



EUROPEAN LANGUAGE DATA SPACE



The European Language Data Space Architecture and Components

Stelios Piperidis (Athena RC, Greece)

spip@athenarc.gr

19-03-2025 LDS Launch Conference
<https://language-data-space.ec.europa.eu>

Long History of Language Data Sharing

META-SHARE LEARN • DISCOVER • PARTICIPATE • CONNECT • LOGIN

Search & exchange language resources

META-SHARE is an open and secure network of repositories for sharing and exchanging language data, tools and related web services

Share your own resources!

[JOIN OUR NETWORK NOW](#)

Already a member? [Log in](#)

Search the META-SHARE inventory

OR [LEARN MORE](#)

- 4,481 users
- 2,887 language resources
- 32% text corpora
- 27,630 number of downloads

Virtual Language Observatory Search Contributors Help CLARIN

CLARIN Virtual Language Observatory

Welcome to the VLO!

Use the **search bar** below to start searching through hundreds of thousands of language resources, or **continue** to browse everything and use **facets** to narrow down to your area of interest or discover new resources.

[See all records](#) [Take a quick tour](#)

Search through 1,030,321 records

Home Browse Resources Help About Register Login

ELRC-SHARE Repository

Type in your keywords, please...

Welcome to the ELRC-SHARE repository!

The ELRC-SHARE repository is used for documenting, storing, browsing and accessing [Language Resources Coordination](#) and considered useful for feeding the CEF Automated Translation (CEF AT) platform.

If you want to contribute resources, all you have to do is [register](#) (new user) or [login](#) (existing user) and go on...

European Language Resource Coordination

EUROPEAN LANGUAGE GRID

Language Technologies

Discover, try out, use and download LT services and resources for all European languages.

Browse ELG and find the LT services, resources, developers and providers you are looking for.

Search the catalogue

Order by: Resource Name A-Z

...	...
...	...
...	...

8000 Corpora 3884 Tools & Services 2812 Conceptual Resources 510 Models & Grammars 1775 Organizations 513 Projects



Build on existing achievements



Respond to language data requirements for multilingual Europe in the AI era



Facilitate the language data economy

Contribute to the upcoming Data Spaces ecosystem

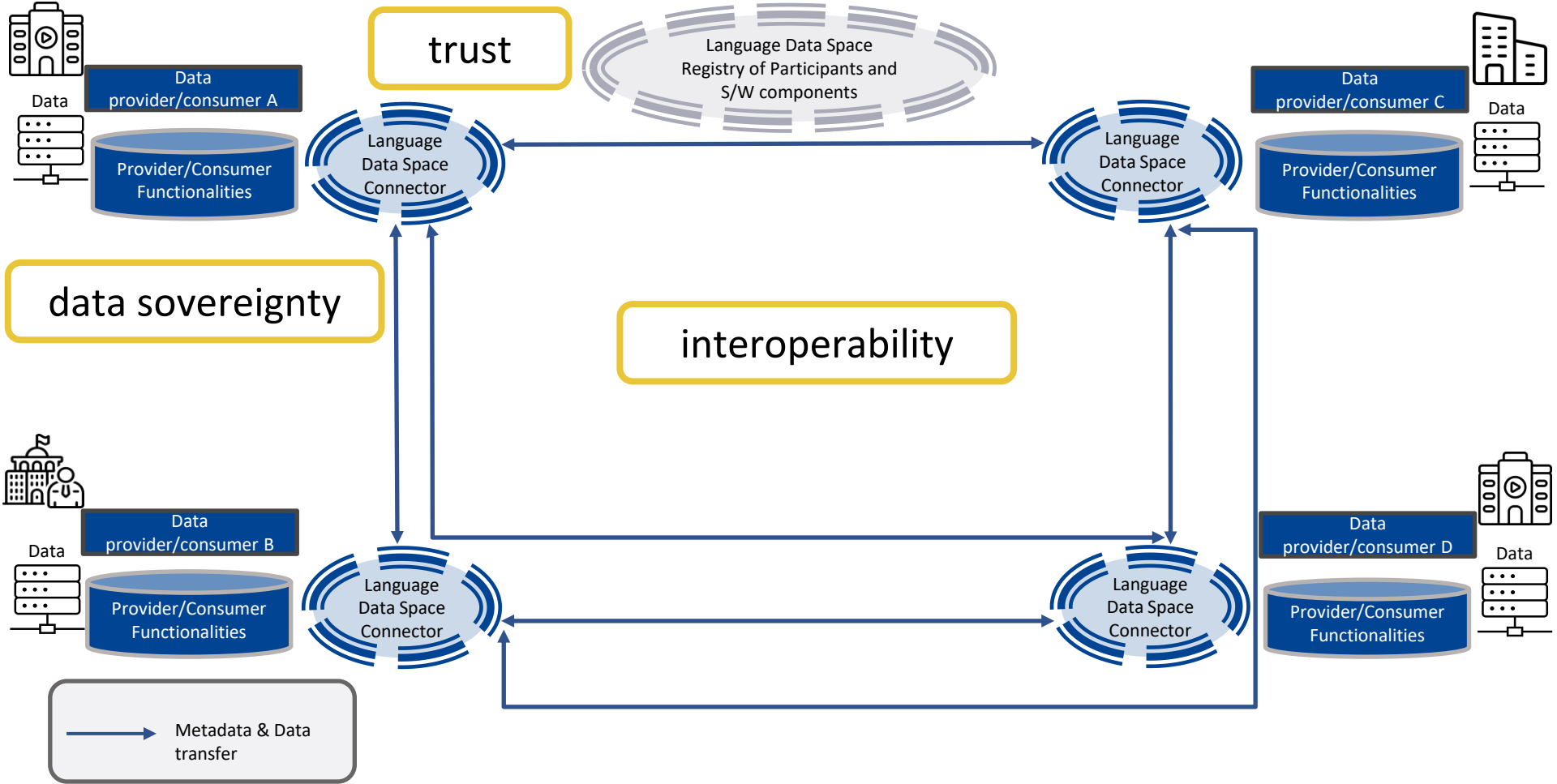
data sovereignty

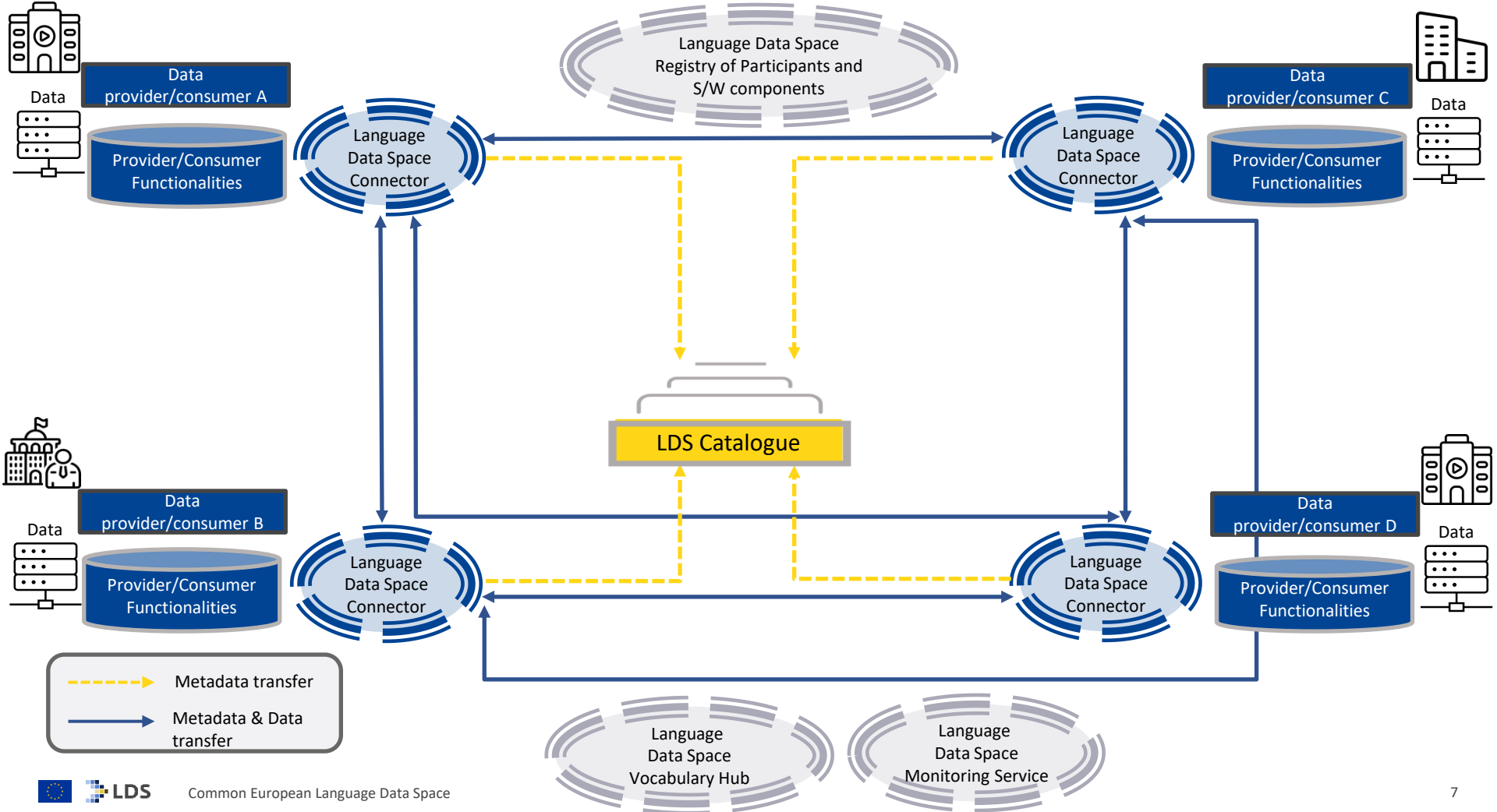
trust

interoperability

LDS Architecture

- Objective 1: **provide technological support for complying with the Data Spaces principles**
 - sovereignty, trust, interoperability
- Objective 2: **provide technological support for LDS workflows** designed
 - for data providers & consumers
 - for the LDS Governance Board





LDS Connector

- Asset: Data/Files+Metadata
- Offer: Asset+Policy



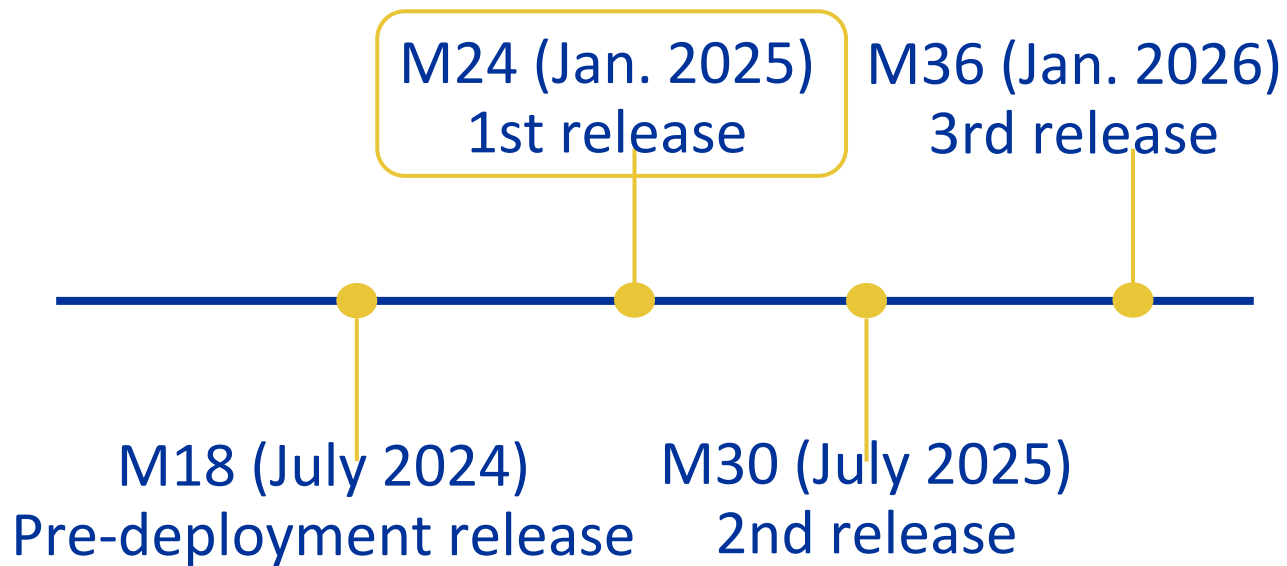
Connector features for

- data providers
- data consumers
- data prosumers

Standards & Technologies

- Eclipse Dataspace Components (EDC)
- Language DCAT-AP
- Open Digital Rights Language (ODRL)
- Dataspace Protocol (DSP)

LDS Release Plan





**EUROPEAN
LANGUAGE
DATA SPACE**

v1.0.0-beta

LDS v1.0.0-beta

1. LDS participant application and onboarding

- Organisation's **legal representative applies** online (<https://ldssetup.ilsp.gr>) and accepts the LDS Terms of Service
- The data are added in the **LDS Registry of participants**
- **LDS Governance Board** assesses the application (to ensure that it complies with the eligibility criteria) → approves or rejects (providing the reasons)
- **Upon approval**, participant downloads, installs & configures the **LDS connector at their own infrastructure** → Issuance of **authentication keys**
- **Registers the Connector** at the LDS Registry
- Configures their **storage solutions**

[Click here to see the demo video!](#)

2. User Identity Access Management (inside the participant's LDS connector)



- **Keycloak solution**
- **Registration** of users by user administrator (with possibility to use active directory)
- Assignment of pre-defined LDS **user roles** for providers & consumers with simple and advanced rights (with possibility to assign multiple roles)

The screenshot displays the LDS user interface. On the left, a navigation menu includes: LDS, Manage, Clients, Client scopes, Realm roles, Users, Groups, Sessions, and Events. The main content area shows a list of user roles:

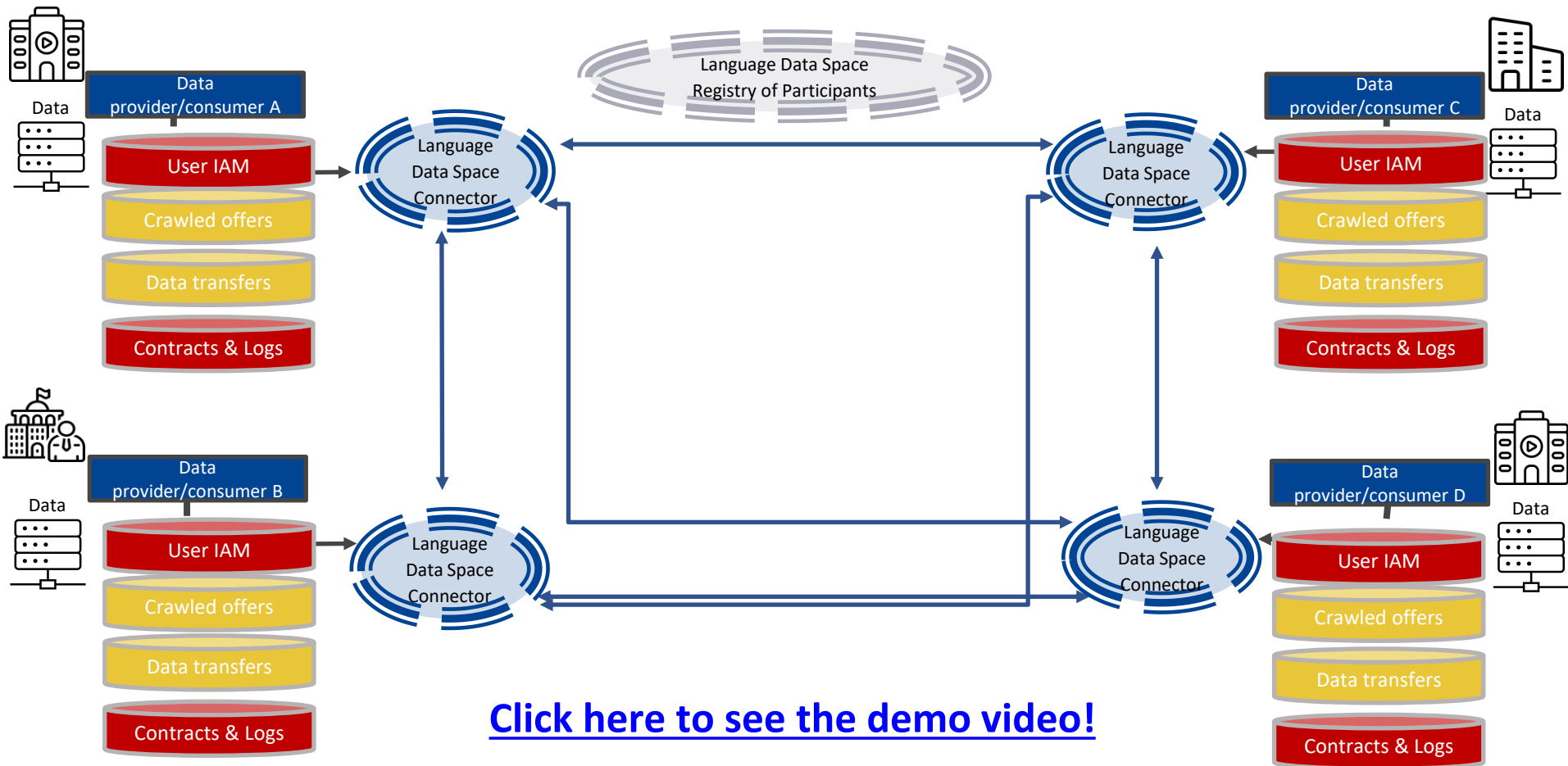
- Data user**: Simple data consumer. Can browse and search for data offerings, and request those offered at no charge. Can access all data acquired by the organisation.
- Data acquisitions manager**: Advanced data consumer. Can search for and request all offers, performing financial transactions to acquire them. Has access to all data acquired by the organisation.
- Business analytics**: Data provider and/or consumer. Generates and manages queries on the organisation's business operations. Has access to the regulations of the organisation.

A central circular graphic labeled **User roles** is connected to these three roles. On the right, a login form is visible with the **EUROPEAN LANGUAGE DATA SPACE** logo and the text "Sign in to your account". The form includes fields for **Username or email** and **Password**, and a **Sign in** button.

3. Data Providers - Creating assets, policies, and offers



4. Data Consumers – offers discovery, negotiation, contracts, download

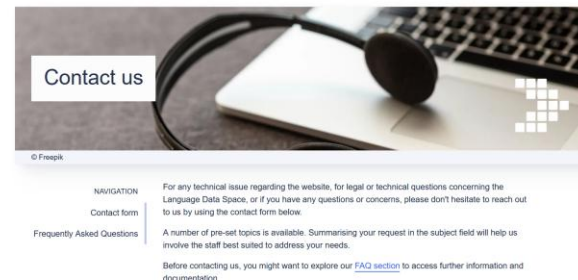


[Click here to see the demo video!](#)

Technical support

- LDS contact https://language-data-space.ec.europa.eu/help_en
- FAQs https://language-data-space.ec.europa.eu/help/faq_en
- [LDS technical documentation](#)
 - for technical administrators
 - for LDS users (providers & consumers)
 - for LDS Governance Board members

Help



The screenshot shows a 'Help' page with a header image of a laptop and headphones. A 'Contact us' button is visible. Below the image, there is a 'NAVIGATION' section with links for 'Contact form' and 'Frequently Asked Questions'. A paragraph of text explains that for technical issues, users should use the contact form, and that pre-set topics are available in the FAQ section.

FAQ



The screenshot shows the 'LDS documentation User Guide' page. The main heading is 'The European Language Data Space Documentation' with the subtitle 'European Language Data Space: the European Marketplace for Language Data'. A 'User Guide' button is present. Below the heading are three icons: 'Share' (a cloud with a plus sign), 'Connect' (two people icons), and 'Benefit' (a leaf icon). Each icon has a corresponding text block describing the benefit.

Share
You will be able to share and monetise your language data, language models and other language resources through a single platform, taking European values and compliance with EU regulations fully into account.

Connect
You will be able to connect and exchange with other stakeholders through the European Language Data Space.

Benefit
You will be supported by the European Language Data Space in the development of multilingual and multimodal language technologies and language-centric AI.

Data Spaces Principles

Trust framework

- applicants assessed by LDS Governance Board
- applicants must consent to LDS Terms of Service
- transactions only between authenticated & authorised connectors
- IAM for individuals inside organisation
- data transfer only upon conclusion of contract (between known contracting parties)

Data sovereignty

- data remain at one's own storage space
- participants define their data policies
- policy enforcement before approval of data request

Provenance & traceability

- contracts stored at both contracting parties' connectors
- logging & monitoring of transactions locally & centrally

Interoperability

- compliant with Dataspace Protocol
- adoption of standard transfer protocols for data transfer
- deployment of standard model for self-descriptions (based on DCAT model)
- deployment of ODRL model for policies
- recommendations for standard models & formats for data

Demo/Testing Setup

EUROPEAN LANGUAGE DATA SPACE

MANAGEMENT

- Home
- History
- Storage Solutions

OPERATIONS

- Assets
- Policies
- Offers

ACCOUNT

- Logout

LDS Connector Management Panel

Here you can create and manage your assets, your policies and your offers and review your contract agreements.

50 Assets	18 Policies	25 Offers
-----------	-------------	-----------

Buttons: Create Asset, View Assets, Create Policy, View Policies, Create Offers, View Offers

Search

Order by: Newest first

1	2024-01-01	Public Domain
2	2024-01-01	Public Domain
3	2024-01-01	Public Domain
4	2024-01-01	Public Domain
5	2024-01-01	Public Domain
6	2024-01-01	Public Domain
7	2024-01-01	Public Domain
8	2024-01-01	Public Domain
9	2024-01-01	Public Domain
10	2024-01-01	Public Domain
11	2024-01-01	Public Domain
12	2024-01-01	Public Domain
13	2024-01-01	Public Domain
14	2024-01-01	Public Domain
15	2024-01-01	Public Domain
16	2024-01-01	Public Domain
17	2024-01-01	Public Domain
18	2024-01-01	Public Domain
19	2024-01-01	Public Domain
20	2024-01-01	Public Domain
21	2024-01-01	Public Domain
22	2024-01-01	Public Domain
23	2024-01-01	Public Domain
24	2024-01-01	Public Domain
25	2024-01-01	Public Domain

EUROPEAN LANGUAGE DATA SPACE

MANAGEMENT

- Home
- History
- Storage Solutions

OPERATIONS

- Assets
- Policies
- Offers

ACCOUNT

- Logout

LDS Connector Management Panel

Here you can create and manage your assets, your policies and your offers and review your contract agreements.

60 Assets	21 Policies	29 Offers
-----------	-------------	-----------

Buttons: Create Asset, View Assets, Create Policy, View Policies, Create Offers, View Offers

Language Technologies

Discover the full, free and download LT services and resources for all European languages.

Review LTG and find the LT services, resources, developers and providers you are looking for.

EUROPEAN LANGUAGE GRID

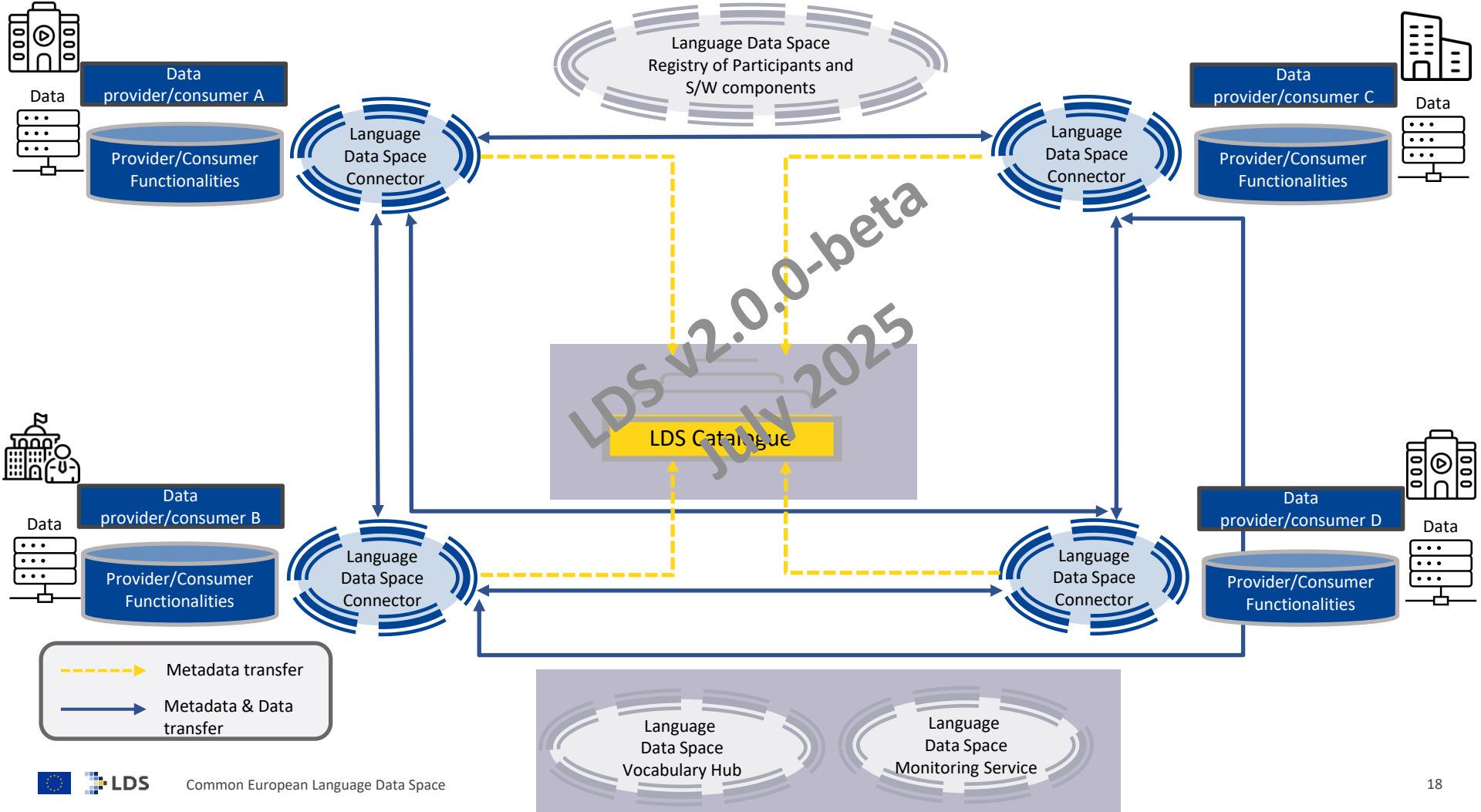
8000	3884	2812	510	1775	513
Assets	Assets	Assets	Assets	Assets	Assets

Apply at

<https://ldssetup.ilsp.gr/>

to join LDS and test it

Vanilla





Common European Language Data Space

Thank you!



A Common European Language Data Space – funded under contract LC-01936389 with the European Union.

Stelios Piperidis (Athena RC, Greece)
spip@athenarc.gr

19-03-2025 LDS Launch Conference
<https://language-data-space.ec.europa.eu>