INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Unraveling the Problems
in Building Public Sector Data Spaces

**Dominik Rozkrut, Barbara Batóg, Mariusz Doszyń, Paweł Majda,
Mirosław Pajor, Monika Rozkrut, Małgorzata Tarczyńska-Łuniewska**

July 3, 2025 — Polish Statistics Congress, Warsaw

Outline

- Introduction
- European Data Spaces
- National Quality Infrastructre
- NQI Data Space

Introduction

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Theses

- Investment in R&D and innovation has a positive impact on the economy
  - however, it is not efficient without the existence of necessary conditions for innovation, including the quality infrastructure
- National quality infrastructure is a critical enabler for firms to access and compete in domestic and foreign markets
  - improve trade and investment opportunities, boosts productivity and innovation, promote competitiveness
  - governments need to stimulate demand for quality, make continuous investments to upgrade their national quality infrastructure

National Quality Infrastructure contributes to policies

- Society
  - well-being, health, safety, consumer protection, food safety, governmental laws and regulations
- Economy
  - industrial development, fair trade, competitiveness, quality, profitability, manufacturing, distribution, purchasing, use, specifications, contracts
- Environment
  - efficient use of natural and human resources, environmental safety, climate change

Why NQI Data Space?

- A QIS is a dynamic system
  - **Dynamic** places the emphasis on actions. These actions are intended to provide results, and they are evaluated according to these results
  - **System** means its parts interact with each other to provide overall QIS results which are greater than could be achieved by the parts working individually

A Coherent Data Ecosystem

- The National Quality Infrastructure (NQI) significantly relies on information
  - possessing substantial data resources and ensuring information security
- Such a data space is essential for analyzing the functioning, influence, and impact of metrology on the economy and society
  - The primary goal is to develop a comprehensive information system for Polish NQI by establishing a new data architecture

European Data Spaces

Data Sharing and Utilisation at Scale

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Driving vision and policy context

- **Access to data** is paramount to achieving ambitious digital objectives.
- The "**Europe fit for the digital age**" priority of the European Commission
  - leveraging digital technologies without compromising fundamental rights, aiming for a human-centred digital transformation, "**Made-in-EU**"
- The European strategy for data, published in 2020, sets a vision to establish common **European Data Spaces** across all strategic societal sectors and domains of public interest

The goal

- to establish a **single market for data in the EU**, promoting trustworthy data sharing aligned with common societal values and existing legal frameworks

- to facilitate the **reuse of data** shared by businesses, citizens, public authorities, and academia, fostering trust and fairness in the data economy

European Data Spaces
Legislative framework

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Data Governance Act

- Introduces horizontal measures to boost trustworthy data sharing by
  - facilitating the reuse of public sector data,
  - regulating data intermediaries,
  - easing data availability for businesses and citizens, and
  - facilitating cross-sectoral and cross-border data sharing

Data Act

- Aims to make more business data available for reuse by
  - defining rules on data access and usage
  - allowing users of connected devices to access their generated data
  - public sector access to private company data in emergencies
  - rebalancing negotiation power for SMEs
  - facilitating switching between cloud data processing services

Implementing Act on High Value Datasets

- Implements the **Open Data Directive** by
  - specifying a list of **public sector datasets** to be made available for free
    - under open access licenses, and in machine-readable formats via APIs.

Digital Markets Act

- Regulates the "**gatekeeper power**" of digital companies
  - setting measures on data access and portability
  - to prevent unfair business practices

Data space regulations

- The horizontal legal frameworks may be complemented by sector-specific legislation, such as the proposed Regulation on the **European Health Data Space**
  - a common framework for health data exchange and access, promoting secure, standardized, and efficient use of health data across EU member states
  - ongoing legislative process, expected implementation: 2025-2026

Artificial Intelligence Act (AI Act)

- It does not grant any rights to individuals
  - but regulates providers of artificial intelligence systems and entities using them for professional purposes
- Low-risk applications are not regulated at all
  - and a voluntary code of conduct is envisaged
- Medium and high-risk systems will require
  - mandatory conformity assessment by the supplier as part of a pre-market self-assessment

European Data Spaces
Definition

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Definition and characteristics

- **Sovereign, trustworthy, and interoperable data sharing** environments where data can flow within and across sectors, respecting European rules and values
- **Bring together relevant data infrastructures and governance** frameworks to facilitate data pooling and sharing
- **Decentralized infrastructures** where actors can share, access, and use data securely, reliably, and trustworthily, following shared governance, organizational, regulatory, and technical mechanisms

Definition and characteristics

- To overcome **legal, organizational, semantic, and technical barriers** to data sharing among diverse actors by combining necessary tools and infrastructures with common rules and standards

- To connect **fragmented and dispersed data** from various ecosystems to support EU priorities, create new business opportunities, and facilitate the use of data for innovation

Definition and characteristics

- Key components
  - **data-sharing tools and services**
    for pooling, processing, and sharing
  - compatible
    data governance structures
  - **improved data**
    availability, quality, and interoperability

Guiding principles

- Mind-set for sharing data in Europe based on common societal values should be propagated throughout the data value chain

- **Key principles**

  - data sovereignty, transparency, security, fairness, consumer protection, fundamental rights, citizen centricity, data altruism, inclusion, sustainability, openness, self-determination, trust, fair competition, and innovation

European Data Spaces
Requirements and features

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Functional requirements

- Data transfer and exchange
- Data publication and discovery
- Data Storage (physical or virtual)
- Identity, authentication, and access control
- Data interoperability (legal, organisational, technical, semantic)
- Data processing and analytics (including AI support)
- Data pooling and collaboration

Functional requirements

- Multi-tier support, federation, and orchestration
- Privacy-preserving mechanisms (e.g., anonymisation)
- Usage control policies
- Compliance and auditing
- Transaction metering and billing
- Data governance
- Data protection (compliance with GDPR)

Non-functional requirements

- Security and confidentiality
- Interoperability
- Maintainability
- Variability
- Scalability
- Performance
- Auditability
- Portability
- Energy efficiency
- Inclusivity
- Fairness
- Sustainability
- Trustworthiness
- Transparency
- Data sovereignty

Features

- Implement requirements such as a secure and privacy-preserving infrastructure for pooling, accessing, sharing, processing, and using data
- Examples
  - federated infrastructures, agreed licensing frameworks, privacy-preserving environments, electronic identification, adoption of open source software, semantic assets, validation services, standard vocabularies

European Data Spaces

Sector-Specific Data Spaces

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Domain-specific common European Data Spaces

- The European strategy for data promotes EU-wide common, interoperable **data spaces** in strategic economic sectors and public interest domains
  - acknowledged to have differences but should be interconnected to simplify cross-sector data exchange and enable cross-exploitation and collaboration for innovation.
  - build upon existing data sharing initiatives to improve interoperability by introducing a harmonised framework for data exchange and a horizontal governance framework

Domain-specific common European Data Spaces

- Health Data Space
- … Manufacturing Data Space, Green Deal Data Space, Mobility Data Space, Financial Data Space, Energy Data Space, Agriculture Data Space, Data spaces for Public Administrations, Skills Data Space …
- European Open Science Cloud (EOSC)

European Data Spaces

Key actors and support structures

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Key actors

- **Relevant stakeholders**
  - from the private sector, public sector, academia, civil society, and individuals, engaged in the creation, maintenance, and governance
  - include data producers, data receivers, data rights holders, data users, data service providers, data intermediaries, and data space enablers

Key actors

- **European Data Innovation Board**
  - established by the Data Governance Act, will support the Commission in issuing guidelines for data space development and identifying relevant standards and interoperability requirements
- **Data Spaces Support Centre**
  - funded by the DIGITAL program, it coordinates actions on sectoral data spaces, provides blueprint architectures, best practices, and common data infrastructure requirements

Support structures

- Logical links exist with regional, national, and transnational initiatives like
  - **Big Data Value Association** (BDVA)
  - **Data Spaces Business Alliance** (DSBA)
  - **International Data Spaces Association** (IDSA)
  - **GAIA-X**
  - **Digital Transport and Logistics Forum** (DTLF)

Funding

- The European Commission supports data space creation through programs like
  - **Digital Europe Programme** (DIGITAL)
  - **Horizon Europe**
  - **Connecting Europe Facility** (CEF)
  - **NextGenerationEU's Recovery and Resilience Facility**

National Quality Infrastructure
Overview

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

National Quality Infrastructure

- A triad, or triple helix of
  - **Quality**
  - **Safety**
  - **Trade**

Definitions

- The set of country institutions that provide the scientific and regulatory framework and services to guarantee the quality and safety of products and services offered in local and foreign markets

- The ecosystem of public and private institutions together with the policies, relevant legal and regulatory framework, and practices needed to support and enhance the quality, safety and environmental soundness of goods, services, and processes.

- A combination of initiatives, institutions, organizations, activities and people, includes a national quality policy and institutions to implement it, a regulatory framework, quality service providers, enterprises, customers and consumers

The complementary nature

- Metrology
- Standardization
- Accreditation
- Conformity assessment
- Market surveillance

Standardization

- Standards are documented agreements that translate desired characteristics into dimensions, tolerances, weights, processes, systems, best practice and other specifics
  - so that products and services that conform to their requirements provide confidence to buyers and users
- Standardization is usually the responsibility of a National Standards Body
  - by using international standards, it ensures that the country benefits from international, state-of-the-art knowledge

Metrology

- Assess products and services in a reliable and objective manner using appropriate measurement and scientific tools
- The science of measurement; accurate measurements and measuring equipment are needed
  - for the protection of health, safety, the environment, and consumers.;
  - vital too in contracts between individual business partners and in world trade in general
- instruments in laboratories need to be properly calibrated so that they can provide reliable measurements

Accreditation

- Ensure that all are playing equitably through competent technical verification
  - formal recognition that a body or person is competent to carry out specific tasks;
  - evaluate the competence of product, management system, and personnel certification bodies, testing laboratories, and inspection bodies
  - often the responsibility of a National Accreditation Body (NAB)

Conformity assessment

- Demonstration that specified requirements relating to a product, process, system, person, or body are fulfilled
  - processes and procedures that are used to demonstrate that a product or a service, a management system, an organization or personnel meet specified requirements
  - these requirements are usually stated in international standards developed by organizations such as ISO (International Organization for Standardization)
- Helps to ensure that products deliver on their promises

Key institutions

- **Polish Committee for Standardization** (PKN)
  - Responsible for developing, publishing and implementing national standards, including standards harmonized with European and international standards.
- **Polish Center for Accreditation** (PCA)
  - National accreditation body, granting accreditation to research and calibration laboratories, certification and inspection bodies
- **Central Office of Measures** (GUM)
  - The highest state institution in the field of metrology, responsible for maintaining and developing national standards of measurement units

Key institutions

- Office of Competition and Consumer Protection (UOKiK)
  - Supervises the conformity assessment system, including CE marking, and protects consumer rights, ensuring compliance with product requirements
- Network of research and calibration laboratories
  - Public and private laboratories with PCA accreditation, conducting quality tests and calibration of measuring instruments
- Certification and inspection bodies
  - Organizations with PCA accreditation, issuing certificates of conformity, conducting inspections, and quality audits

NQI Data Space
Technical Aspects

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Software stack for a data space

- No universal solution; choices involve **proprietary monolithic** applications versus **modular open-source** tools
  - open source is recommended due to its transparency, facilitation of innovation, and early adoption of new standards
- The criteria for selection
  - support for international standards, vitality of the software community, clear contribution guidelines, and ease of maintenance and evolution
- The **European Union Public License** is recommended for new software.

Data interoperability

- **Data interoperability** crucial for data reuse and is achieved through adopting common specifications and open standards for documentation, modelling, encoding, and sharing data

- The **European Interoperability Framework** provides a solid foundation across legal, organisational, semantic, and technical dimensions

Licensing

- Licensing information should always be available for all datasets
  - **Standard licenses**, like Creative Commons (CC) (e.g., CC0 or CC-BY), are recommended to maximise reuse benefits
  - **Machine-readable licenses** facilitate automated data exchange and use

Promoting standards

- **Adherence to the requirements** of selected standards is essential for efficient data sharing
- **Automated validation techniques** using specialised tools are favoured over self-declaration for compliance with technical specifications and legal requirements.

Data discovery

- **Standardised metadata** (machine and human-readable, e.g., SDMX) and data/metadata sharing through **Application Programming Interfaces** (APIs) substantially improve findability
- Providing data in multiple encodings, including web-native formats like **HTML**, enables search engine indexing
- Annotating data with commonly agreed-upon vocabularies (e.g., **SEMIC Action**, schema.org, SDMX) further improves web findability

Data standardization

- Prioritize well-known standards adopted by global communities and developed by established, trusted international standards development organisations (SDOs)
- Consider the maturity of standards, preferring those with a large user base, multiple implementations, and strong software support
- "De facto" standards should also be taken into consideration

APIs

- Web-based **Application Programming Interfaces** (APIs) act as "glue" connecting different actors and systems
- APIs should
  - self-describe possible interactions, accessible data, and conditions of use
  - following widespread industry standards like OpenAPI specifications

Privacy-enhancing technologies

- Privacy-preserving techniques (PETs) must be used to protect personal and sensitive business data in a data space
- Examples
  - Pseudonymisation, Secure Multi-Party Computation (SMPC), Differential Privacy (DP), Homomorphic Encryption (HE), Federated Learning (FL), Zero Knowledge Proofs, Trusted Execution Environments (TEEs)
- Selection requires assessing the risk-benefit ratio and the Technology Readiness Level (TRL) of each solution

Data enhancement

- **Synthetic data** artificially generated from original datasets, maintaining similarity while not including original records, thus preserving analytical value without breaching privacy
- Help overcome challenges with
  - **privacy constraints** (e.g., health data)
  - **gaps in representativeness** (e.g., minority demographics)
  - **data augmentation for AI training** (e.g., autonomous vehicles)

Essential data operations

- Data Transfer and Exchange
- Data Processing and Analytics
- Data Pooling and Collaboration
- Data Publication and Discovery
- Usage Control Policies

NQI Data Space
Conclusion

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Robust governance framework

- A practical governance framework is essential for configuring and coordinating the necessary actions by different organizations to achieve the data space's objectives
- It must outline clear duties, standards, and responsibilities, ensuring data protection while actively supporting data sharing and mobility
- This framework incorporates legal, organizational, semantic, and technical interoperability dimensions, drawing upon frameworks like the European Interoperability Framework
- The NQI project will develop recommendations for data acquisition and processing methods, specifically for key stakeholders such as the Main Office of Measures (GUM) and other NQI entities

INSTITUTE OF ECONOMICS
AND FINANCE
UNIVERSITY OF SZCZECIN

Thank you!