Finansē Eiropas Savienība NextGenerationEU

2027 Nacionālais attīstības plāns

Data Curators at the Intersection of IT, Research, and Data – Latvian Experience of Building the Research Ecosystem

1862 RĪGAS TEHNISKĀ UNIVERSITĀTE

RĪGAS STRADIŅA UNIVERSITĀTE

LATVIJAS UNIVERSITĀTE

Latvijas Biozinātņu un tehnoloģiju universitāte

VPC AUGSTĀKĀS IZGLĪTĪBAS UN ZINĀTNES IT KOPLIETOŠANAS PAKALPOJUMU CENTRS

Project «Support for implementing Open Science, developing solutions for shared use of research data, and participation in the EOSC» (RRF 2.1.3.1.i.)

About Rīga Stradiņš University

- Rīga Stradiņš University (RSU) has been recognized as the university with the best reputation in Latvia for almost a decade now
- Over the past seven decades, we have evolved from being a local medical institute to being a modern European university in the fields of **medicine, healthcare, and social sciences**

9 FACULTIES

- Faculty of Pharmacy
- Faculty of Dentistry
- Faculty of Law
- Faculty of Medicine
- Faculty of Public Health and Social Welfare
- Faculty of Communication
- Faculty of Rehabilitation
- Faculty of European Studies
- Faculty of Residency

11 755 students	61 study programmes	~700 students undertaking Graduate Medical Training	82 student source countries
8 programmes in English	3068 international students	>150 doctoral students	35 study bases

RĪGAS STRADIŅA UNIVERSITĀTE

Data Steward Team at RSU

Viola Daniela Kiseļova

Līna Lāže

Laura Badūne

Agate Jarmakoviča

Rīgas Stradiņa universitāte

datukuratori@rsu.lv

About Us - Data Stewards

Lina Laze

Now:

- Data steward at RSU
- Pharmacology lecturer at RSU

Previously:

- Senior expert in State Agency of Medicines of the Republic of Latvia
- Senior expert in Latvian Council of Science
- Community and hospital pharmacist

Education:

- Degree in pharmacy and clinical pharmacy

Viola Daniela Kiselova

Now:

- Data steward at RSU
- Lecturer at Institute of Public health, RSU

Previously:

- Data analyst at Institute of Microbiology and Virology, RSU
- Health promotion coordinator at Centre for Disease Prevention and Control of Latvia

Education:

- Currently – MS in Biostatistics
- BSc in healthcare science (Public health)

Main Objective- Open Science

The ethics of open science are based on the belief that wider access to controlled scientific data will contribute to the accumulation of richer knowledge

Goal: to make scientific research, data, code, and publications **freely available to everyone without restrictions**

Result: research **transparency, collaboration** and **reproducibility** to increase the **accessibility and impact of scientific knowledge**

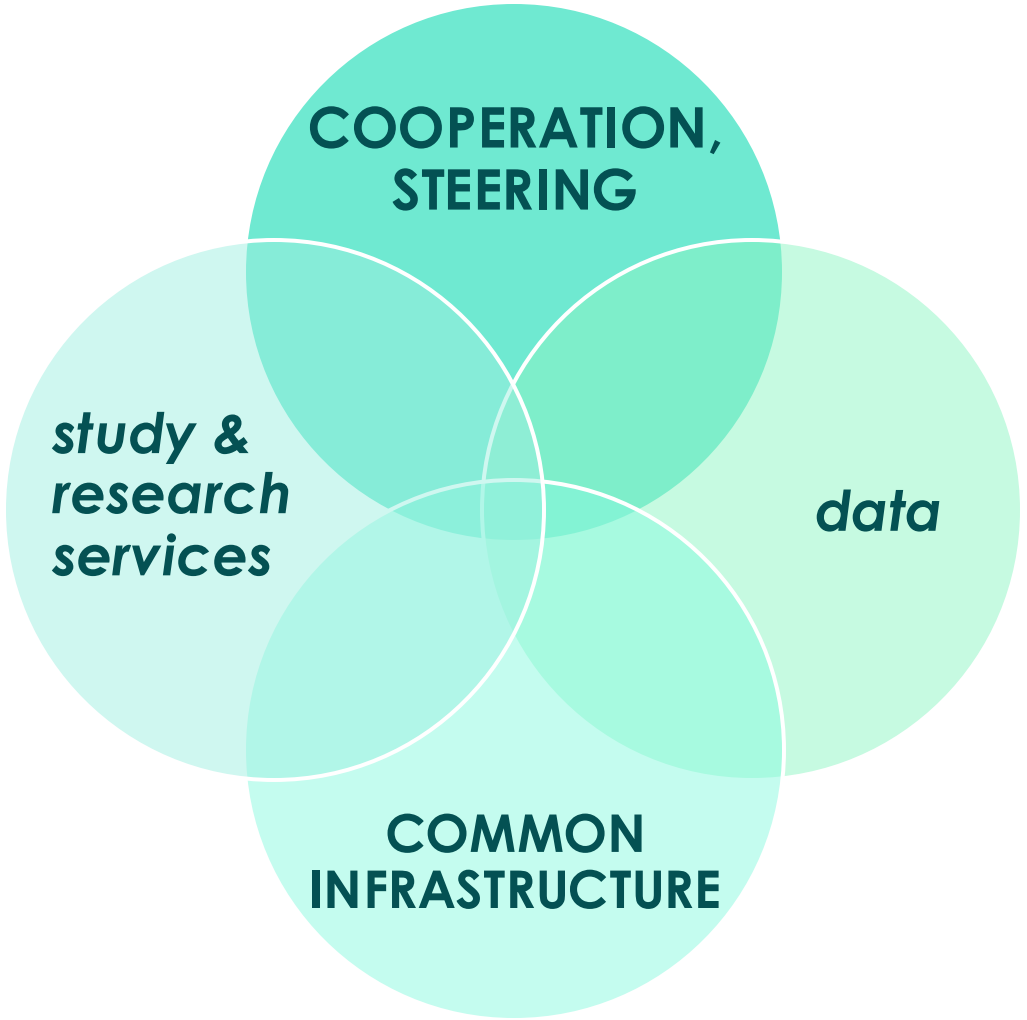
Open Science in Europe

- FAIR Data
- Research Integrity
- Next Generation Metrics
- Future of Scholarly Communication
- Citizen Science
- Education and Skills
- Rewards and Initiatives
- European Open Science Cloud (EOSC)

What is VPC?

VPC.LV – organisation of four Latvian Research Universities to ***integrate, develop and provide high-quality shared IT services in higher education***

similar to csc.fi, surf.nl, snd.se



VPC Vision for Comprehensive Data Services

Searching

- identifying data needs
- searching for available data
- working with data sources

Formalities

- data specifications
- data requests, permits
- contracts, procedures, requirements
- payments and reports
- GDPR, IP, etc

Preparing

- merging, cleaning, anonymizing
- preparing datasets

Processing

- data storage and access
- tools: analytical, visualization, programming, AI, HPC
- Secure Processing Environments

Repository, reuse

- research data repository *(DataverseLV)*
- data quality
- promoting available data
- data catalogues

Support

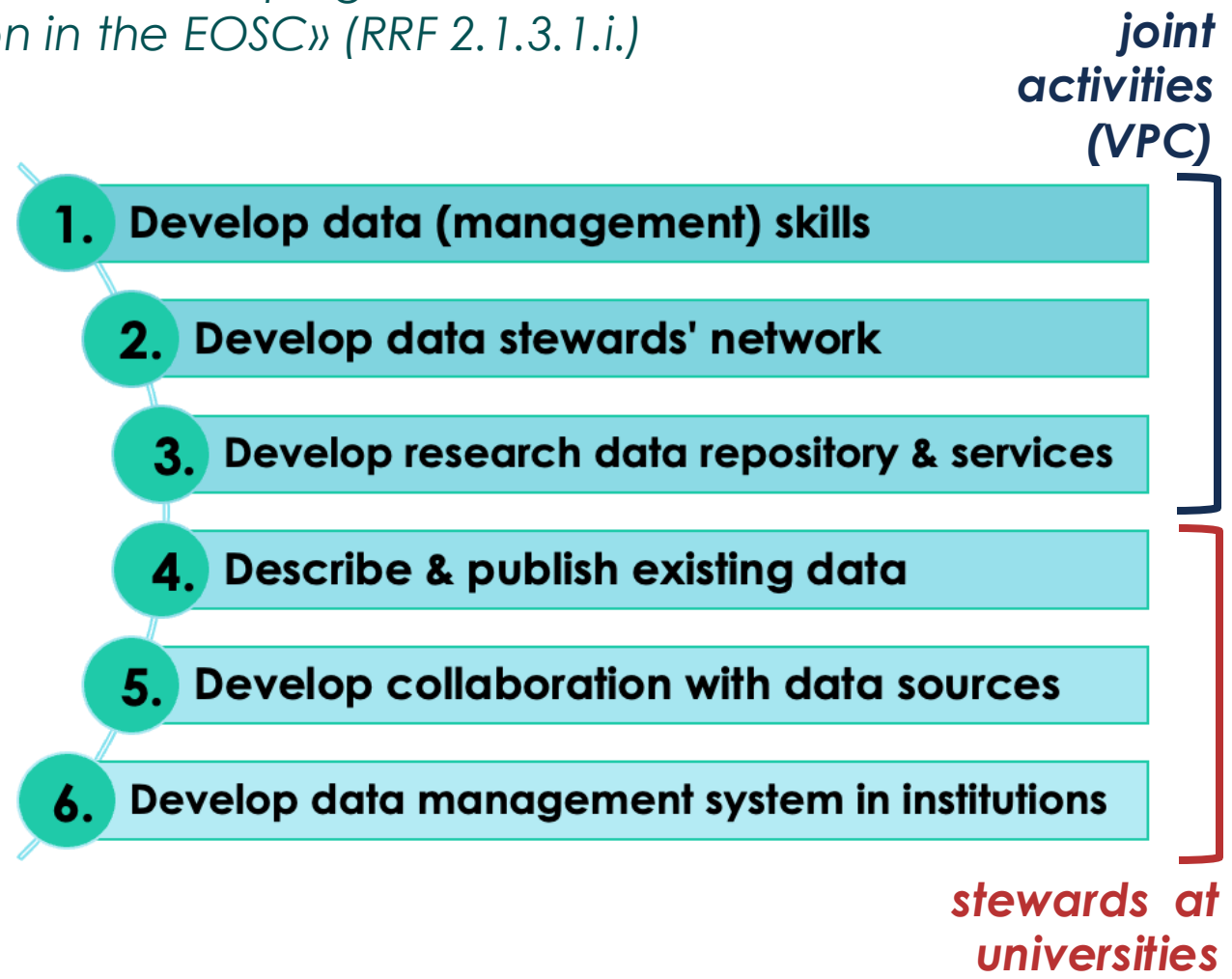
- consulting, training, outreach, collaborations
- standard data policies, procedures, requirements, solutions
- standard contracts, payments, processes

EOSC National Node integration w/ EU Data Spaces

“Data Stewards” project

project «Support for implementing Open Science, developing solutions for shared use of research data, and participation in the EOSC» (RRF 2.1.3.1.i.)

- goals:
  - repository
  - data management systems
  - capacity development
  - data stewards’ network
- 2024-2026, €3.4m
- 4-5 stewards at each university, 5-6 at VPC



The Role of Data Stewards

Development of data management guidelines and regulations

Bridge between the researchers and other departments (IT, legal ect.)

Organising trainings, seminars

Data quality assassurance

Support for researchers

eosc

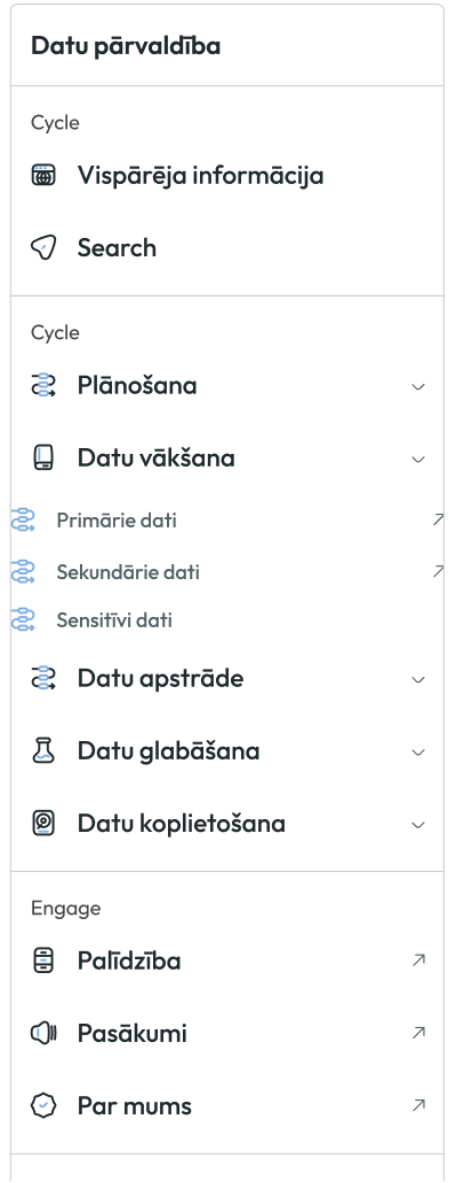
International collaboration and networking

Dataverse

DataverseLV university instance administration

What's Done

- ✓ the data stewards' network established
- ✓ self-studies, materials ~100hr
- ✓ ~490 DMPs, ~250 consultations, ~10 events
- ✓ RDM situation analysis + plans in 4 universities
- ✓ institutional RDM regulations, information *(RDM guidebook)*
- ✓ DatavarseLV *(pilot)*
- ✓ training at the DCC , University of Edinburgh



Data Steward Competency Training in Edinburgh

THE UNIVERSITY of EDINBURGH

DCC

What's Next

- developing specific RDM topics:
  - anonymisation, pseudonymisation
  - IT tools – availability, regulations, instructions, support
  - repositories, metadata, documenting data
  - data quality & security
- special data *(GDPR, health, public sector, classified, IP, commercial, etc.)*
- outreach, international networking
- preparing and depositing datasets
- CoreTrustSeal certification for DataverseLV

DS services for other institutions:
- DM policies & regulations
- IT tools
- DM consultations

datukuratori@vpc.lv

Our Network of Data Stewards

RĪGAS TEHNISKĀ UNIVERSITĀTE

datukuratori@rtu.lv

RĪGAS STRADIŅA UNIVERSITĀTE

datukuratori@rsu.lv

LATVIJAS UNIVERSITĀTE

datukuratori@lu.lv

Latvijas Biozinātņu un tehnoloģiju universitāte

datukuratori@lbtu.lv

VPC

AUGSTĀKĀS IZGLĪTĪBAS UN ZINĀTNES IT KOPLIETOŠANAS PAKALPOJUMU CENTRS

datukuratori@vpc.lv

What Happens When Data is Unorganized?

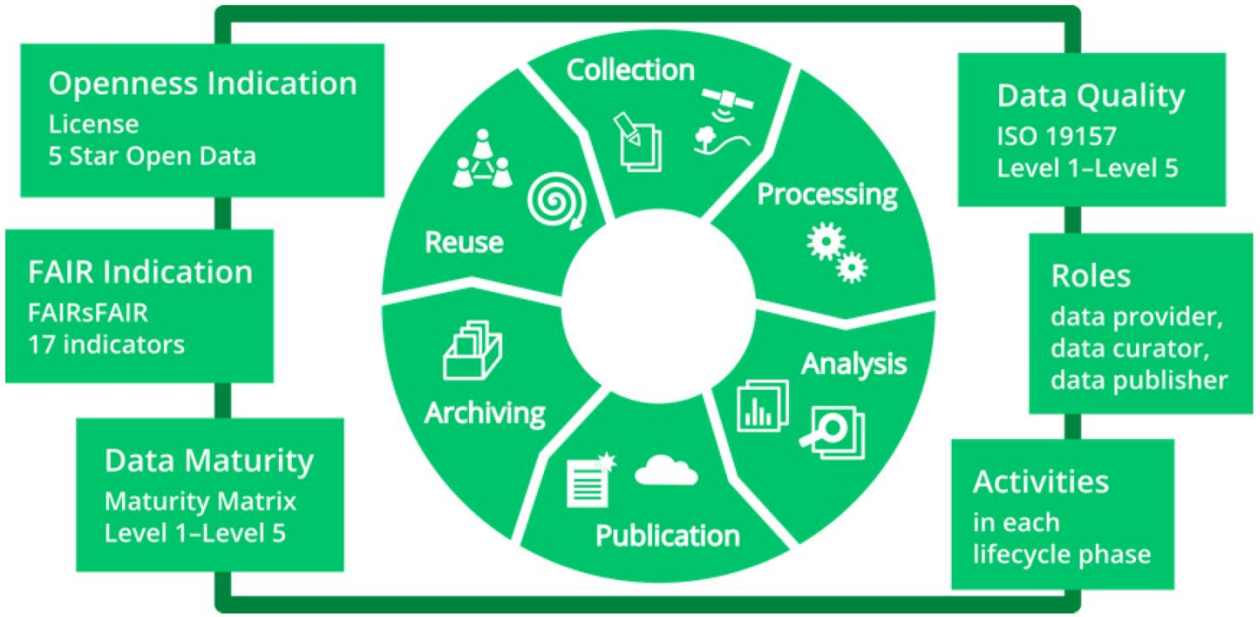
"Without a systematic way to start and keep data clean, bad data will happen." /Donato Diorio/

What is a Data Steward?

Data steward is a professional responsible for managing, organizing, and ensuring the quality and accessibility of data throughout its lifecycle.

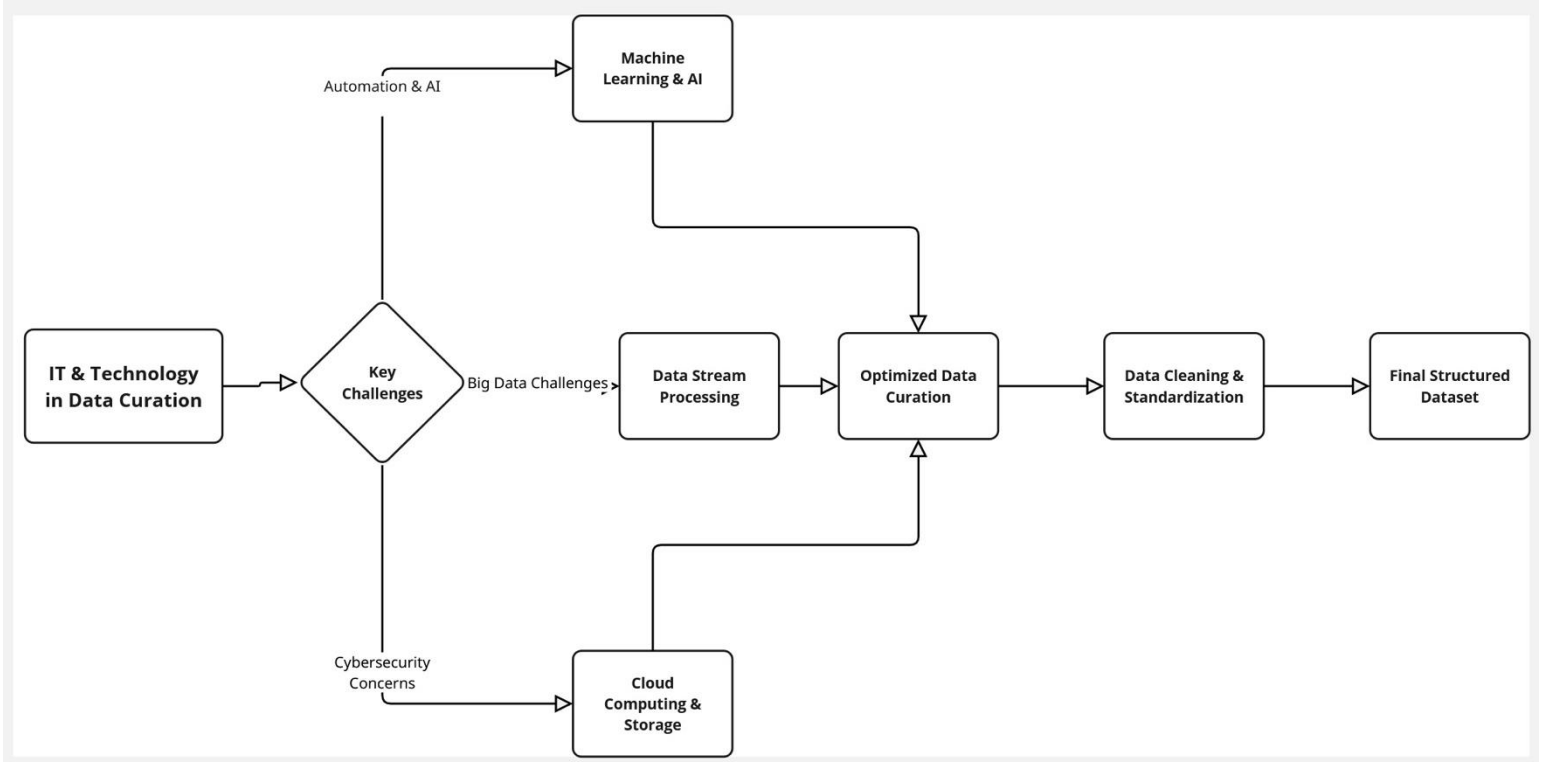
Role of Data Stewards in the Research Ecosystem

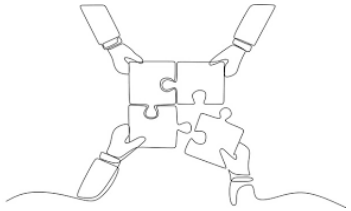
- Data Organization & Management
- Quality Control
- Metadata Creation
- FAIR Principles
- Data Preservation
- Collaboration with Researchers



IT and Innovation in Data Curation

How IT-driven solutions streamline data curation, making it more efficient, scalable, and secure in modern data-intensive environments?





Challenges

Scalability issues

Metadata & Interoperability

Data privacy & Security

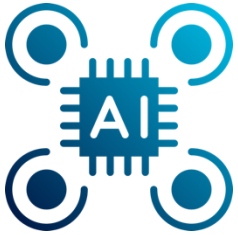
Long-Term Data Preservation

Data Quality & Integrity

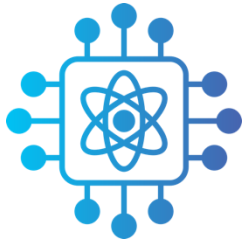
Data Bias & Representation

Cost & Resource Constraints

Future Directions & Innovations



AI-driven Automation



Interoperability & Standardization

Blockchain for Data Integrity

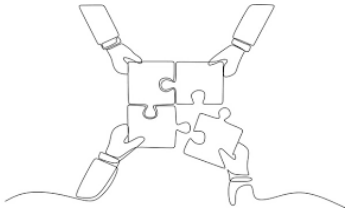
Quantum Computing for Big Data Curation

Edge Computing in Data Curation

Conclusions

- Ensuring Research Integrity & Reproducibility
- Facilitating Collaboration & Accessibility
- Managing Big Data Challenges
- Enhancing Security & Compliance
- Shaping the Future of Research

"Data curation is the foundation of reliable, accessible, and impactful research. Investing in technology-driven curation ensures scientific progress and global collaboration."



RDM practices at Rīga Stradiņš University

Finansē Eiropas Savienība NextGenerationEU

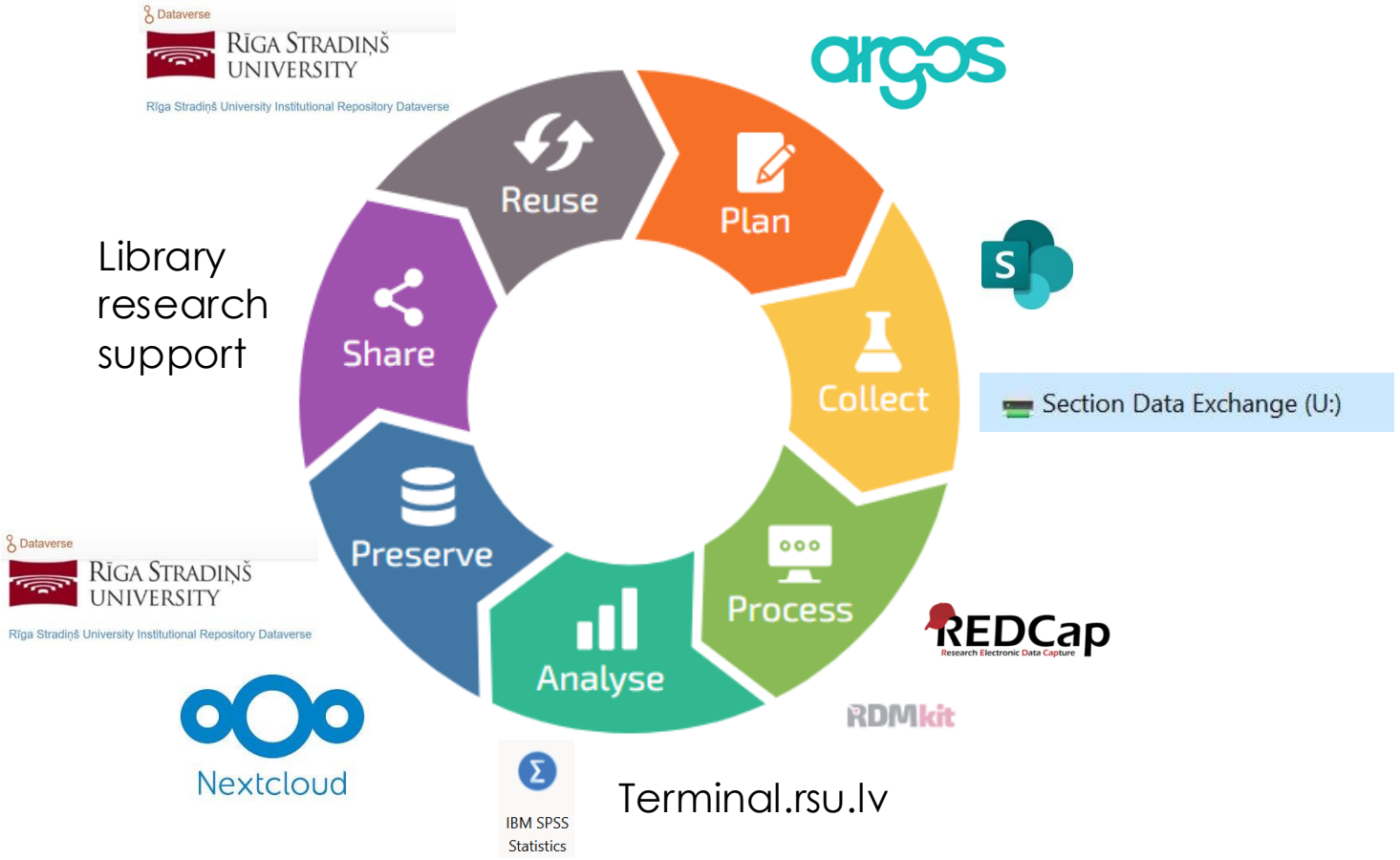
2027 Nacionālais attīstības plāns

Rīgas Stradiņa Universitāte

Consultations for Researchers

From September 1, 2024 – **over 150 consultations**

Support in all data life cycle stages



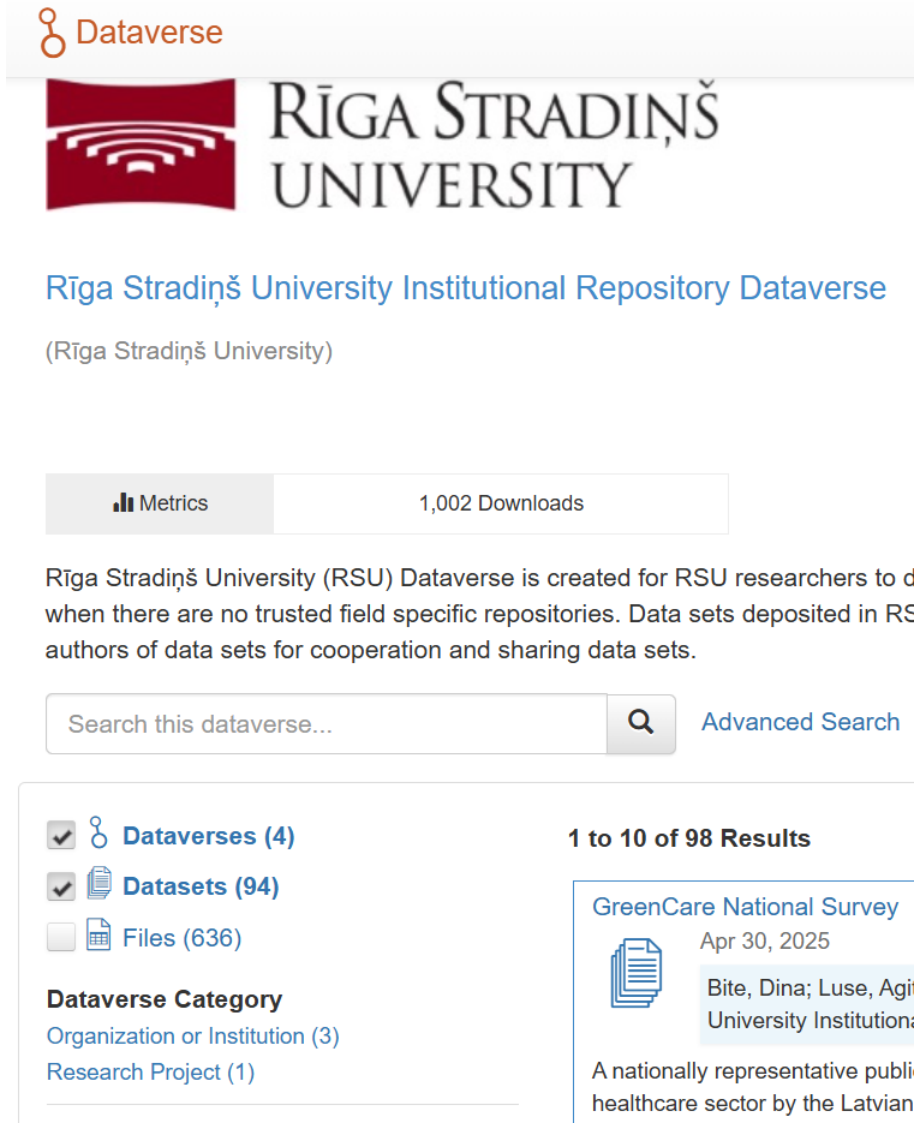
Short-Term Data Storage

During the research, data can be stored:

- On RSU internal server infrastructure;
- In secure cloud services – for example, SharePoint, Nextcloud

RSU Dataverse – Institutional Repository

- Open-source code
- Open metadata standard (DDI)
- Backups
- Automatic licensing (Creative Commons)
- Unique identifier (DOI)
- Open and restricted access
- Automatic versioning



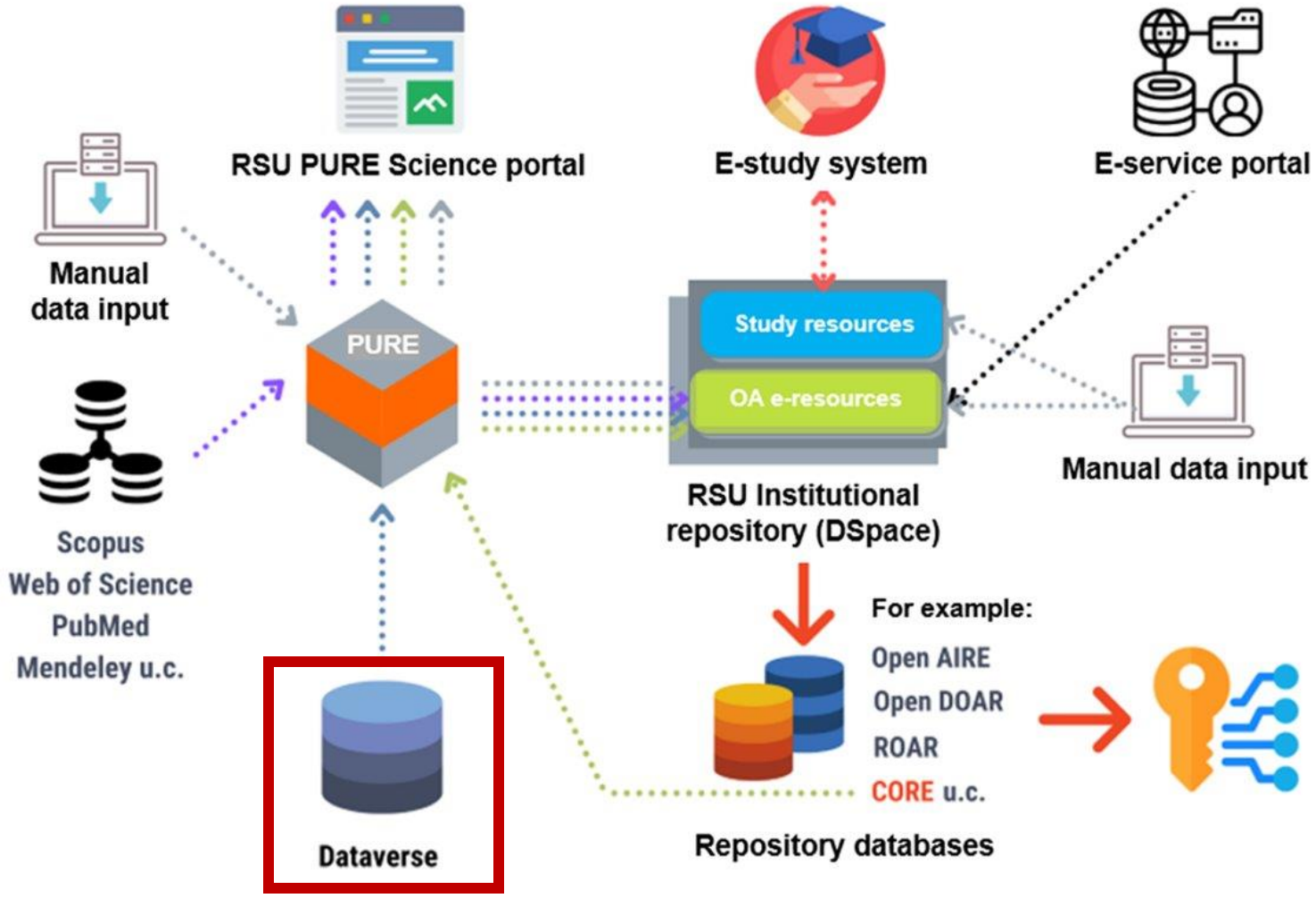
re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

EUROPEAN OPEN SCIENCE CLOUD

DataCite

OpenAIRE

Dataverse in RSU IT Ecosystem



RSU Software and IT Tools Catalog

- Research Preparation Stage
- Data Collection and Collaboration Stage
- Data Analysis and Processing Stage
- Results Interpretation and Visualization
- Publication Preparation and Documentation
- Data and File Storage/Accessibility
- Cloud Solutions
- Project-Acquired Tools
- Supporting Tools
- Repository
- Research Information Management System

Research Information Management System

ZDIS PURE is part of the Latvian Research Information System (ZDIS), based on Elsevier's Pure platform. It provides:

- Centralized storage of research information: publications, projects, conferences, and other research-related metadata.
- Integration with international registries and systems: ORCID, SCOPUS, and others.
- Implementation of open science principles: supporting open repositories and data management plans.
- Ability to track scientific activity and publication impact: citations, metrics, and other statistical indicators.

Seminars for Doctoral School

- Introduction to RDM
- Practical seminar – how to prepare a DMP
- Data documentation
- Short-term and long-term data storage

Soon

Dataset quality assurance. Preparing for depositing

New Instutitional Guidelines and Procedures

- Guidelines for the protection of patients' personal data in research
- Guidelines for the processing and storage of special category data
- Informed Consent Form
- Research Data Management Procedure

DMP template
Metadata questionnaire
Application form for access to research data from RSU Dataverse

Finansē Eiropas Savienība NextGenerationEU

2027 Nacionālais attīstības plāns

Thank you for your attention!

RĪGAS TEHNISKĀ UNIVERSITĀTE

RĪGAS STRADIŅA UNIVERSITĀTE

LATVIJAS UNIVERSITĀTE

Latvijas Biozinātņu un tehnoloģiju universitāte

VPC
AUGSTĀKĀS IZGLĪTĪBAS UN ZINĀTNES IT KOPLIETOŠANAS PAKALPOJUMU CENTRS

Project «Support for implementing Open Science, developing solutions for shared use of research data, and participation in the EOSC» (RRF 2.1.3.1.i.)