More and more research results are achieved

by using **already existing** research data

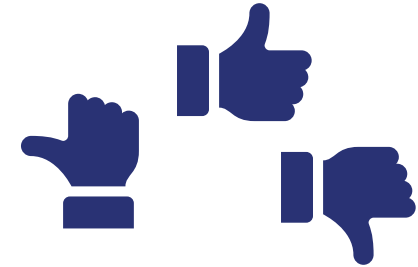# Currently, data are often only stored

decentrally

temporarily

with non-standardized meta-data

with varied quality

Thus, significant efforts are still required to

find the right data set, **understand** it and

use it for the purpose at hand

→ Need for better **research data management** in compliance with the **FAIR principles**

FAIR Data Principles

Findable – meaningful (meta-)data with unique identifier

Accessible – Communication protocol, supports authentication & rights management

Interoperable – common terminologies and ontologies

Reusable – Rich contextual (meta)data, provenance information

FAIR Data Spaces

German National Research Data Infrastructure (NFDI)

As of September 2022

216
Members

> 800
Involved
Persons

19
Consortia

4
Sections

NFDI Consortia & DFG Scientific Disciplines
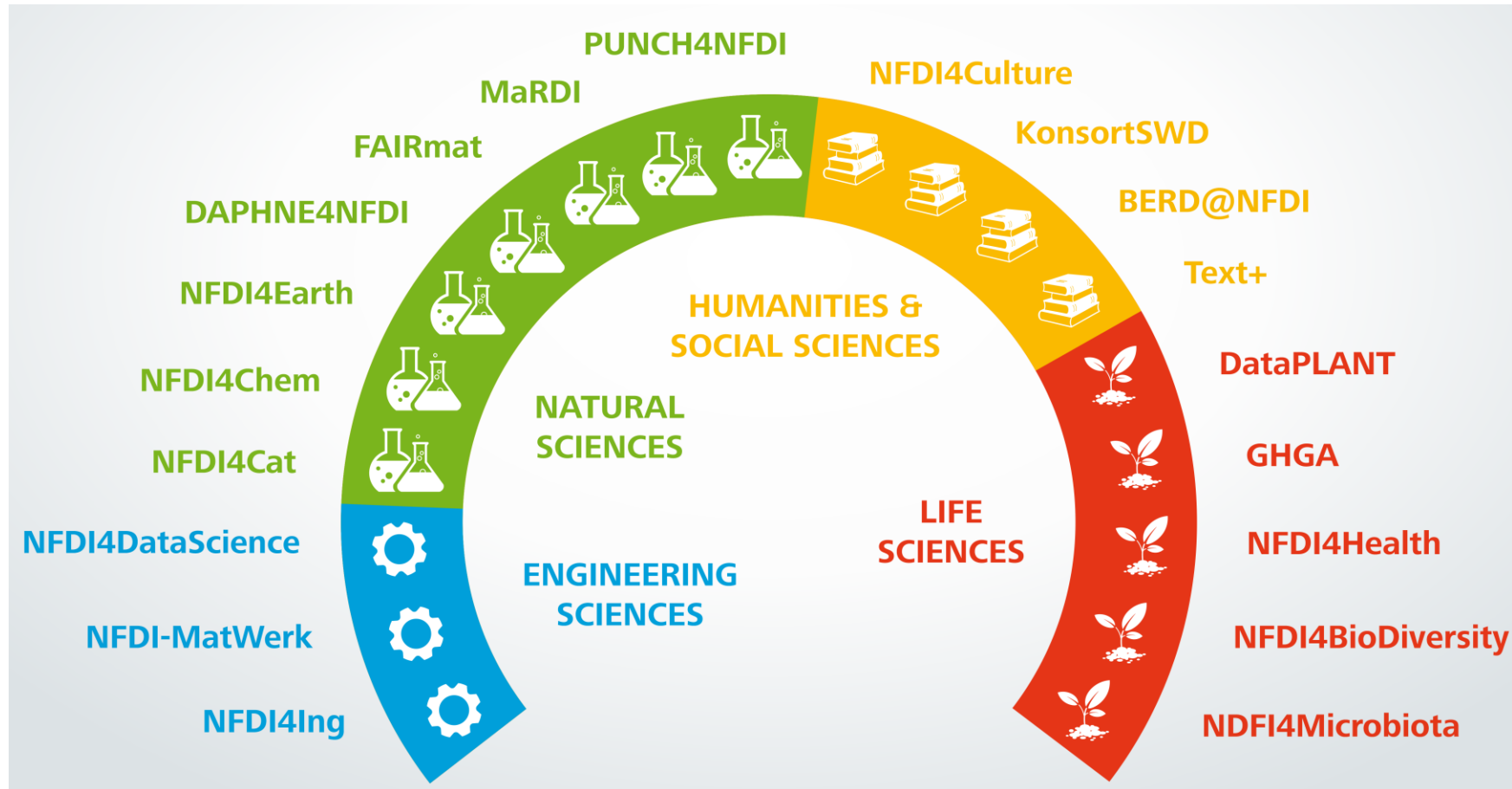
FAIR Data Spaces

NFDI network
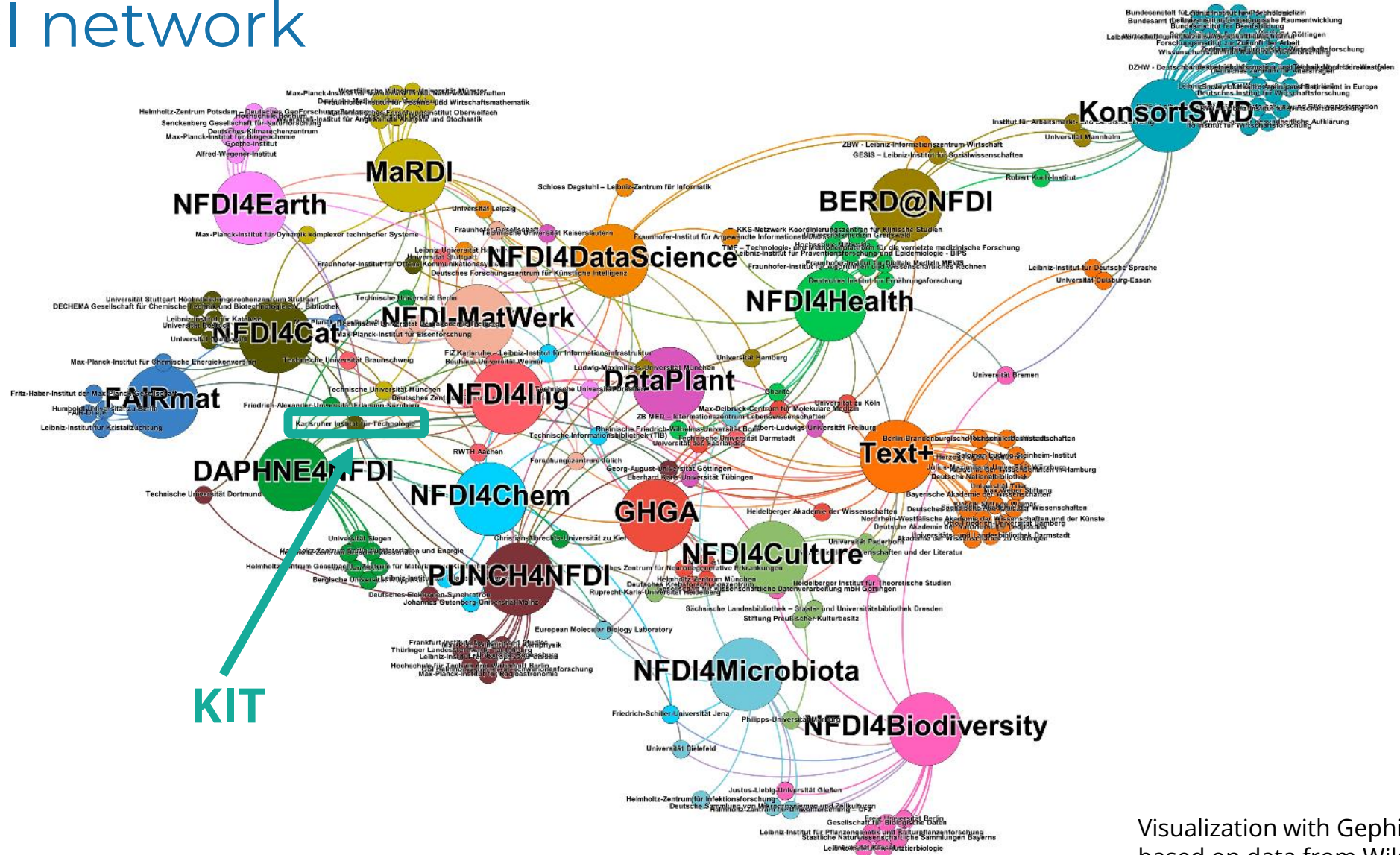
Visualization with Gephi
based on data from Wikidata

FAIR Data Spaces

8

Consortia with KIT participation
Examples

Currently 9 consortia with KIT participation

- NFDI4Ing
    - **Aim**: develop, disseminate, standardize and provide methods and services to make engineering research data FAIR

- NFDI4Cat
    - **Aim**: promote and support cross-disciplinary research in the field of catalysis

- FAIRmat
    - **Aim**: cover the full breadth of the Condensed Matter Section of the German Physical Society

- NFDI-MatWerk
    - **Aim**: integrate decentralized (meta)data, experimental and numerical workflows; establish a materials ontology

- **PUNCH4NFDI (P**articles, **U**niverse, **NuC**lei & **H**adrons for the NFDI)
    - **Aim**: establish a federated and FAIR science data platform that serves PUNCH community and entire NFDI

- furthermore: NFDI4Chem, NFDI4Microbiota, NFDI4Earth, DAPHNE4NFDI

PUNCH4NFDI website: https://www.punch4nfdi.de/
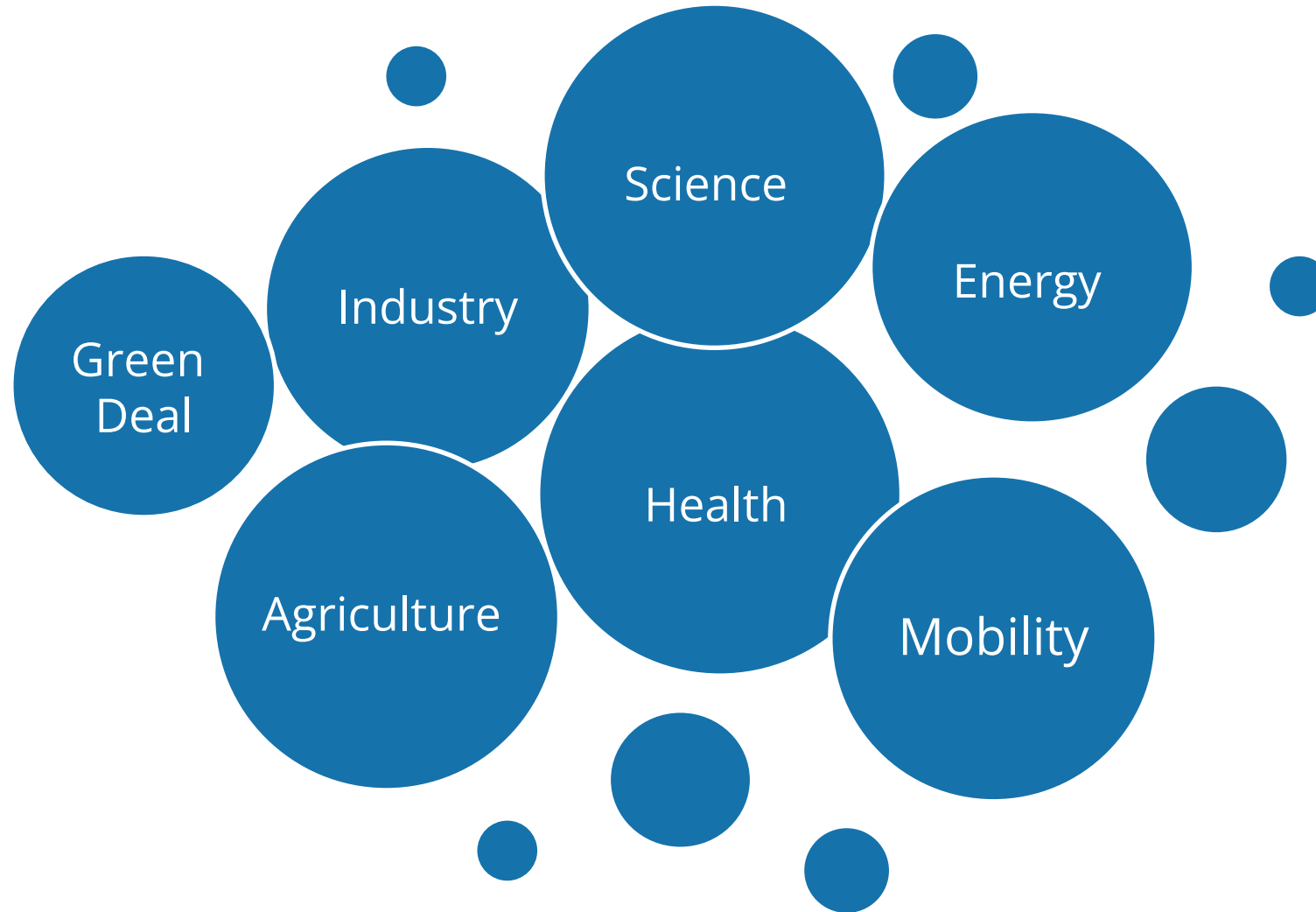
FAIR Data Spaces

European Data Strategy

Data can **flow** within the **EU** and **across sectors**

European **rules and values** are fully **respected**

Rules for access to and use of data are **fair, practical and clear**
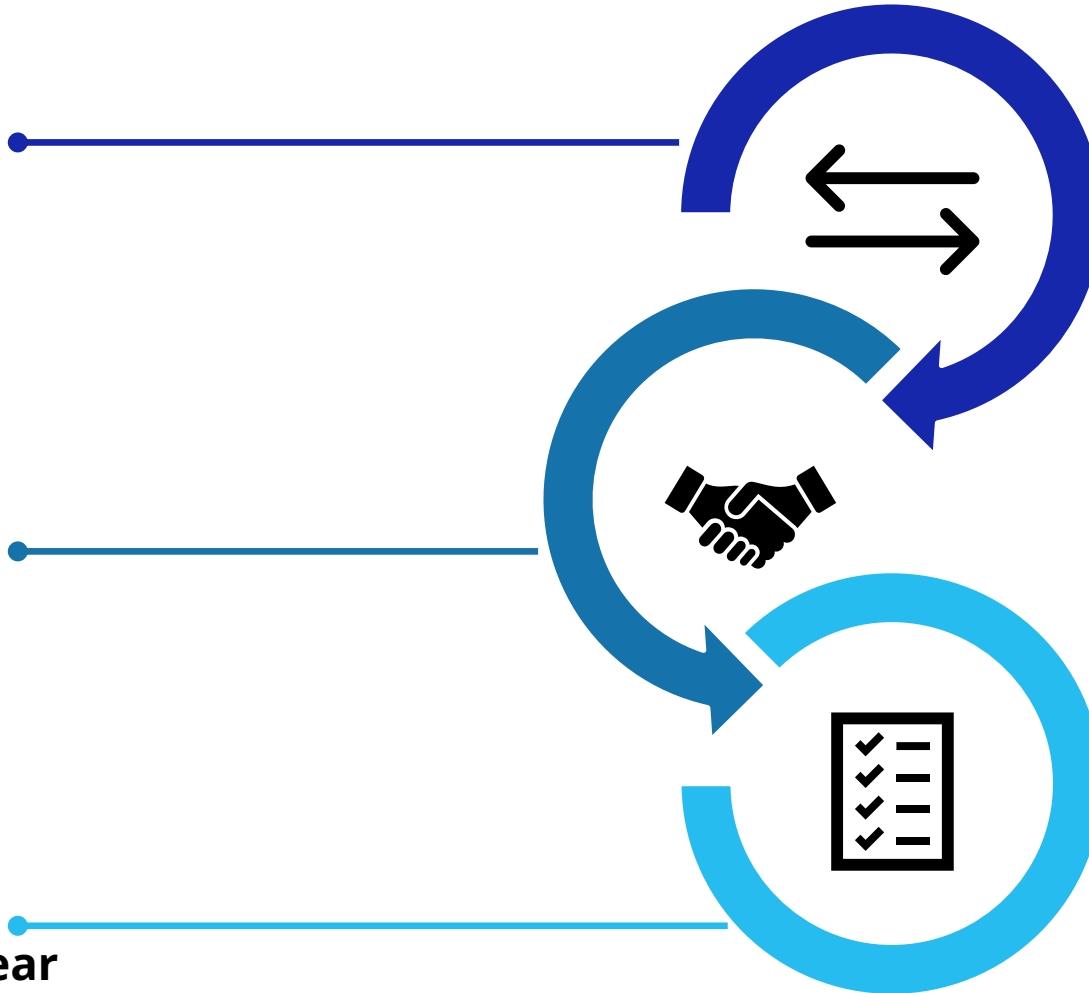
FAIR Data Spaces

Data spaces

Towards creating a data space economy
Gaia-X

- Gaia-X is a EU initiated project with 350+ participants from industry and academia

- Create an **open, transparent** and **secure, federated** data infrastructure by
  - Developing a software framework of control and governance
  - Implementing a common set of policies and rules
  - Deploying the framework on top of any existing cloud that joins the Gaia-X network

- **Linking cloud service providers** and **data providers** to form data spaces where participants exchange data in a trusted environment

- **Aligns with** the objectives of the **European data strategy**

Gaia-X website: https://gaia-x.eu/

FAIR Data Spaces

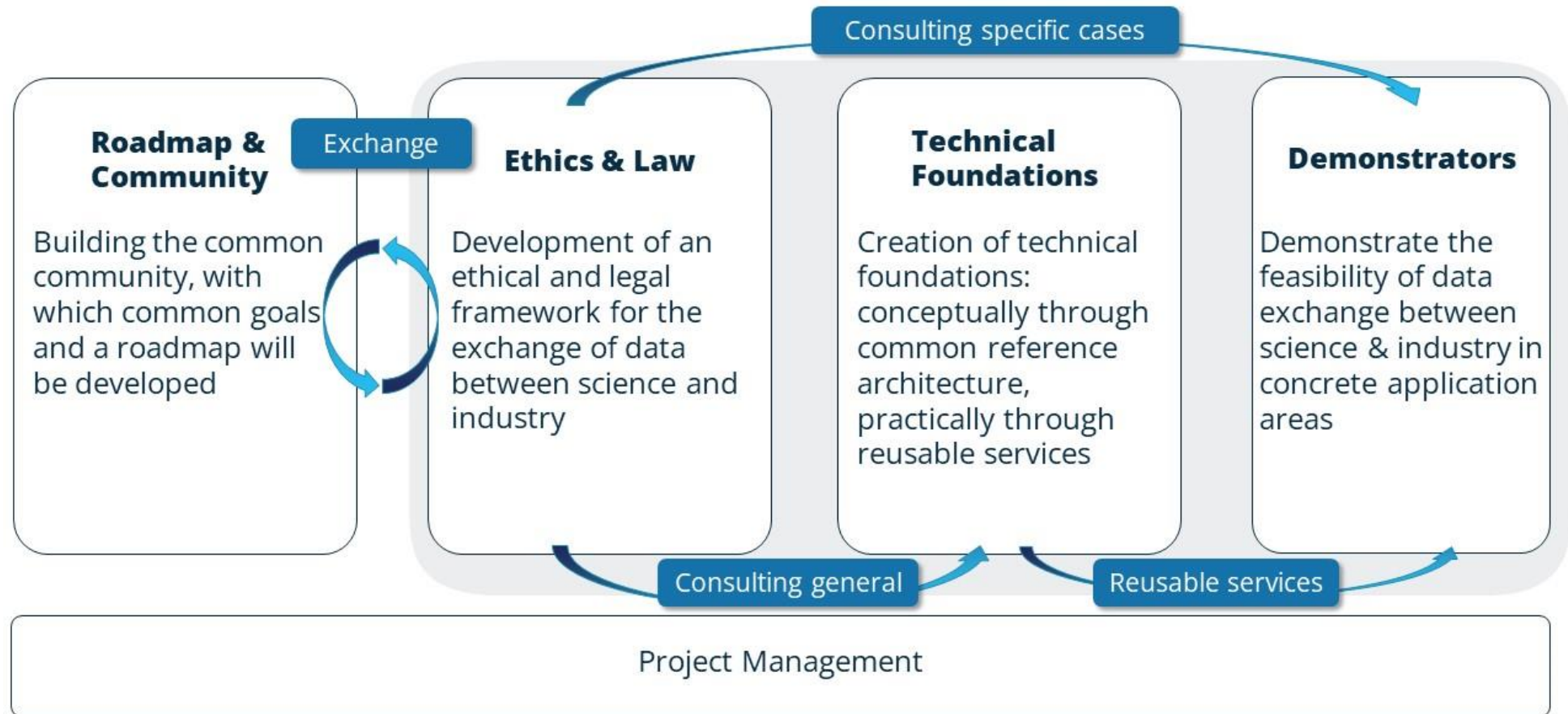# FAIR Data Spaces project
Connecting domains

- **Vision**: Development of a common cloud-based data space for industry and science (based on the FAIR Data principles) by linking the two initiatives Gaia-X and NFDI

- **Mission**: Create and expand synergies between existing technologies and communities

- **Goals**:
  - Interlinking the two domains by **clarifying legal and ethical issues**
  - Providing **technical foundations**
  - **Demonstrating** and promoting a sovereign exchange of data between industry and science both nationally and in the EU in concrete applications
  - Building a **common community** within the project and beyond

NFDI

Science

Industry

Gaia-X
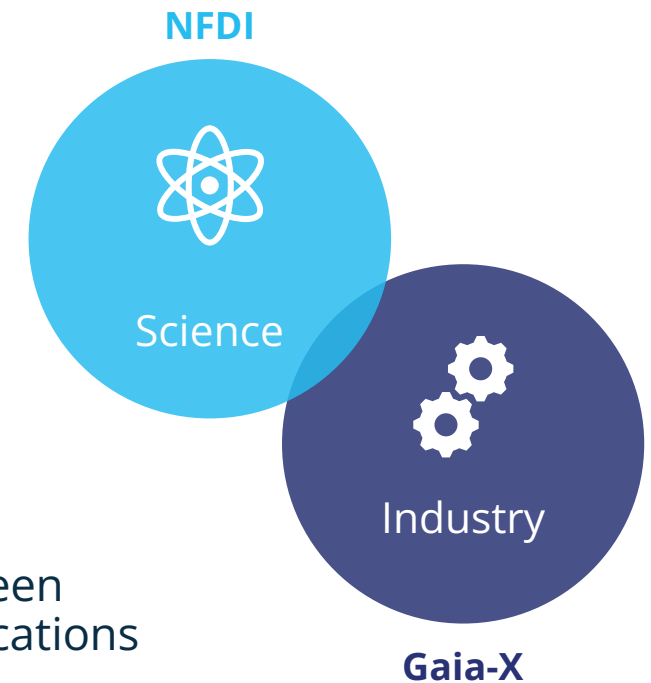
FAIR Data Spaces

FAIR Data Spaces at a glance

FAIR Data Spaces at a glance

What I do →

Consulting specific cases

Roadmap & Community — Exchange

Building the common community, with which common goals and a roadmap will be developed

Ethics & Law

Development of an ethical and legal framework for the exchange of data between science and industry

Technical Foundations

Creation of technical foundations: conceptually through common reference architecture, practically through reusable services

Demonstrators

Demonstrate the feasibility of data exchange between science & industry in concrete application areas

Consulting general

Reusable services

Project Management

Linking domains
Community Building

For a fruitful collaboration between domains, there needs to be

- a common **vision** and a **roadmap** with common goals
- a **balance** of the different **perspectives** of industry and research
- a **common community** to enlarge and strengthen the scope of the activities

This is achieved by

- building a **communication infrastructure**
- organizing **workshops** and events (nationally and on EU level)
- **communicating** project results to NFDI consortia and Gaia-X working groups

FAIR Data Spaces

FAIR Data Spaces Demonstrators
Showcase data exchange

- NFDI4Biodiversity & Gaia-X

  - **Aim**: Combine data from science and industry through Gaia-X compatible clouds using biodiversity and geodata

  - **Challenges**: Wide range of data acquisition methods and data formats

- FAIR Data Quality Assurance and Workflows (NFDI4Ing)

  - **Aim**: Develop automatic quality assurance of (largely) manually recorded data

  - **Challenges**: many participants, different programming languages and data formats

- Cross-Platform FAIR Data Analysis (NFDI4Health)
  - **Aim**: Demonstrate cross-platform privacy-compliant data analysis on distributed data sets without sharing the data itself
  - **Challenges**: data distributed across multiple sites, different data formats, sensitive personal data

FAIR Data Spaces

Summary

- Different **data economy ideas** and corresponding ecosystems are currently **forming**

- **Linking** the **developments** is important on domain and cross-domain level to benefit from each other

- By **agreeing on common principles and guidelines**, a fair and sovereign data exchange is possible

Daniela.Mockler@nfdi.de

#FAIRDataSpaces @FAIRDataSpaces

https://www.nfdi.de/fair-data-spaces/

https://www.nfdi.de/fair-data-spaces-newsletter/

https://zenodo.org/communities/fair-ds/

FAIR Data Spaces