www.egi.eu
@EGI_eInfra

A Common Approach to Data Ecosystems and Data Spaces: Looking at GAIA-X and EOSC

Mark Dietrich
Senior Advisor
mark.dietrich@egi.eu

The work of the EGI Foundation
is partly funded by the European Commission
under H2020 Framework Programme

Overview

Swirling Terms and Technologies

Definitions

Proposing a Consistent Framework: Data Exchange Approach

- Key Characteristics
- A Spectrum of Data Exchange Approaches

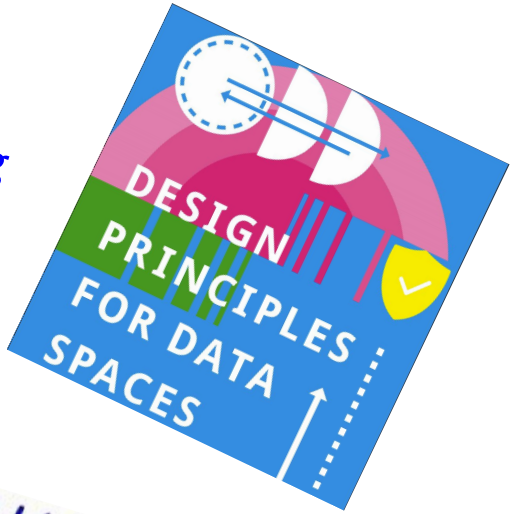
Exploring the Spectrum

Swirling Terms and Technologies

Can we make sense of them all?

data sharing space

Data Sharing Domain



INTERNATIONAL DATA SPACES ASSOCIATION

Open Data Ecosystem

gaia-x

DAWEX

Defining Data Ecosystem, Data Space...

...and their relationship

Data Ecosystem: *a purposeful collaboration or partnership consuming, producing and providing interoperable data and related resources*.

- Multiple data ecosystems possible -- they already exist:
  - E.g. in agriculture: Djust Connect, API-Agro, DKE Agrirouter, JoinData, Agrimetrics, Aladin.farm, DataConnect
  - "Science communities" in different areas of research

Data Space: *a collection of FAIR, quality data and related resources consumed, produced and provided by identified participants, each respecting societal values and operating within an explicit framework of trust and governance*.

- EC refers to **Public** Data Space
  - Does not identify any collaborating actors or a community
  - Is not "purposeful"
- EC aspires to a "single common European data space"
- Domain Data Spaces, collecting data and data-services relevant to a domain or sector.

Data Ecosystem = Private Data Space + Community + Purpose

Data Exchange Approach (DEA)

A Unifying Concept

Data Exchange Approach: defined as

- a set of (1) organizational policies and roles and
- (2) technical specifications,
- required to enable the trustworthy exchange of data between two parties.

Data Ecosystems and Data Spaces are types of DEA.

DEAs can flexibly describe different characteristics of exchange structures.

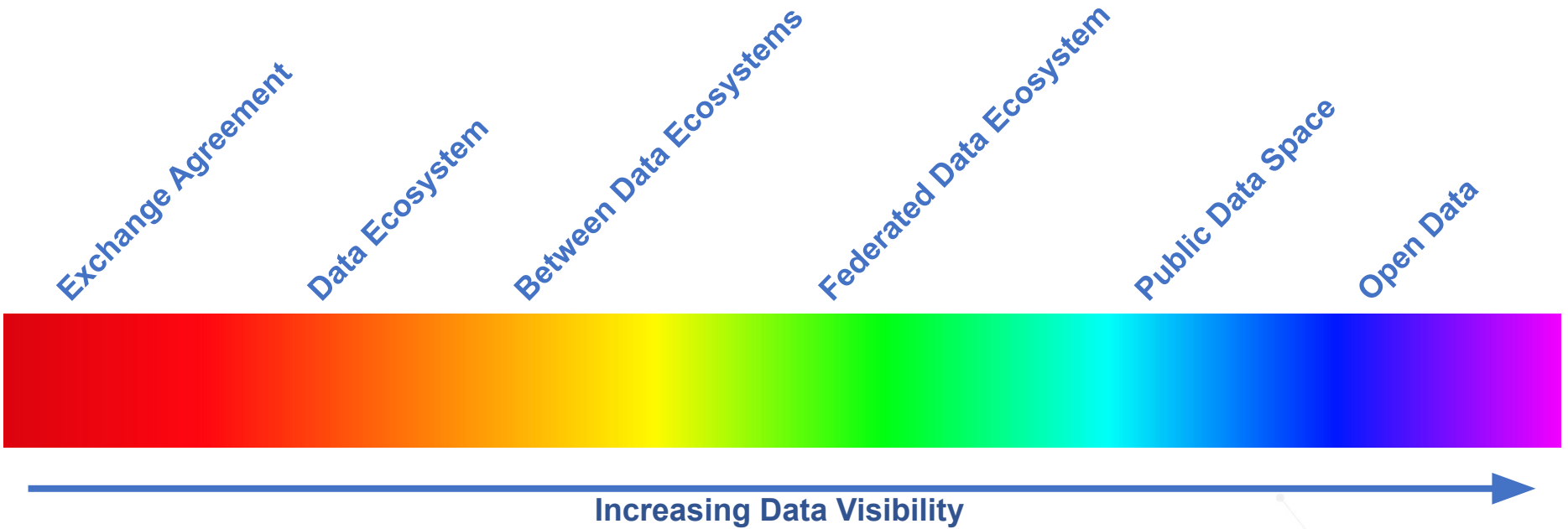
- E.g. the **visibility** of data, which can be limited or controlled:
  - exposure to only one other party
  - combined exposure of metadata and data to limited groups
  - public exposure of both data and metadata

Key Characteristics of a DEA

Dimensions for Comparison

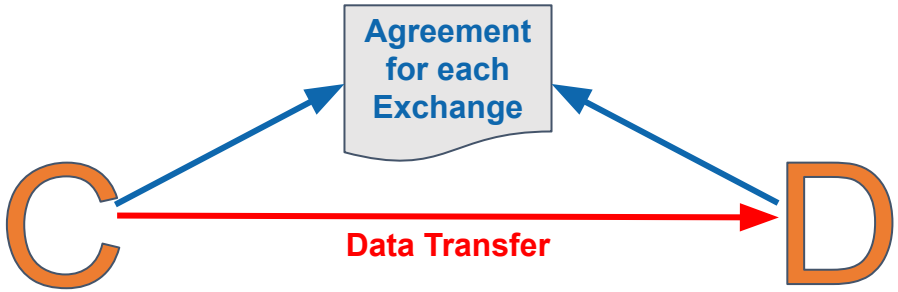
- Visibility and Security
- Findability
- Accessibility
- Interoperability
- Reusability
- Quality
- Data Sovereignty Mechanism(s)
- Community
- Purpose
- Governance

Spectrum of DEAs



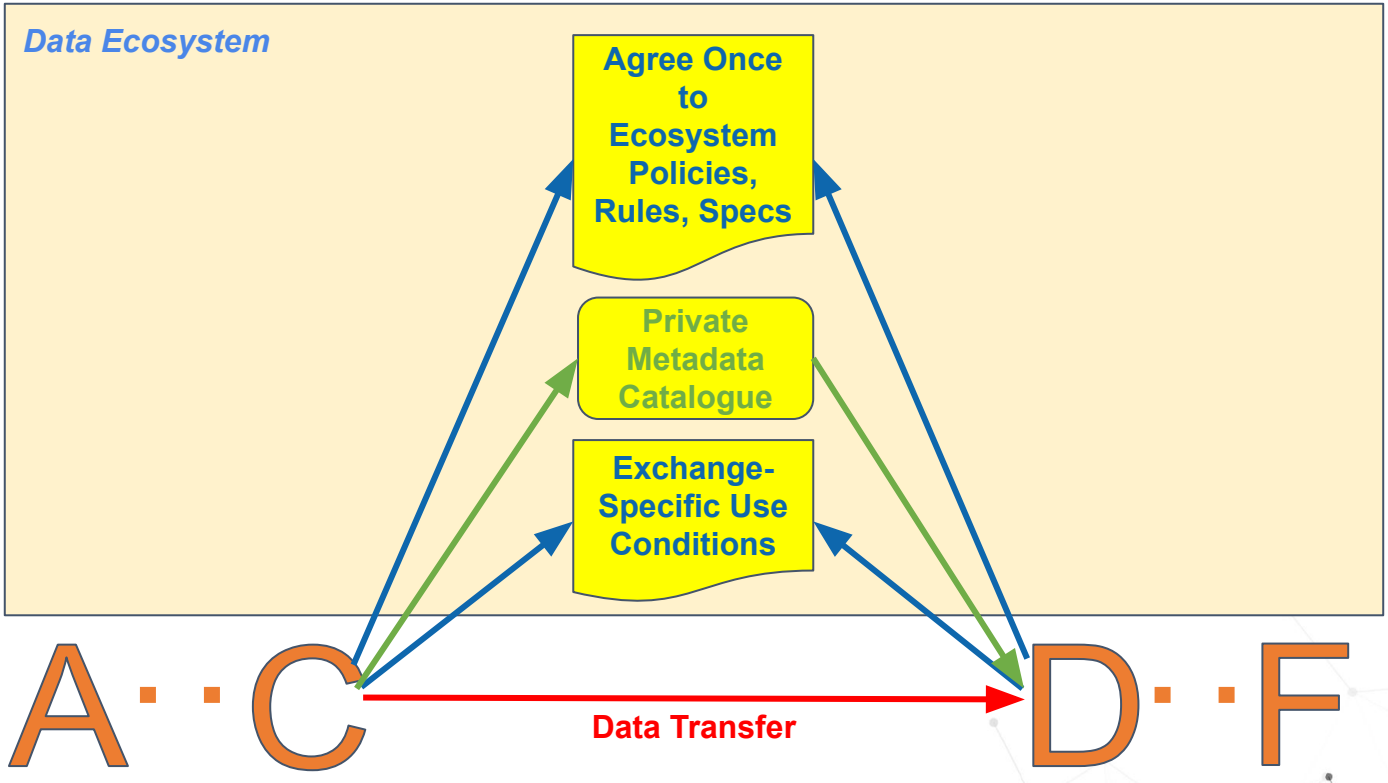
With an Exchange Agreement

Data Exchange between Parties C and D



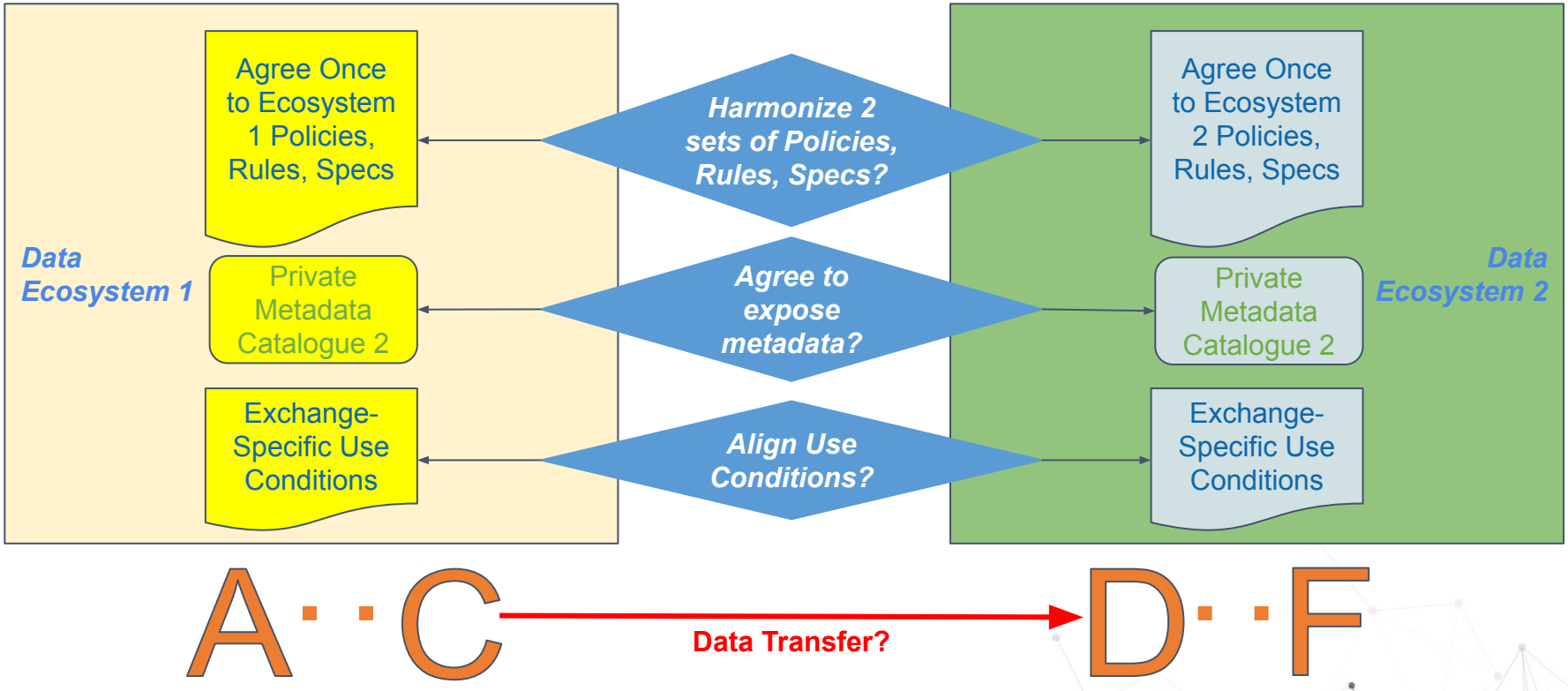
In a Data Ecosystem

Between Community Members A to F



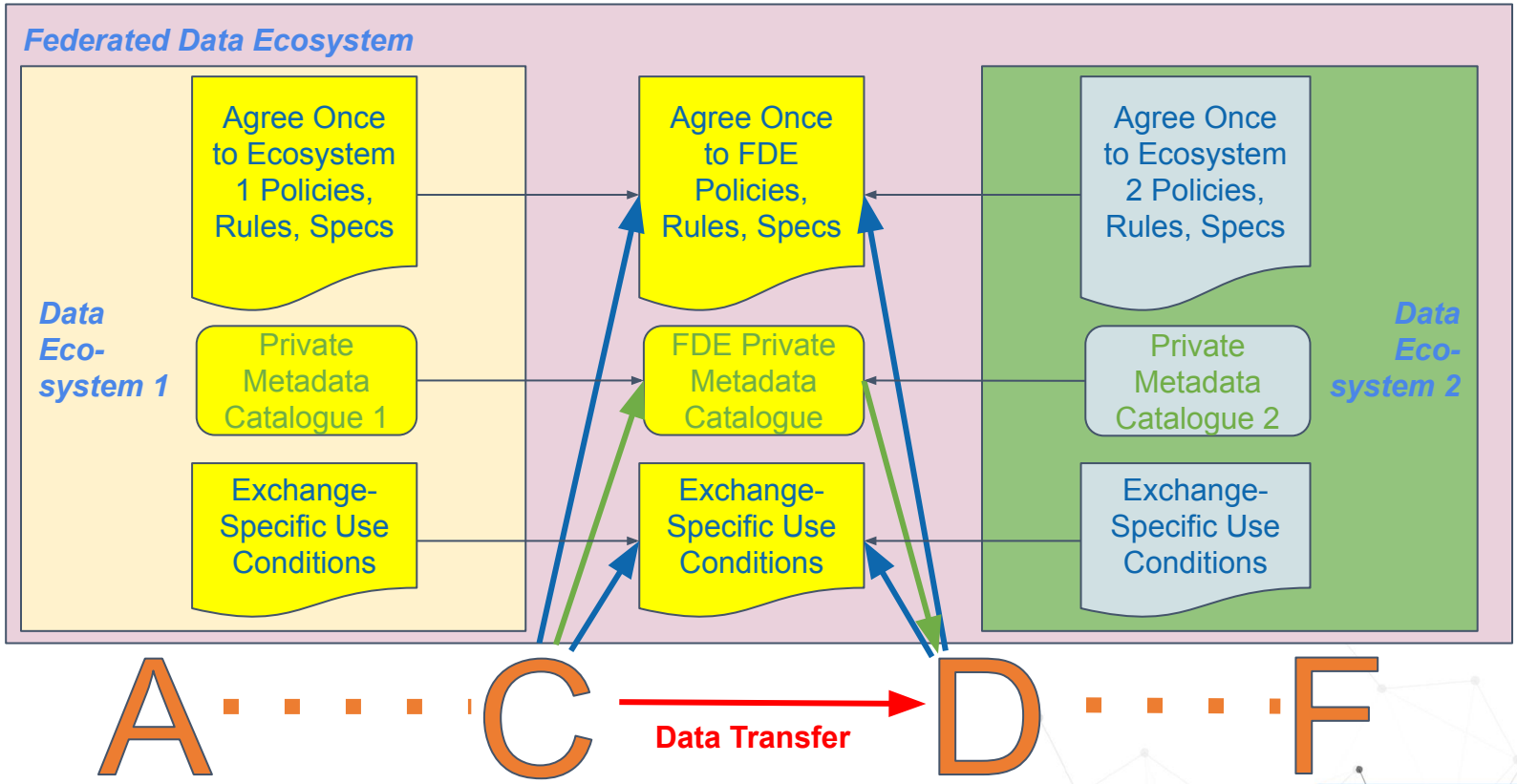
Across 2 Different Ecosystems

Several questions must be answered



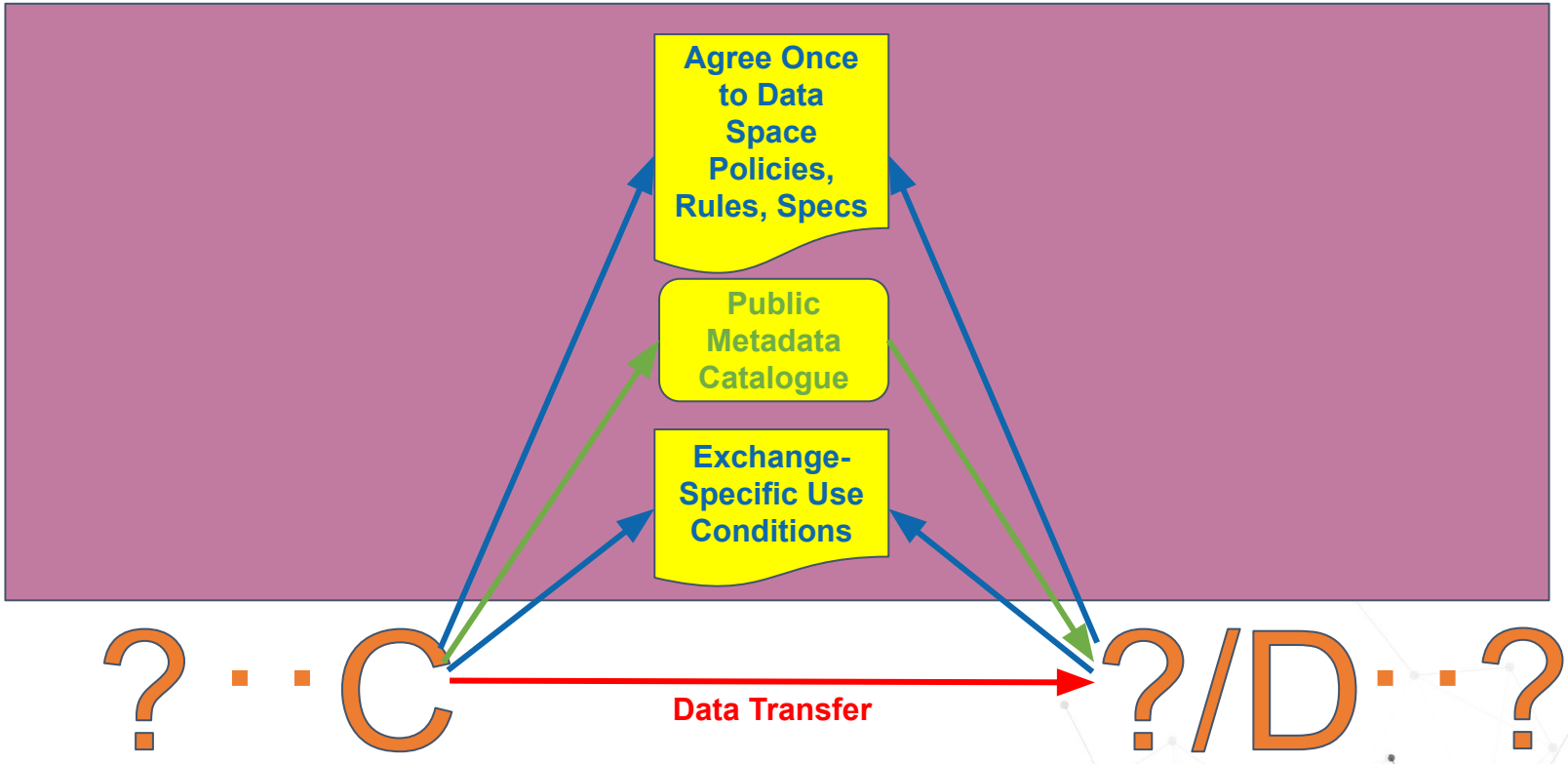
In a Federated Data Ecosystem

Systematic Agreement to those Questions



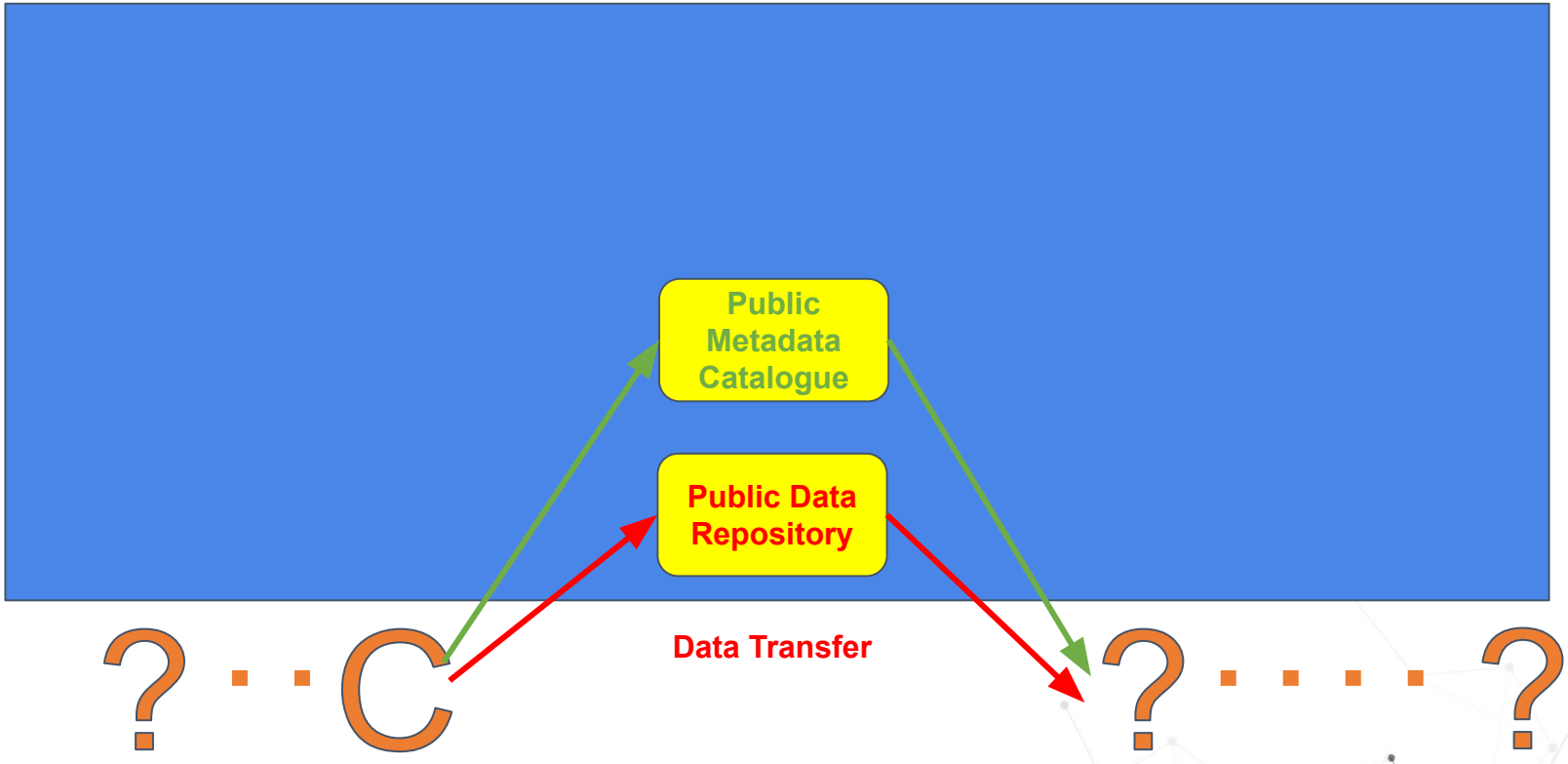
In a Public Data Space

Data Consumers only identified as part of exchange



By Publishing Open Data

Data and metadata visible to anonymous users



Important Issues

Data Sovereignty

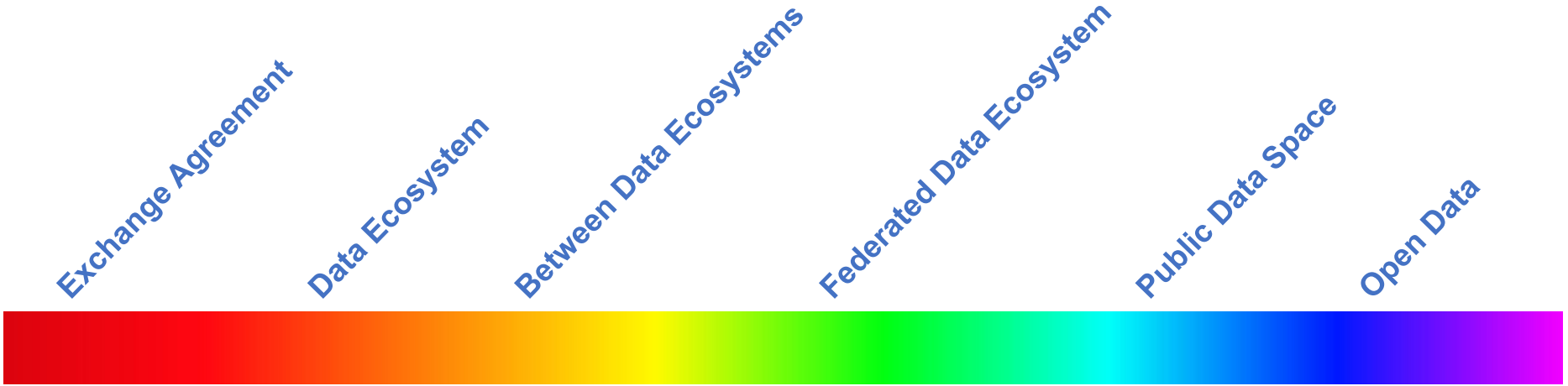
DEAs are separate and distinct.

Data is exchanged -- not shared, except when it is published openly.

Possible role of Data Intermediaries

Bringing compute to the data → a service ≠ not data exchange.

Specific DEAs along the Spectrum



many examples in both research and industry

EUROPEAN OPEN SCIENCE CLOUD

European Commission

data.europa.eu

gaia-x

EGI: Advanced Computing for Research

www.egi.eu
@EGI_eInfra

Thank you for your attention.

Questions?

This work was supported by the European Commission H2020 grant number 101017567 EGI-ACE (https://www.egi.eu/projects/egi-ace/), and H2020 grant number 871920 H-CLOUD (https://www.h-cloud.eu/).