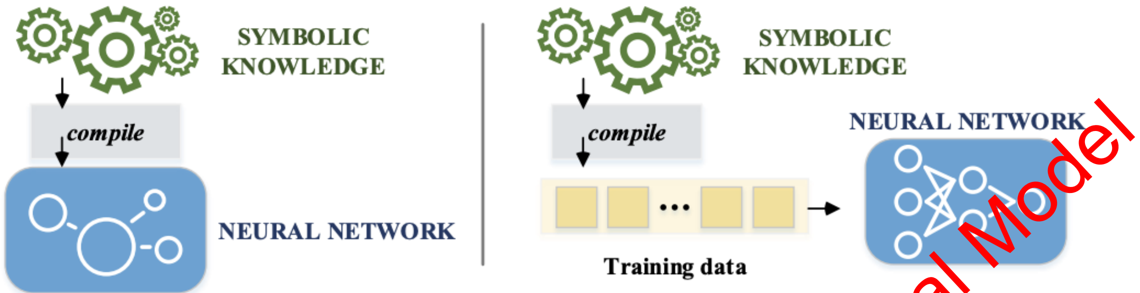


Fig. 5. Type 4: Neuro:Symbolic→Neuro.

Foundational Model

Wang, Wenguan, and Yi Yang. "Towards data-and knowledge-driven artificial intelligence: A survey on neuro-symbolic computing." *arXiv preprint arXiv:2210.15889* (2022)

What is a Foundation Model?

Definition

- *Any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks*

- Foundational models can be applied to a single modality while others are multimodal
    - BERT, GPT-3, GPT-4, LLaMA, Segment Anything Model, CLIP and Stable Diffusion

- **Emergence:** the behaviour of a system to be implicitly induced rather than explicitly constructed
- **Homogenization:** the consolidation of methodologies for building machine learning systems across a wide range of applications

On the Opportunities and Risks of Foundation Models

Rishi Bommasani* Drew A. Hudson Ehsan Adeli Russ Altman Simran Arora
Sydney von Arx Michael S. Bernstein Jeannette Bohg Antoine Bosselut Emma Brunskill
Erik Brynjolfsson Shyamal Buch Dallas Card Rodrigo Castellon Niladri Chatterji
Annie Chen Kathleen Creel Jared Quincy Davis Dorottya Demszky Chris Donahue
Moussa Doumbouya Esin Durmus Stefano Ermon John Etchemendy Kawin Ethayarajh
Li Fei-Fei Chelsea Finn Trevor Gale Lauren Gillespie Karan Goel Noah Goodman
Shelby Grossman Neel Guha Tatsunori Hashimoto Peter Henderson John Hewitt
Daniel E. Ho Jenny Hong Kyle Hsu Jing Huang Thomas Icard Saahil Jain
Dan Jurafsky Pratyusha Kalluri Siddharth Karamcheti Geoff Keeling Fereshte Khani
Omar Khattab Pang Wei Koh Mark Krass Ranjay Krishna Rohith Kuditipudi
Ananya Kumar Faisal Ladhak Mina Lee Tony Lee Jure Leskovec Isabelle Levent
Xiang Lisa Li Xuechen Li Tengyu Ma Ali Malik Christopher D. Manning
Suvir Mirchandani Eric Mitchell Zanele Munyikwa Suraj Nair Avanika Narayan
Deepak Narayanan Ben Newman Allen Nie Juan Carlos Niebles Hamed Nilforoshan
Julian Nyarko Giray Ogut Laurel Orr Isabel Papadimitriou Joon Sung Park Chris Piech
Eva Portelance Christopher Potts Aditi Raghunathan Rob Reich Hongyu Ren
Frieda Rong Yusuf Roohani Camilo Ruiz Jack Ryan Christopher Ré Dorsa Sadigh
Shiori Sagawa Keshav Santhanam Andy Shih Krishnan Srinivasan Alex Tamkin
Rohan Taori Armin W. Thomas Florian Tramèr Rose E. Wang William Wang Bohan Wu
Jiajun Wu Yuhuai Wu Sang Michael Xie Michihiro Yasunaga Jiaxuan You Matei Zaharia
Michael Zhang Tianyi Zhang Xikun Zhang Yuhui Zhang Lucia Zheng Kaitlyn Zhou
Percy Liang*[1]

Center for Research on Foundation Models (CRFM)
Stanford Institute for Human-Centered Artificial Intelligence (HAI)
Stanford University

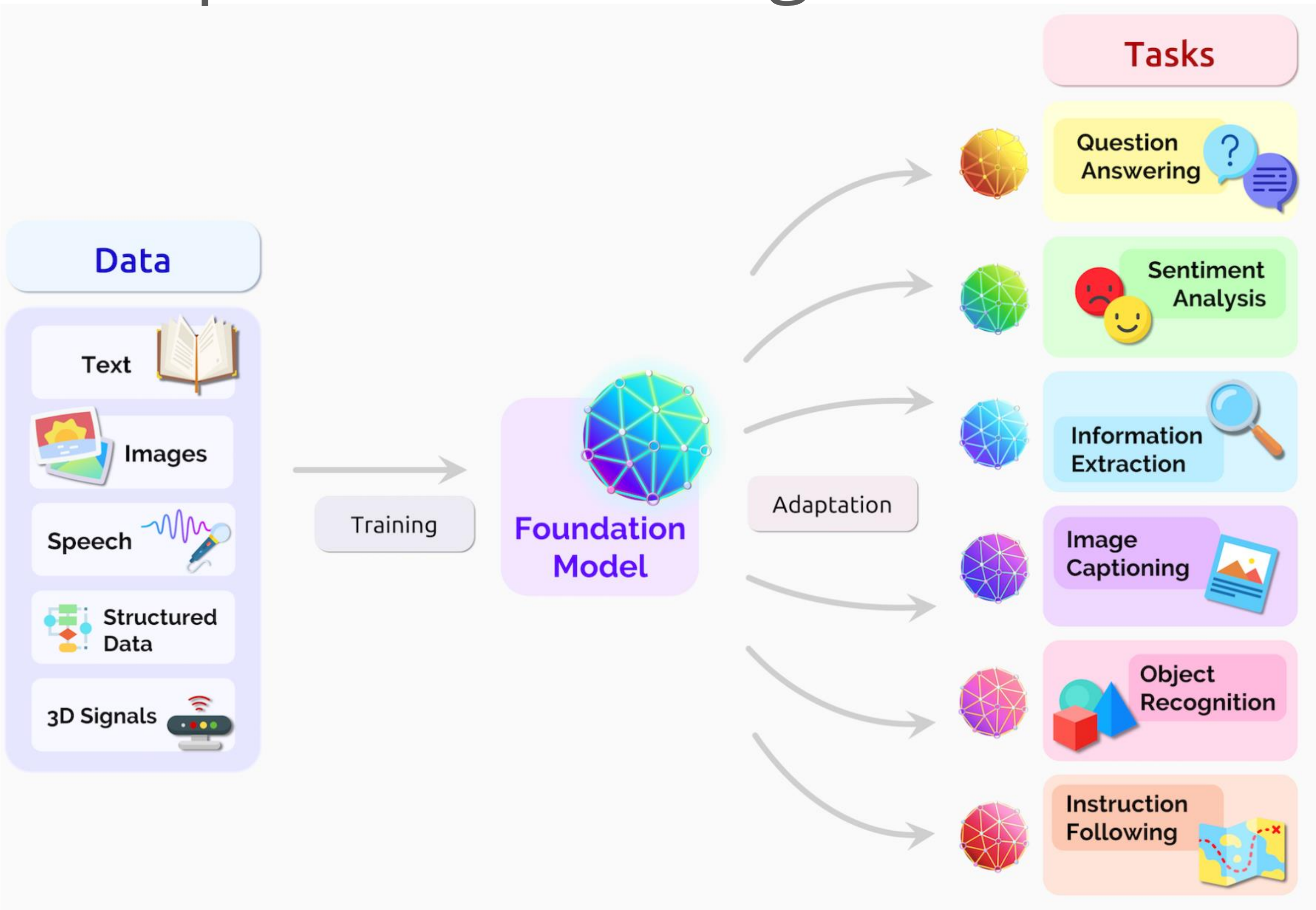
AI is undergoing a paradigm shift with the rise of models (e.g., BERT, DALL-E, GPT-3) trained on broad data (generally using self-supervision at scale) that can be adapted to a wide range of downstream tasks. We call these models foundation models *to underscore their critically central yet incomplete character. This report provides a thorough account of the opportunities and risks of foundation models, ranging from their capabilities (e.g., language, vision, robotic manipulation, reasoning, human interaction) and technical principles (e.g., model architectures, training procedures, data, systems, security, evaluation, theory) to their applications (e.g., law, healthcare, education) and societal impact (e.g., inequity, misuse, economic and environmental impact, legal and ethical considerations). Though foundation models are based on standard deep learning and transfer learning, their scale results in new emergent capabilities, and their effectiveness across so many tasks incentivizes homogenization. Homogenization provides powerful leverage but demands caution, as the defects of the foundation model are inherited by all the adapted models downstream. Despite the impending widespread deployment of foundation models, we currently lack a clear understanding of how they work, when they fail, and what they are even capable of due to their emergent properties. To tackle these questions, we believe much of the critical research on foundation models will require deep interdisciplinary collaboration commensurate with their fundamentally sociotechnical nature.*

[1] Corresponding author: pliang@cs.stanford.edu

*Equal contribution.

Bommasani, R. et al. 2021. On the Opportunities and Risks of Foundation Models. (Aug. 2021).

Centralize information from various modalities, then adapt to a wide range of downstream tasks….

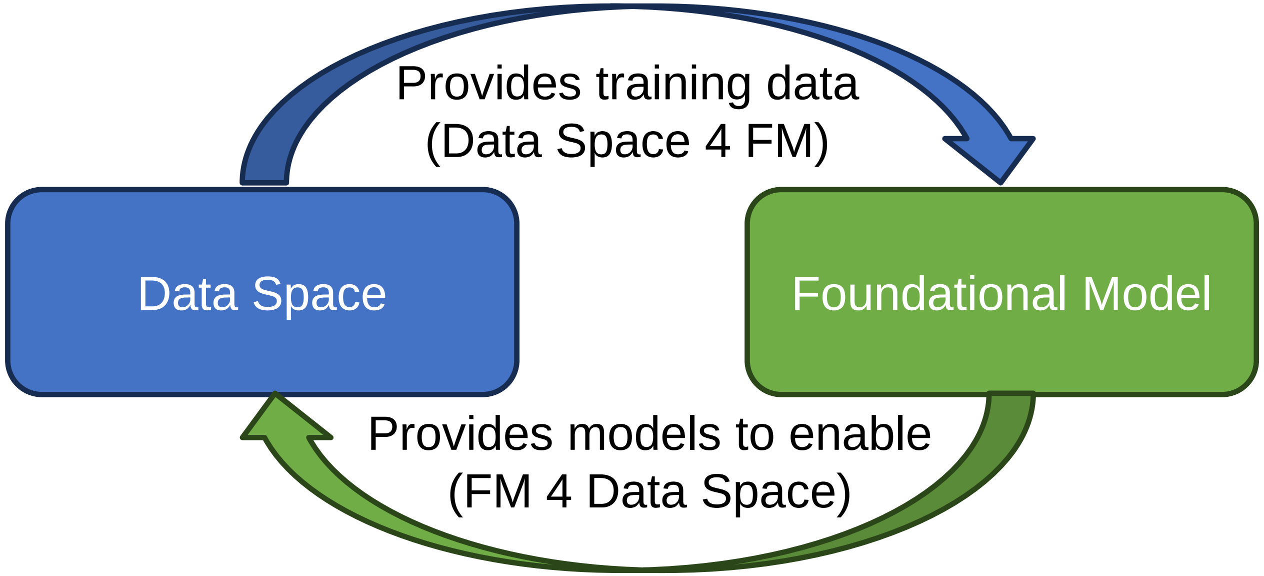


- Trained on extensive datasets to establish a knowledge foundation
- Leverage transfer learning and scale in various downstream tasks.

Bommasani, R. et al. 2021. On the Opportunities and Risks of Foundation Models. (Aug. 2021).

A Symbiotic Relationship…..

High Quality Data
Community of Data Users/Owners
Data Goverance/compliance



Provides training data
(Data Space 4 FM)

Data Space

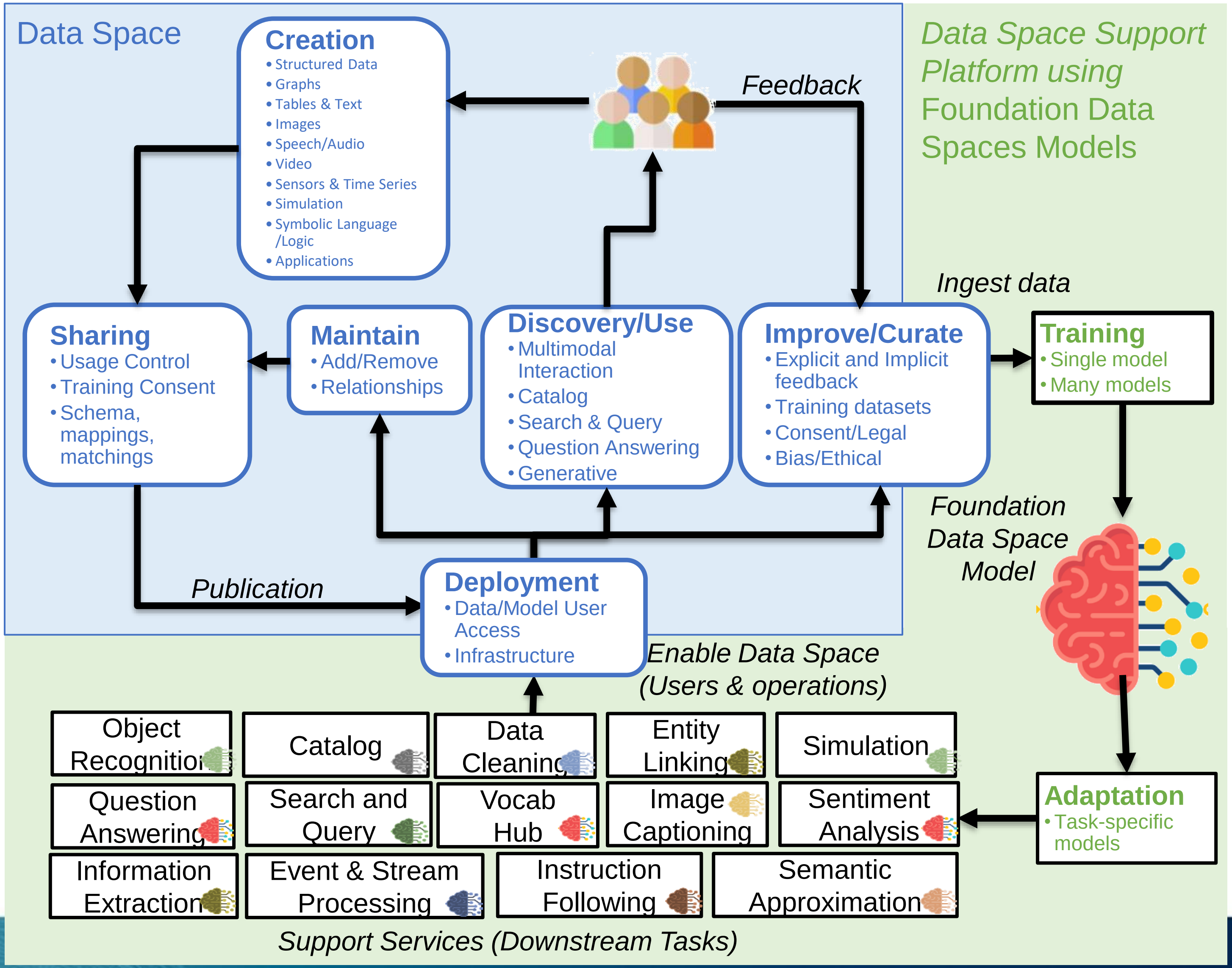
Foundational Model

Provides models to enable
(FM 4 Data Space)

Faster innovation cycles
Reduce cost with “good enough”
Supports Long Tail of Data

Basis of a sustainable business model?

A Unified Life Cycle…



Data Space

Creation
- Structured Data
- Graphs
- Tables & Text
- Images
- Speech/Audio
- Video
- Sensors & Time Series
- Simulation
- Symbolic Language /Logic
- Applications

Feedback

Sharing
- Usage Control
- Training Consent
- Schema, mappings, matchings

Maintain
- Add/Remove
- Relationships

Discovery/Use
- Multimodal Interaction
- Catalog
- Search & Query
- Question Answering
- Generative

Improve/Curate
- Explicit and Implicit feedback
- Training datasets
- Consent/Legal
- Bias/Ethical

Publication

Deployment
- Data/Model User Access
- Infrastructure

Enable Data Space (Users & operations)

Data Space Support Platform using Foundation Data Spaces Models

Ingest data

Training
- Single model
- Many models

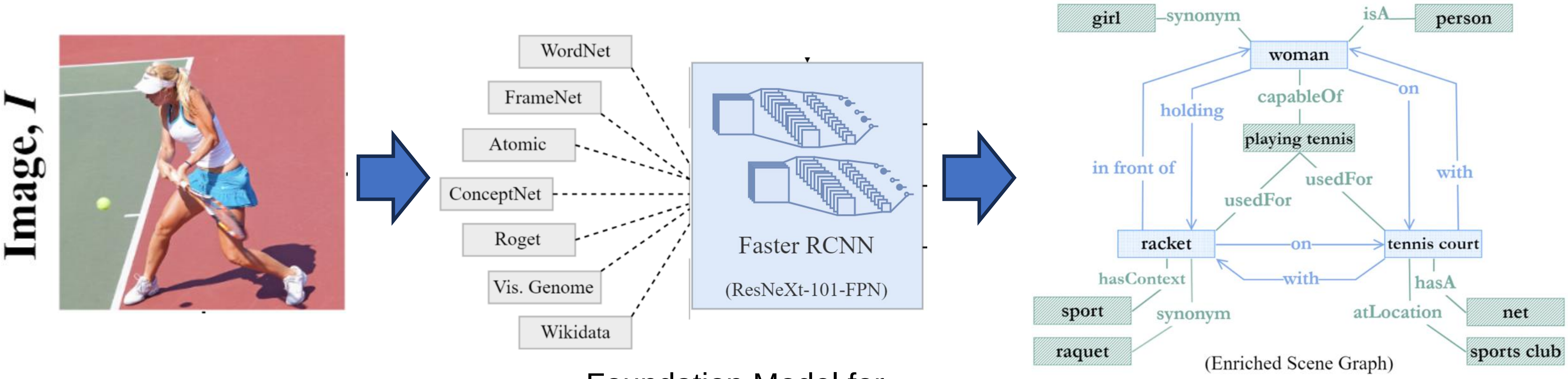
Foundation Data Space Model

Adaptation
- Task-specific models

Object Recognition	Catalog	Data Cleaning	Entity Linking	Simulation
Question Answering	Search and Query	Vocab Hub	Image Captioning	Sentiment Analysis
Information Extraction	Event & Stream Processing	Instruction Following	Semantic Approximation	

Support Services (Downstream Tasks)

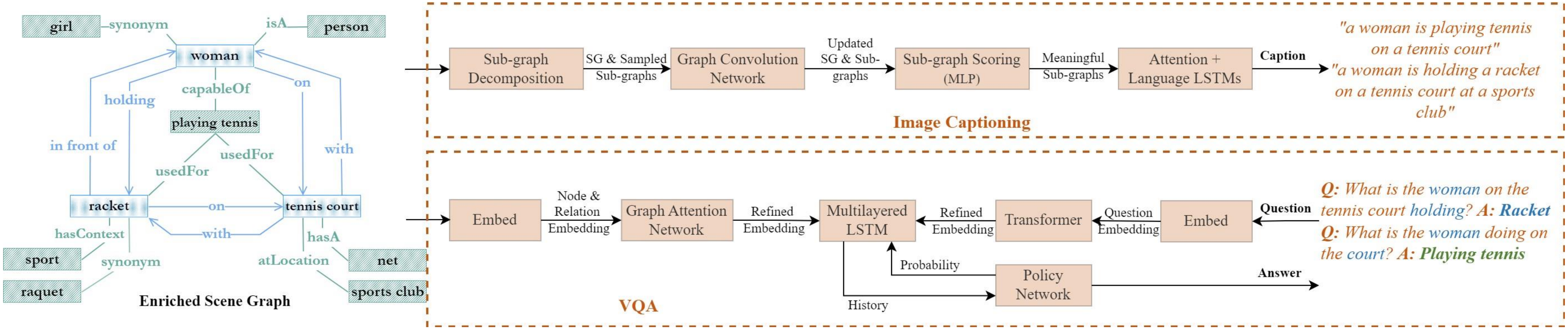
Foundation Model for Scene Graphs



Foundation Model for Scene Graphs

Graphs support the relationships between data space resources

Downstream Tasks – Image Captioning + Multimodal Question Answering



girl
synonym
isA
person
woman
holding
capableOf
on
playing tennis
in front of
usedFor
with
usedFor
racket
on
tennis court
with
hasContext
hasA
sport
synonym
atLocation
net
raquet
sports club
Enriched Scene Graph

Sub-graph Decomposition
SG & Sampled Sub-graphs
Graph Convolution Network
Updated SG & Sub-graphs
Sub-graph Scoring (MLP)
Meaningful Sub-graphs
Attention + Language LSTMs
Caption

"a woman is playing tennis on a tennis court"

"a woman is holding a racket on a tennis court at a sports club"

Image Captioning

Embed
Node & Relation Embedding
Graph Attention Network
Refined Embedding
Multilayered LSTM
Refined Embedding
Transformer
Question Embedding
Embed
Question
Probability
Policy Network
History
Answer

Q: *What is the woman on the tennis court holding?* ***A: Racket***

Q: *What is the woman doing on the court?* ***A: Playing tennis***

VQA

.J. Khan, J.G. Breslin, E. Curry, "Expressive Scene Graph Generation using Commonsense Knowledge Infusion for Visual Understanding and Reasoning," Extended Semantic Web Conference (ESWC) 2022, Crete, Greece, May-June 2022.

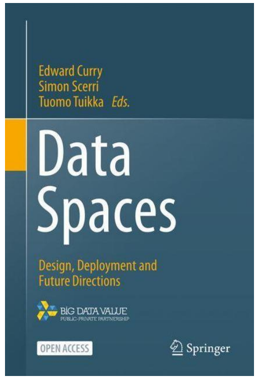
Conclusions

OPEN ACCESS

- Common European Data Spaces will require new data management paradigms in order to facilitate large-scale data pooling and sharing

- Semantics have never been more challenging
  - Scale, content space, multimodal, subjectivity, trust, cost, …

- Foundational / Language models and hybrid Neuro-Symbolic approaches provide promising avenues to develop new paradigms
  - Computationally driven
  - Knowledge Graphs + Deep Learning
  - Exact + Approximate best effort results



Edward Curry
Simon Scerri
Tuomo Tuikka *Eds.*

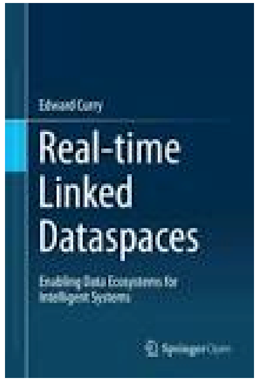
Data Spaces

Design, Deployment and Future Directions

BIG DATA VALUE

OPEN ACCESS

Springer

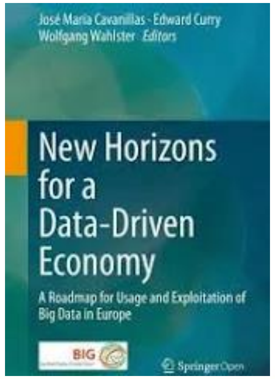


Edward Curry

Real-time Linked Dataspaces

Enabling Data Ecosystems for Intelligent Systems

SpringerOpen



José María Cavanillas · Edward Curry
Wolfgang Wahlster *Editors*

New Horizons for a Data-Driven Economy

A Roadmap for Usage and Exploitation of Big Data in Europe

BIG

SpringerOpen

SEMIC Specifications, services and trainings in support of Data Spaces

Alexandra Balahur

SEMIC Specifications, services and trainings in support of Data Spaces

Alexandra Balahur

DG DIGIT, Interoperability Unit

6 October 2023

Index

What is the index of this presentation?

01 **Why?** Understanding the policy context

02 **What?** Understanding the role and services from DG DIGIT

03 **Focus on Semantic Interoperability –** Understanding the role and services of SEMIC to support data spaces

04 **Where?** Reaching out to us through a single entry-point

Why

Understanding the policy context

Policy Context

- Europe’s digital future will be enabled by a **data-driven economy** and the use of **Artificial Intelligence**, fully respecting EU values and regulations. The public sector also needs to become more data-driven; improve the capability of developing policies and services through the management, sharing and use of data.

- The European Data Strategy aims to create a single market for data through common European data spaces that benefit from common standards and interoperability protocols.

- AI legislation and coordinated plan to foster the development and use of AI in Europe, highlighting the public sector as a trailblazer for using AI.

- The Interoperable Europe Act complements the EU data and digital policy landscape on data availability and data exchange, from a public sector angle. It implements interoperability by design and fosters the sharing and reuse of interoperable solutions.

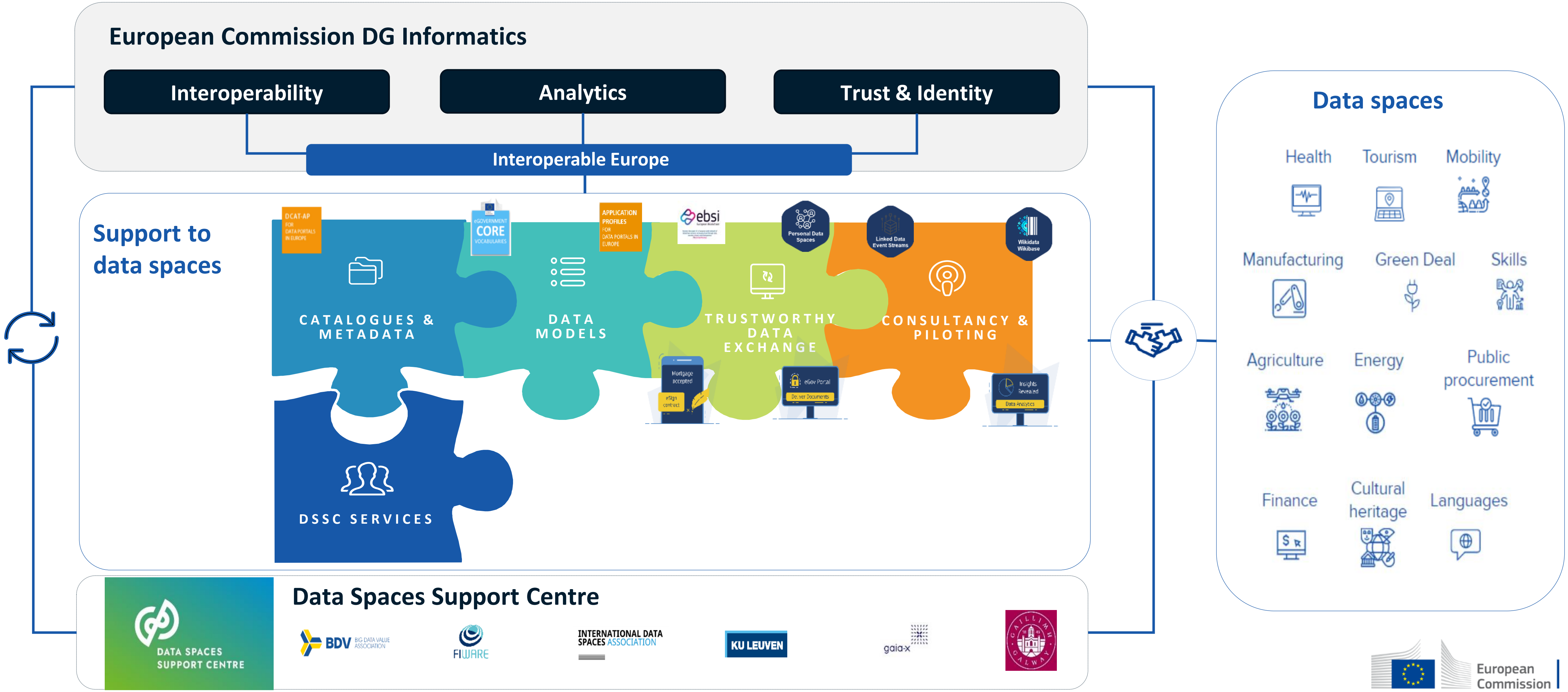
What

Understanding the role and services from DG DIGIT

European Commission

DG DIGIT’s role and services

DG Informatics is ready to support data spaces with existing assets and services, as well as to establish synergies with stakeholders active in this field to provide a more comprehensive support.



Focus on Semantic Interoperability

Understanding the role and services
from SEMIC supporting data spaces

European
Commission

SEMIC Service offering

SEMIC's goal is to deliver pragmatic support to help build an interoperable Europe.

Specifications

Publication and maintenance of open and free-to-reuse data models, with regular updates

Pilots

Developing specific solutions for public administrations to scale up their interoperability maturity

Toolkit

Provision of an accessible European Toolchain for data extraction, transformation and loading

Knowledge Hub

Training materials, guidelines and events to foster interoperability and share knowledge of its benefits

European Commission

SEMIC specifications

SEMIC specifications enable interoperability by:

- Making data transparent and available
- Supporting coherent implementation of laws and policies
- Helping implement cost efficiencies
- Helping digitalisation and harmonising processes

Core Vocabularies

The cornerstone for semantic interoperability

Core Vocabularies provide a standardised approach for describing key concepts such as locations, businesses, organisations and natural persons.

Application Profiles

A tailored data model for specific applications

Application Profiles make use of vocabularies for a detailed set of use cases to define mandatory relations, constraints and relationships.

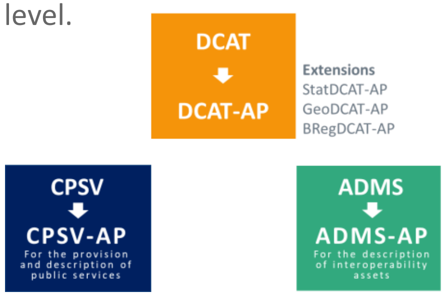
We adopt a balanced approach providing flexibility, customisation and solid indications to ensure a high degree of semantic interoperability.

Examples

SEMIC Core Vocabularies are reused in various implementations and extensions across domains and countries, from INSPIRE regulation to national databases (e.g. organisation registers) and private initiatives (Smart Data Models from Fiware).

DCAT-AP is a standard model to describe data catalogues. It is used by all open data portals in Europe. We have already made available extensions for statistics, geospatial and base registries.
We are currently supporting the European Health Data Space and the Mobility Data Space to extend DCAT-AP in their respective domains. This will ensure that minimum data descriptions will flow across data spaces.

CPSV-AP is a reusable and extensible data specification used for harmonising the way public services are described in a machine-readable format. Public administrations and service providers use this approach to describe their services and guarantee a level of cross-domain and cross-border interoperability at European, national and local level.


DCAT
DCAT-AP
Extensions
StatDCAT-AP
GeoDCAT-AP
BRegDCAT-AP
CPSV
CPSV-AP
For the provision and description of public services
ADMS
ADMS-AP
For the description of interoperability assets

Zooming into DCAT-AP

Objectives of DCAT-AP

Supporting the discovery of/access to (open) data in a cross-border and cross-domain environment, by describing metadata to be harvested across a distributed network of portals. DCAT-AP is based on the Data Catalog Vocabulary (DCAT) from W3C and

- expresses constraints and usages on DCAT properties and classes, and
- includes additional properties and usages of controlled vocabularies

Domains of applications

Open data portals with an extension for statistics and geospatial data.

Base registries metadata descriptions

Machine Learning with MLDCAT-AP

Data spaces
DCAT-AP can make a significant contribution in connecting data catalogues of data spaces. Stakeholders that we have discussed with have shared their interest in it for several European common data spaces. Examples include:

- NAPCORE-Mobility
- HealthDCAT-AP
- …

SEMIC DCAT-AP Support in 2023

7 November 2023

DCAT-AP working group webinar on status and governance

21 November 2023

DCAT-AP working group webinar on technical issues

Ongoing

Public review DCAT-AP 3.0.0

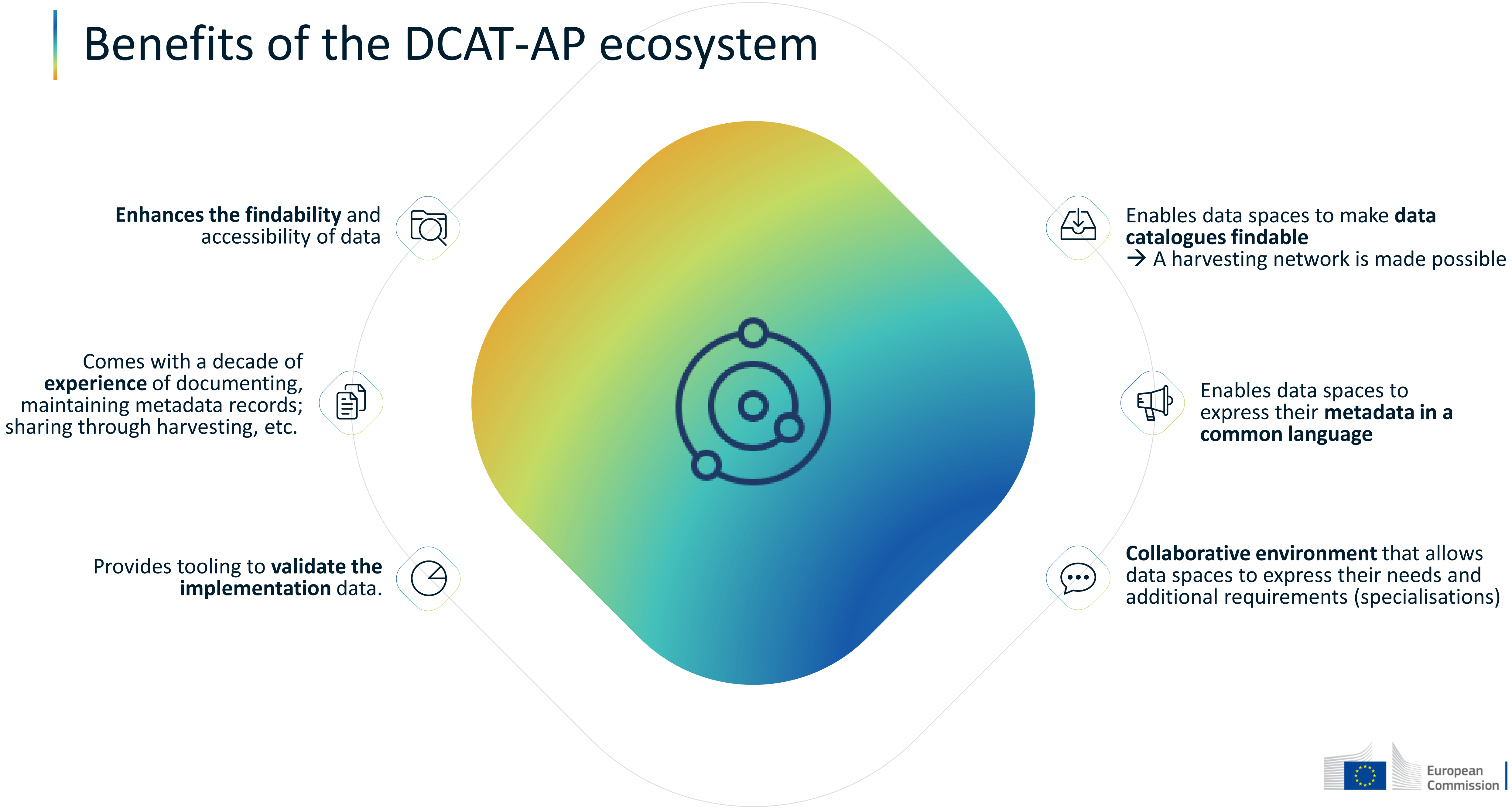
December 2023

Training materials on DCAT-AP, dedicated to Data Spaces

December 2023

Guidelines on profiling and extending DCAT-AP

Benefits of the DCAT-AP ecosystem



Enhances the findability and accessibility of data

Comes with a decade of **experience** of documenting, maintaining metadata records; sharing through harvesting, etc.

Provides tooling to **validate the implementation** data.

Enables data spaces to make **data catalogues findable**
→ A harvesting network is made possible

Enables data spaces to express their **metadata in a common language**

Collaborative environment that allows data spaces to express their needs and additional requirements (specialisations)

SEMIC pilots

SEMIC sets up pilots to showcase the value of new approaches and ecosystems, which can be leveraged across public administrations to scale up their interoperability maturity. Pilots usually involve participants from several Member States and sector-specific DGs co-creating solutions with SEMIC's support.

Artificial Intelligence

Elevating the application of AI tools

SEMIC wants to become the reference in AI for interoperability and interoperability in AI for public organisations in Europe.

Linked Data Event Streams

A new data publishing approach

A publishing strategy by which a data provider allows multiple third parties to stay in sync with the latest or historical versions of the data source in a cost-effective manner.

Personal Data Spaces

Empowering individuals and supporting the data economy

There is a need for more coordination and synergies between Personal Data Spaces implementations to ensure interoperability between existing and potential solutions in this emerging market.

Wikidata / Wikibase

Supporting community-driven efforts

Tools to enable the co-creation of semantic data models emerging from a community of users.

Examples

Linked Data Event Streams

The European Railways Agency has deployed a solution based on the LDES technology which allowed them to save costs and efficiently transfer only the data that has changed, together with what has changed and when, for historic records.

Proof-of-concept

Proof-of-concepts serve as practical examples of how AI can complement and automate existing ways of working, facilitating more efficient work and promoting interoperability within the European Union's stakeholders. These tools demonstrate the potential for AI to improve data processing, analysis, and interoperability. For example, the NLP-based proof-of-concept for automatically tagging web pages.

Studies and research

Studies and research conducted under SEMIC are research endeavours aimed at investigating various aspects of AI, interoperability, and advanced technologies. For example, SEMIC is currently creating a conference paper on fine-tuning of Large Language Models in Tourism domain.

Zooming into LDES

Use cases

Replicate the data: An LDES allows all data users to replicate the data from the unique source of truth (i.e. a base registry).

Stay in-sync: Stay up-to-date on the latest changes of the event source.

Grow your collection of objects: Each time an event happens, a new object will be added to the LDES allowing you to update your data on an event-base.

Benefits

No more maintenance hell
The data owner does not have to maintain all the querying APIs on their side.

Less chance for replication hell
There is no need for a local copy of the data source since it easy to connect and stay up-to-date.

Publishing data in a scalable manner

Since objects are published each time an event happens, the LDES publication technology is scalable by nature.

The role of SEMIC in LDES

A Linked Data Event Stream (LDES) is a collection of immutable objects whereby you do not change the data itself but simply add new data record to the stream. For business purposes, it is a publication strategy to share your data. It allows data users to:

- Have up to date data
- Be aware of changes
- Access historic data
- Relate historic data to current data

SEMIC assists Members States and organisations to publish their data in a manner that ensures interoperability, using LDES:

- Aid in the implementation of pilots
- Aid in the development of Strategy and roadmap
- Share knowledge

SEMIC Knowledge Hub

SEMIC acts as an enabler to exchange experiences, good practices and insights. By sharing knowledge we aim to facilitate the development and use of data specifications, as well as to discuss the latest technological trends, and present expert views on semantic interoperability topics.

Community building

Experience exchange and building consensus

Organisation of webinars, workshops and other events to understand user and market needs, foster experience and good practice exchange, and help reach consensus between stakeholders.

Training

Providing online learning materials

SEMIC contributes to advancing digital skills in the area of interoperability to support policy, service delivery and impact evaluation.

Studies & other materials

Setting trends and providing guidance

Over the years SEMIC has published many reports, studies and other guidance materials revolving around semantic interoperability.

Examples

SEMIC has been developing many online training materials that contribute to advancing digital skills. Some of the training developed are particularly relevant to data spaces.

The training developed by SEMIC can be grouped into three main categories:

- Core semantics
- DCAT and DCAT-AP
- Other semantic solutions for public administration

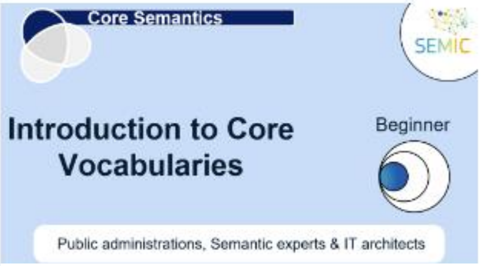
The training materials developed by SEMIC are all publicly available on EU Academy.

Online trainings with relevance to data spaces

The training materials developed by SEMIC are all publicly available on [EU Academy](#).

CORE SEMANTICS

[Introduction to Core Vocs](#)
[Introduction to CPSV-AP](#)
[ABR eLearning course](#)

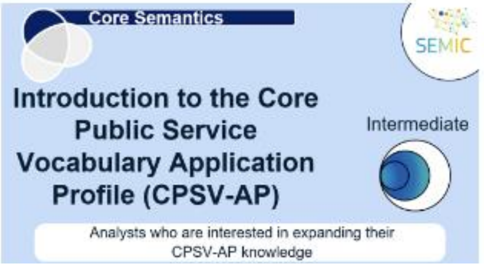

Core Semantics

SEMIC

Introduction to Core Vocabularies

Beginner

Public administrations, Semantic experts & IT architects

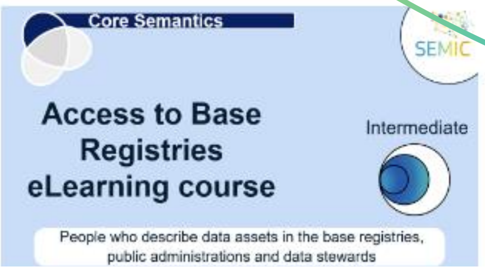

Core Semantics

SEMIC

Introduction to the Core Public Service Vocabulary Application Profile (CPSV-AP)

Intermediate

Analysts who are interested in expanding their CPSV-AP knowledge


Core Semantics

SEMIC

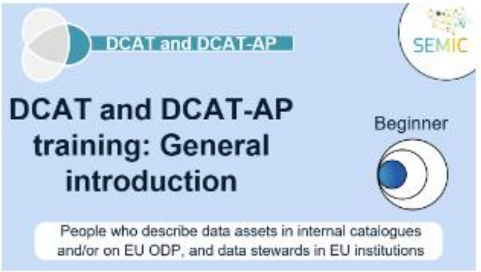
Access to Base Registries eLearning course

Intermediate

People who describe data assets in the base registries, public administrations and data stewards

DCAT AND DCAT-AP

[Introduction](#)
[Basic user](#)
[Advanced user](#)

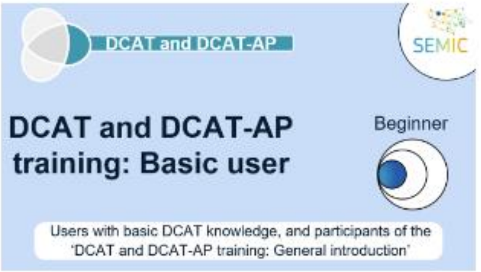

DCAT and DCAT-AP

SEMIC

DCAT and DCAT-AP training: General introduction

Beginner

People who describe data assets in internal catalogues and/or on EU ODP, and data stewards in EU institutions

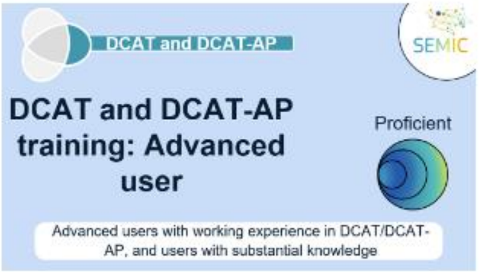

DCAT and DCAT-AP

SEMIC

DCAT and DCAT-AP training: Basic user

Beginner

Users with basic DCAT knowledge, and participants of the 'DCAT and DCAT-AP training: General introduction'


DCAT and DCAT-AP

SEMIC

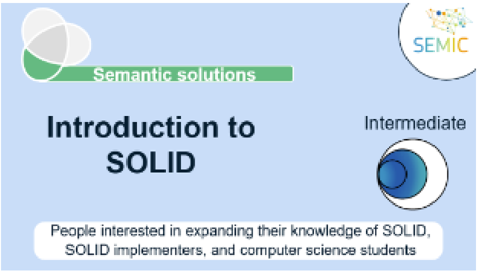
DCAT and DCAT-AP training: Advanced user

Proficient

Advanced users with working experience in DCAT/DCAT-AP, and users with substantial knowledge

OTHER SEMANTIC SOLUTIONS FOR PUBLIC ADMINISTRATIONS

[SOLID](#)
[Wikibase](#)
[LDES](#)

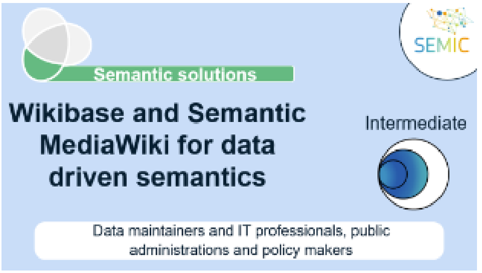

Semantic solutions

SEMIC

Introduction to SOLID

Intermediate

People interested in expanding their knowledge of SOLID, SOLID implementers, and computer science students

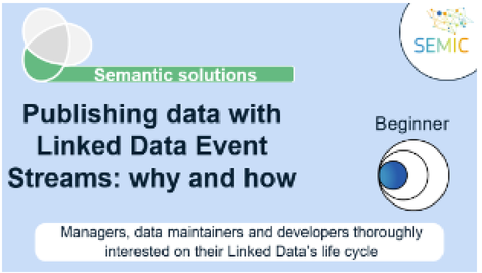

Semantic solutions

SEMIC

Wikibase and Semantic MediaWiki for data driven semantics

Intermediate

Data maintainers and IT professionals, public administrations and policy makers


Semantic solutions

SEMIC

Publishing data with Linked Data Event Streams: why and how

Beginner

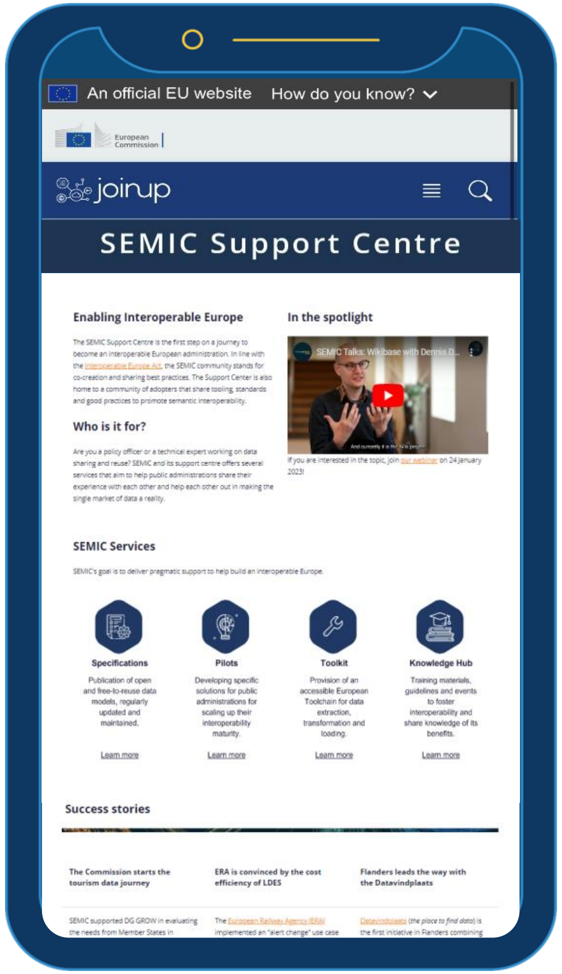
Managers, data maintainers and developers thoroughly interested on their Linked Data's life cycle

Where

Reaching out to us through a single entry point

SEMIC Support Centre

Your one-stop-shop for all digital interoperability challenges


An official EU website
How do you know?
European Commission
joinup
SEMIC Talks: Wikibase with Dennis D...
And currently it is the Wiki project

SEMIC Support Centre

Enabling Interoperable Europe

The SEMIC Support Centre is the first step on a journey to become an interoperable European administration. In line with the Interoperable Europe Act, the SEMIC community stands for co-creation and sharing best practices. The Support Center is also home to a community of adopters that share tooling, standards and good practices to promote semantic interoperability.

Who is it for?

Are you a policy officer or a technical expert working on data sharing and reuse? SEMIC and its support centre offers several services that aim to help public administrations share their experience with each other and help each other out in making the single market of data a reality.

In the spotlight

If you are interested in the topic, join our webinar on 24 January 2023!

SEMIC Services

SEMIC's goal is to deliver pragmatic support to help build an interoperable Europe.

Specifications
Publication of open and free-to-reuse data models, regularly updated and maintained.
Learn more

Pilots
Developing specific solutions for public administrations for scaling up their interoperability maturity.
Learn more

Toolkit
Provision of an accessible European Toolchain for data extraction, transformation and loading.
Learn more

Knowledge Hub
Training materials, guidelines and events to foster interoperability and share knowledge of its benefits.
Learn more

Success stories

The Commission starts the tourism data journey
SEMIC supported DG GROW in evaluating the needs from Member States in

ERA is convinced by the cost efficiency of LDES
The European Railway Agency (ERA) implemented an "alert change" use case

Flanders leads the way with the Datavindplaats
Datavindplaats (*the place to find data*) is the first initiative in Flanders combining

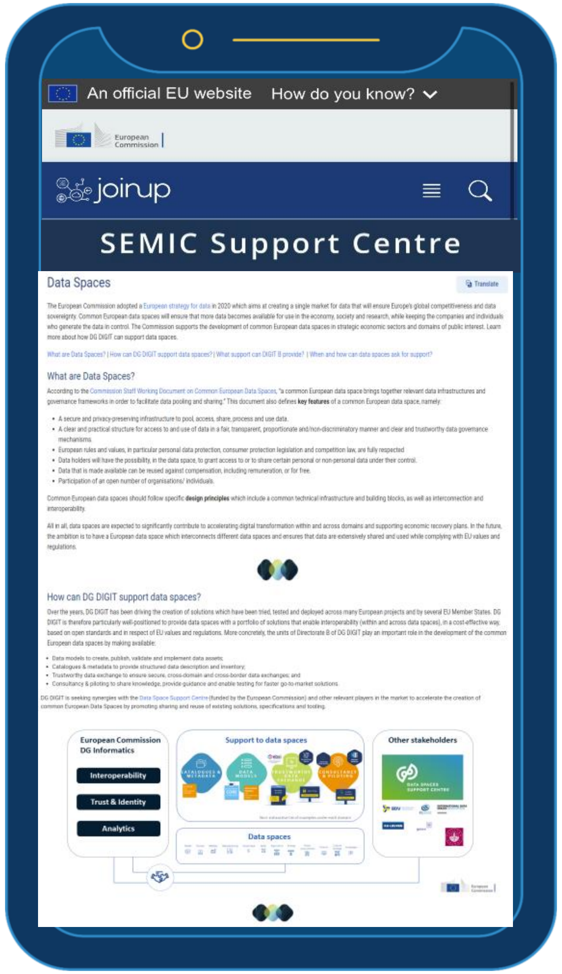
What is it?

- Concrete and direct support to all stakeholders through our helpdesk and GitHub
- Collects good practices and serves as a knowledge hub on interoperability issues

Who is it for?

- Public officers and technical experts
- Any public organisation in its interoperable journey

SEMIC Support Centre – Data Spaces



An official EU website How do you know?

European Commission

joinup

SEMIC Support Centre

Data Spaces

Translate

The European Commission adopted a European strategy for data in 2020 which aims at creating a single market for data that will ensure Europe's global competitiveness and data sovereignty. Common European data spaces will ensure that more data becomes available for use in the economy, society and research, while keeping the companies and individuals who generate the data in control. The Commission supports the development of common European data spaces in strategic economic sectors and domains of public interest. Learn more about how DG DIGIT can support data spaces.

What are Data Spaces? | How can DG DIGIT support data spaces? | What support can DIGIT B provide? | When and how can data spaces ask for support?

What are Data Spaces?

According to the Commission Staff Working Document on Common European Data Spaces, "a common European data space brings together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing." This document also defines **key features** of a common European data space, namely:

- A secure and privacy-preserving infrastructure to pool, access, share, process and use data.
- A clear and practical structure for access to and use of data in a fair, transparent, proportionate and/non-discriminatory manner and clear and trustworthy data governance mechanisms.
- European rules and values, in particular personal data protection, consumer protection legislation and competition law, are fully respected
- Data holders will have the possibility, in the data space, to grant access to or to share certain personal or non-personal data under their control.
- Data that is made available can be reused against compensation, including remuneration, or for free.
- Participation of an open number of organisations/ individuals.

Common European data spaces should follow specific **design principles** which include a common technical infrastructure and building blocks, as well as interconnection and interoperability.

All in all, data spaces are expected to significantly contribute to accelerating digital transformation within and across domains and supporting economic recovery plans. In the future, the ambition is to have a European data space which interconnects different data spaces and ensures that data are extensively shared and used while complying with EU values and regulations.

How can DG DIGIT support data spaces?

Over the years, DG DIGIT has been driving the creation of solutions which have been tried, tested and deployed across many European projects and by several EU Member States. DG DIGIT is therefore particularly well-positioned to provide data spaces with a portfolio of solutions that enable interoperability (within and across data spaces), in a cost-effective way, based on open standards and in respect of EU values and regulations. More concretely, the units of Directorate B of DG DIGIT play an important role in the development of the common European data spaces by making available:

- Data models to create, publish, validate and implement data assets;
- Catalogues & metadata to provide structured data description and inventory;
- Trustworthy data exchange to ensure secure, cross-domain and cross-border data exchanges; and
- Consultancy & piloting to share knowledge, provide guidance and enable testing for faster go-to-market solutions.

DG DIGIT is seeking synergies with the Data Space Support Centre (funded by the European Commission) and other relevant players in the market to accelerate the creation of common European Data Spaces by promoting sharing and reuse of existing solutions, specifications and tooling.

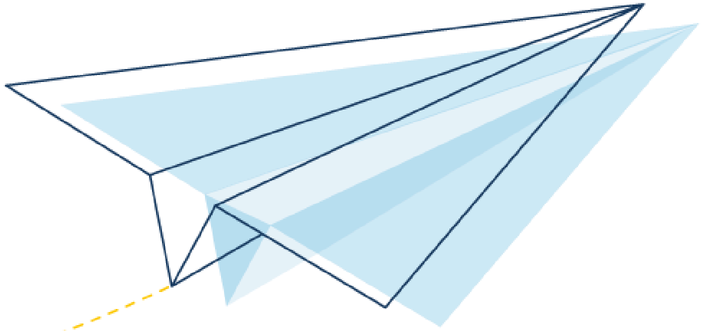
What can be found on the page?

- Explanation about what data spaces are
- Information about how DG DIGIT supports data spaces
- Specific support provided by DIGIT B
- Contact details to ask for support

Contact us

Find more information and contact us through a single entry-point:

https://joinup.ec.europa.eu/collection/semic-support-centre/data-spaces



Thank you !

Questions & Answers

data.
europa
academy

Please provide your feedback!

data.europa academy

Stay up-to-date on our
2023 activities!

data.europa academy

Join our next webinars!

Workshop

How to use open data for your research?

data. europa academy

19 October 2023

10.00 – 12.00 CET

WEBINAR

Open Data Maturity 2022: Diving deeper into the impact dimension

data. europa academy

27 October 2023

10.00 — 11.30 CET

data. europa academy

Fill in our user survey

data.europa academy

Sign up for the newsletter: data.europa.eu/newsletter

Follow us on social media:

EU_opendata

Publications Office of the European Union

data.europa.eu

data.europa academy

Thank you

data.europa academy