Final Deliverable Report (Milestone 3)

FAIR Data Spaces Demonstrator for Cross-Organizational Governance and Usage of Persistent Identifiers in Healthcare & Pharma for faster drug development

Date: 2025-01-21
Contacts: Heiner Oberkampf (CEO), Helge Krieg (Senior Data Engineer)

Overview of Deliverables

Deliverable 1 report:
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/blob/main/doc/FAIR%20DS%20Deliverable%20Report%201%20ACCURIDS.pdf

Deliverable 2 report:
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/blob/main/doc/FAIR%20DS%20Deliverable%20Report%202%20ACCURIDS.pdf

Deliverable 3 report (this document):
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/blob/main/doc/FAIR%20DS%20Deliverable%20Report%203%20ACCURIDS.pdf

Architecture documentation:
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/blob/main/doc/FAIR%20DS%20Architecture%20Document%20ACCURIDS.pdf

Web App:
https://pharma.accurids.com/ and https://hospital.accurids.com/

Docker Image:
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/pkgs/container/accurids-demonstrator

CI/CD:
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/actions

Video:
https://github.com/FAIR-DS4NFDI/accurids-demonstrator/raw/refs/heads/main/doc/accurids-demonstrator.mp4

Contents

Introduction

The FAIR Data Spaces program is exciting and ground-breaking, because it helps to connect data from research and industry according to FAIR data principles to advance innovation and the digital transformation in various sectors. In the demonstrator implemented by ACCURIDS, three different use cases for the data exchange between pharma and research have been investigated and one use case has been implemented as a demonstrator to show how clinical trials can be accelerated with the FAIR Data Spaces.
Next, we provide a summary of the results before we present more detailed descriptions in subsequent sections.

Executive Summary

Documented Pharma Use Cases

As part of the project, we've performed interviews with different people in a selected German pharmaceutical company covering roles in IT/Data, Clinical, Regulatory and Early Research. Based on these interviews three use case descriptions have been created that highlight challenges of cross-organizational collaboration and data exchange between

- Pharma and clinical/hospitals – use case 'Clinical Trials' (Use Case 1))
- Pharma and academic research institutes in biology and genetics in case of the target identification / cell culture research – use case 'Early Research on Cell Cultures' (Use Case 2)
- Pharma, health authorities, pharmacies, hospitals – use case 'Drug Shortage Monitoring' (Use Case 3)

Demonstration

We implemented the scenario of use case 1 about clinical trials with a setup of two ACCURIDS instances with corresponding configuration, customization and user groups.

Pharma Instance

It is available at https://pharma.accurids.com as a demo instance to capture the basic clinical trial information and requirements for collecting human bio sample data in trials performed at clinical research institutions such as a university hospital.

Hospital Instance

It is available at https://hospital.accurids.com as a demo instance to master patient information and collection of bio samples as part of the clinical trial that is performed by a hospital for the sponsoring pharma company.

Practical Demonstration

We demonstrated the scenario live at the final event on December 3rd, 2024. There is a detailed description in section of this document.
We created a short video about the scenario, the video is available at https://github.com/FAIR-DS4NFDI/accurids-demonstrator/raw/refs/heads/main/doc/accurids-demonstrator.mp4.

Technical Implementation

Setup of ACCURIDS and Eclipse Dataspace Components Connector (EDC Connector)

The system architecture of the demonstrator follows best practices of modularization and separation of concerns. We have prepared a detailed diagram that helps to understand the components of the technical implementation. The diagram is included in the architecture document that is available at https://github.com/FAIR-DS4NFDI/accurids-demonstrator/blob/main/doc/FAIR%20DS%20Architecture%20Document%20ACCURIDS.pdf.

The FAIR Data Space Demonstrator consists of the following high-level building blocks:

- Two ACCURIDS platforms were deployed for the FAIR Dataspace demonstrator. They were added as instances to the Accurids cluster in the Accurids cloud.
- We prepared an EDC launcher similar to one of the examples of the Eclipse Dataspace components (https://github.com/eclipse-edc/Connector/tree/main/launchers/sts-server).
- We implemented a Python script that manages the communication with the International Dataspace (IDS) protocol.

Detailed Description of the Deliverables

Documented Pharma Use Cases

We have performed interviews with different people in a German pharmaceutical company covering roles in IT/Data, Clinical, Regulatory and Early Research. Based on these interviews, three use case descriptions have been created that highlight the challenges of cross-organizational collaboration and data exchange between a pharma company and (Use Case 1) clinical/hospital (Use Case 2) university biotech research and (Use Case 3) health authorities. The three use cases below are documented as part of the presentation of the final event on December, 3rd 2024. The implementation of Use case 1 was presented as a live demonstration during the final event.

Use Case 1: Clinical Trials – Connecting Pharma with a University Hospital

Challenge

The execution of a clinical trial is a race against time to bring novel medicines to the market. Over three phases, a pharmaceutical company must scientifically prove the efficacy of a new treatment for a target patient group. This process generates vast amounts of data, collected by diverse research organizations, including university hospitals. Currently, this data is often submitted to pharma companies in the form of numerous Excel sheets through contract research organizations (CROs), lacking standardized formats, making integration and reliable identification and traceability of each data point almost impossible.

Solution Outline

We implemented a connector for the FAIR Data Spaces with ACCURIDS to combine a trusted infrastructure with our existing Data Registry software that is already used within big pharma. Our solution offers the following benefits:

1. *ACCURIDS Registry Instance:* Each participant operates an ACCURIDS registry, linked to the FAIR Data Spaces infrastructure for standardized, interoperable data sharing.
2. *Unique Data Identifiers:* Every data point has a globally unique, persistent identifier, ensuring traceability and integration.
3. *Standardized Metadata Sharing:* Metadata follows industry standards, ensuring consistent and reliable data exchange.
4. *Controlled Data Granularity:* Data originators control how much detail is shared, ensuring security and customization.
5. *Seamless Data Integration:* Pharma companies integrate external and internal data automatically, reducing manual work.
6. *Data Traceability and Quality:* Full traceability and automatic quality checks maintain data integrity and reliability.
7. *Real-Time Data Analysis:* Enables immediate analysis, helping pharma companies adjust trials faster and accelerate drug delivery.

Result: The Demonstrator highlights the feasibility of faster, cheaper and better execution of clinical trials. We can bring novel medicines faster to the market.

Use Case 2: Early Research on Cell Cultures – Connecting Pharma with University Research in Biotech.

Challenge

Early-stage research for novel medicines relies on *in vitro* testing of RNA constructs to identify viable candidates. However, collaboration between academia and industry faces the following challenges limiting collaboration efficiency and speed:

- *Increased number of different Cell Cultures:* Advances in biotechnology have led to an explosion in cell culture variants and modifications.
- *Diverse Analytical Techniques:* Increasingly complex methods make it more difficult to interpret results effectively.
- *Data Exchange Issues:* Current practices use Excel sheets and ad hoc numbering or naming, creating ambiguities in identifying cell cultures, RNA constructs, and plasmids.
- *Knowledge Silos:* Reliance on individual expertise, e.g., a PhD student conducting the joint research project, risks losing critical insights during transitions.
- *Plasmid Complexity:* Plasmid designs involve intricate metadata, yet current systems poorly handle this complexity.

Solution Outline

We ensure the unique identification of key research objects and standardize their metadata. The solution offers the following benefits:

1. *Standardized Taxonomy:* Classify key research objects like cell cultures (e.g., by lineage, phenotype) and plasmids (e.g., by design and function) for consistency across institutions.
2. *Minimal Metadata Standards:* Define essential attributes for cell cultures (e.g., genetic modifications, growth conditions) and plasmids (e.g., sequence details, modification history) to ensure traceability.
3. *Unique Identification:* Use globally unique IDs based on standardized metadata agreed normalization functions and public references (e.g., GenBank, UniProt) for unambiguous tracking of cell cultures, RNA constructs, and plasmids.
4. *Seamless Metadata Integration:* Enable structured metadata sharing to automate integration into pharma systems, reducing manual work and ensuring data consistency.
5. *Real-Time Analysis and Governance:* Leverage integrated data for immediate analysis to refine research focus, accelerate RNA construct validation, and establish governance for scalable, standardized collaborations.

Target Result: This solution enhances collaboration, accelerates research timelines, and improves reproducibility for faster drug discovery.

Use Case 3: Drug Shortage Monitoring – Connecting Pharma with Health Authorities.

Challenge

Due to the global disruption of supply chains, drug shortages have become a problem also for countries of the European Union, even for critical medicines.
To avoid drug shortages, health authorities such as the European Medicines Agency (EMA) or the German Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) collect information about the inventory of available medicinal products from pharmaceutical companies, hospitals and pharmacies.
The problem however is, that the information across the different organizations and European countries is not standardized which prevents health authorities from aggregating information effectively. For example, the same product is marketed under different brands by different companies in different pack sizes with different strengths in different countries.
In addition, critical information about manufacturing locations, manufacturers, and input materials is not standardized or tracked, impeding proactive shortage mitigation efforts.

Solution Outline

The FAIR Data Spaces and ACCURIDS Registry with data standardization pipelines for medicinal product information provide a trusted and standardized infrastructure for harmonizing and sharing medicinal product data across stakeholders in the pharmaceutical supply chain. The solution offers the following benefits:

1. *Harmonization of Master Data:* Establish a unified framework to standardize medicinal product data based on ISO Identification of Medicinal Products (IDMP) and the IDMP Ontology, including attributes for drug name, formulation, manufacturing locations, and materials.
2. *Trusted Data Exchange:* Implementation of HL7 Fast Healthcare Interoperability Resources (FHIR) using the FAIR Data spaces infrastructure in combination with FHIR messaging hubs for data exchange between participants using agreed upon reference and master data for contextualizing information in a standardized manner.

With this basis in place, health authorities can setup an inventory monitoring system based on standardized processes for real-time inventory tracking of critical medicines, leveraging FAIR-compliant infrastructure to detect and mitigate shortages proactively.

Result: Drug shortage monitoring and drug shortage prevention can be done effectively based on standardized medicinal product data, and incrementally along multiple dimensions:

(a) *Scope of medicines:* We start with critical ones and expand.
(b) *Jurisdictions and countries*: We focus on a first country and afterwards repeat for further ones.
(c) *Scope of shared information*: We start with the inventory of packaged medicinal products and expand to manufacturing and supply chain information.

Demonstrator of Use Case 1: Clinical Trails

The demonstrator implements healthcare and pharma use case 1. We show how a pharma company and a hospital can exchange data about human bio samples in the context of a clinical trial sponsored by a pharma company in an automated manner allowing to speed up clinical trial data analysis and decision making. It uses the implemented EDC connector for ACCURIDS to realize a trusted data exchange and sharing of data collection requirements with a setup of two ACCURIDS instances with corresponding configuration, customization and user groups.

Pharma Instance

It us available at https://pharma.accurids.com as a demo instance to capture the basic clinical trial information and requirements for collecting human bio sample data in trials performed at clinical research institutions such as a university hospital.

Hospital Instance

It is available a https://hospital.accurids.com as a demo instance to master patient information and collection of bio samples as part of the clinical trial perform for the sponsoring pharma company.

Demonstration

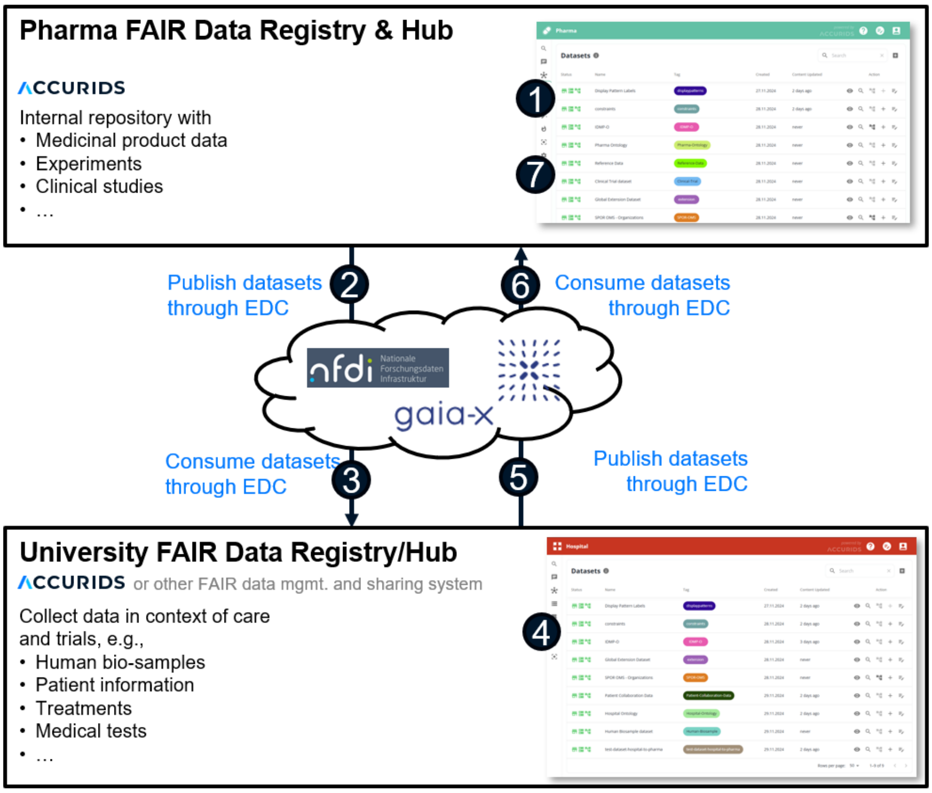

Pharma FAIR Data Registry & Hub
ACCURIDS
Internal repository with
• Medicinal product data
• Experiments
• Clinical studies
• …
Pharma
Datasets
1
7
Publish datasets through EDC
2
6
Consume datasets through EDC
nfdi Nationale Forschungsdaten Infrastruktur
gaia-x
Consume datasets through EDC
3
5
Publish datasets through EDC
University FAIR Data Registry/Hub
ACCURIDS or other FAIR data mgmt. and sharing system
Collect data in context of care and trials, e.g.,
• Human bio-samples
• Patient information
• Treatments
• Medical tests
• …
Hospital
Datasets
4

Figure 1: Demonstrator setup for Use Case 1 with Pharma and University instances of ACCURIS connected to the FAIR Data Spaces through the EDC connectors

The diagram above provides an overview of the high-level setup of the demonstrator including the following demonstration steps:

1. Manage FAIR clinical trial data in pharma instance
2. Provide (meta-)data standards through EDC
3. Consume (meta-)data standards through EDC
4. Capture trial related bio samples in hospital instance
5. Provide bio samples datasets through EDC
6. Consume bio samples datasets
7. Analyze clinical trial data

The Accurids Data Registration Software was used as a distributed governance platform to register and resolve data objects managed in different organizations (pharma, university hospitals) within a secure environment. The publication of semantic data schema for different type of objects, e.g., clinical trial or human bio samples with corresponding lab tests are shared between pharma and hospital instances together with corresponding reference vocabularies.

The demonstrator implements the FAIR Data Spaces system and architecture through an open-source EDC connector. This approach ensures interoperability with other FAIR Data spaces demonstrators.

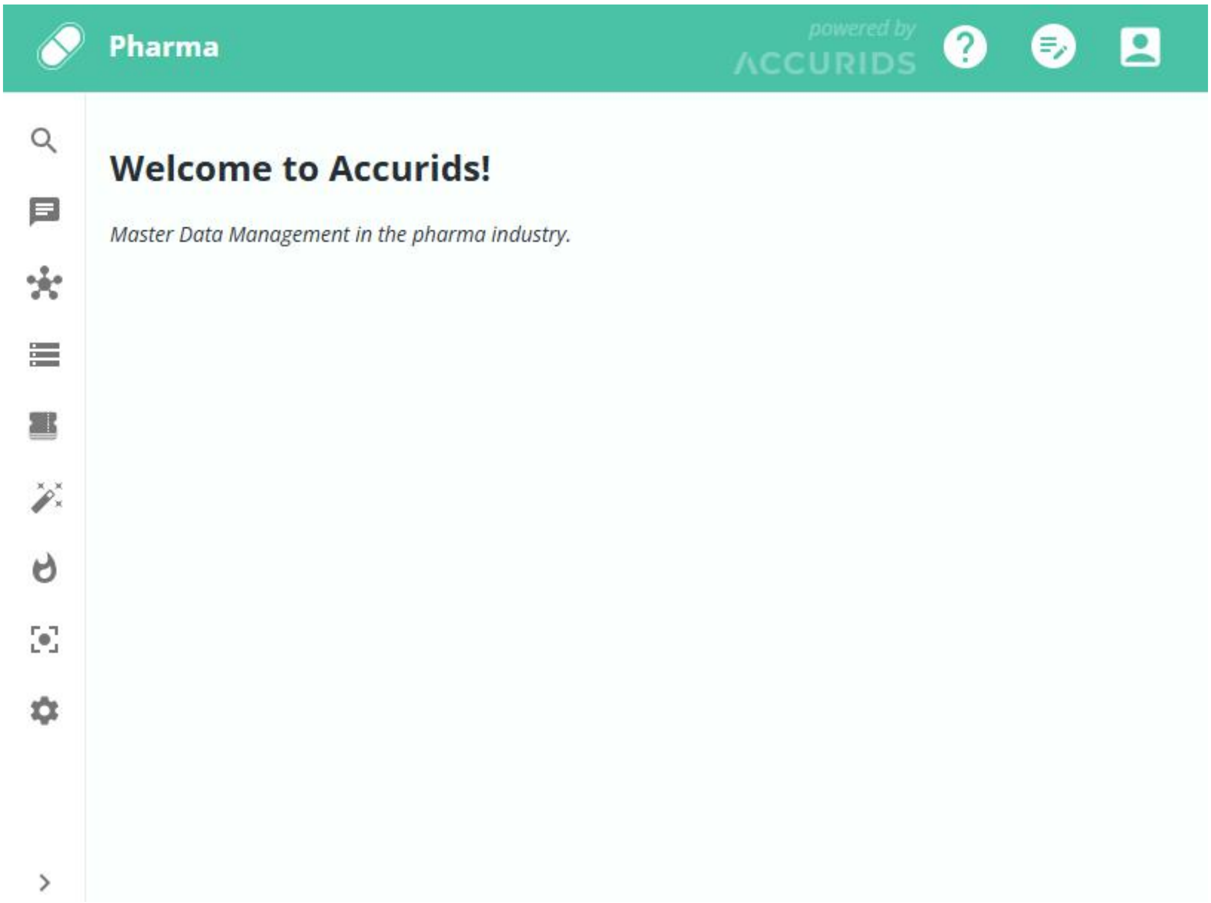
The demonstrator may be expanded in the future to create synergies with nfdi4health use cases and demonstrators, such as the Personal Health Train that already used ACCURIDS software.

ACCURIDS Demonstrator Setup

Pharma Instance

We have one Accurids instance that represents the Pharma side of the demonstrator.

Landing page

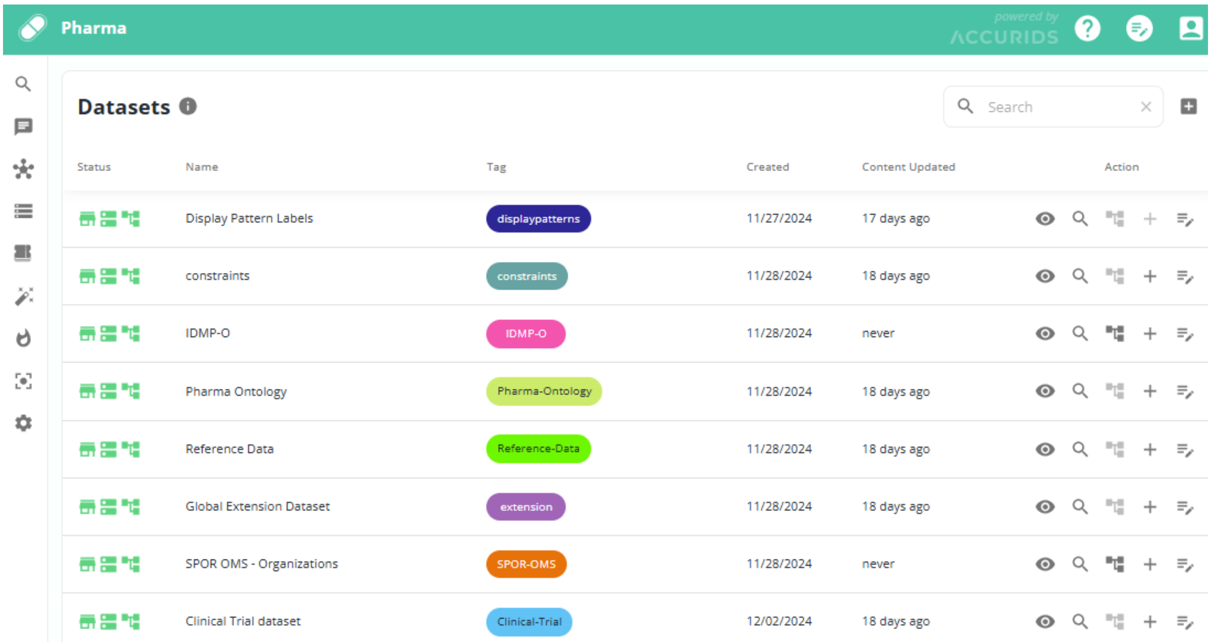


Welcome to Accurids!

Master Data Management in the pharma industry.

The landing page of the Pharma instance offers the standard layout and buttons as included in the Accurids platform out of the box.
We have added custom styles and logo to highlight the Pharma side.

Datasets



Datasets

Status	Name	Tag	Created	Content Updated	Action
	Display Pattern Labels	displaypatterns	11/27/2024	17 days ago	
	constraints	constraints	11/28/2024	18 days ago	
	IDMP-O	IDMP-O	11/28/2024	never	
	Pharma Ontology	Pharma-Ontology	11/28/2024	18 days ago	
	Reference Data	Reference-Data	11/28/2024	18 days ago	
	Global Extension Dataset	extension	11/28/2024	18 days ago	
	SPOR OMS - Organizations	SPOR-OMS	11/28/2024	never	
	Clinical Trial dataset	Clinical-Trial	12/02/2024	18 days ago	

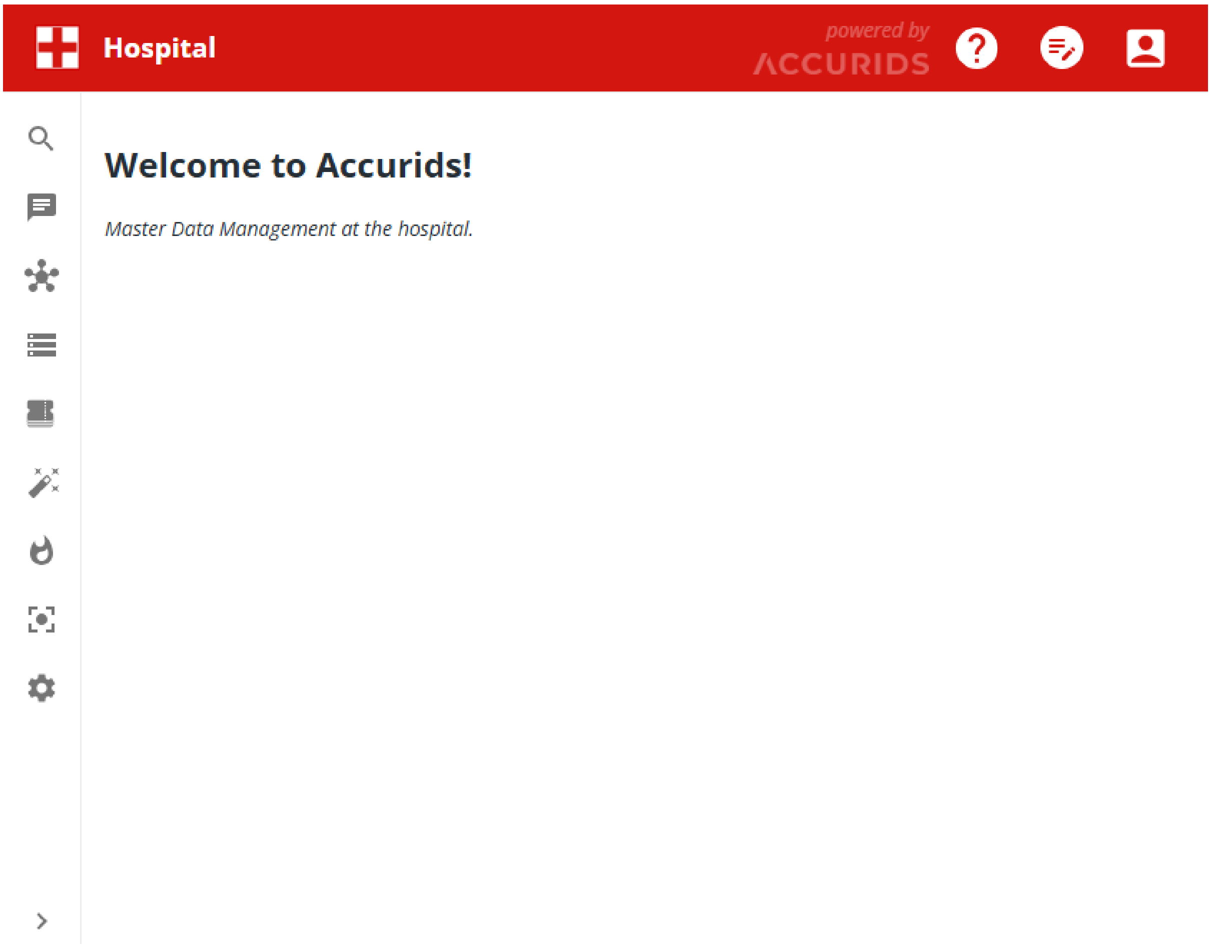
We have prepared the following datasets for demonstrating the use case:

- Pharma Ontology
  - It is the vocabulary of the Pharma company to describe information about clinical trials and human bio samples.
  - The Ontology contains terms like idmp-mprd:ClinicalTrial, pharma:TrialPhase, pharma:TrialSite, pharma:HumanBiosample.
    “pharma:” is the namespace of the Pharma company. The vocabulary can be aligned with public terminology like the IDMP-O (https://spec.edmcouncil.org/idmp/ontology). For example, the namespace “idmp-mprd:” belongs to the IDMP ontology.
- Reference Data
  - It is the standardized reference data from the Pharma company.
  - We include information like standard types of bio samples (e.g. cells, tissue, blood), typical phase of clinical trials (e.g. preclinical phase, phase 1, 2, 3, etc.)
- Public terminology
  - Interoperability and reusability are two of the core principles of FAIR data. Consequently, we reuse public terminology like IDMP-O, or standard SPOR datasets.
- Clinical Trial dataset
  - This dataset contains the Pharma information about the clinical trials. Our goal is to connect the clinical trial data with the information content coming from hospitals taking part in the clinical trial.
- Constraints
  - The dataset contains the specification of the information model of the Pharma company represented by data shapes following the SHACL specification.

Hospital Instance

We have one Accurids instance that represents the Hospital side of the demonstrator.

Landing page

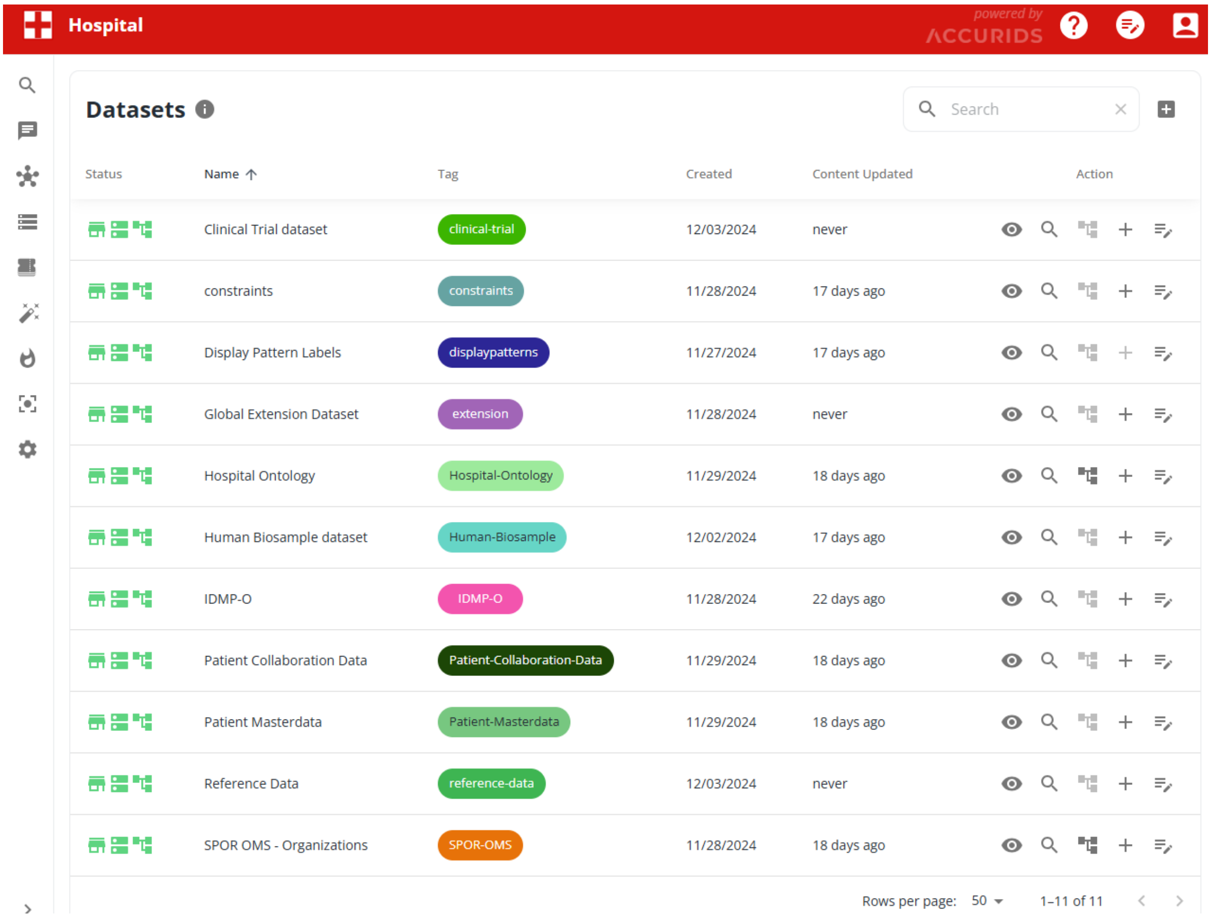


Welcome to Accurids!

Master Data Management at the hospital.

The landing page of the Hospital instance offers the standard layout and buttons as included in the Accurids platform out of the box. We have added custom styles and logo to highlight the Hospital side.

Datasets



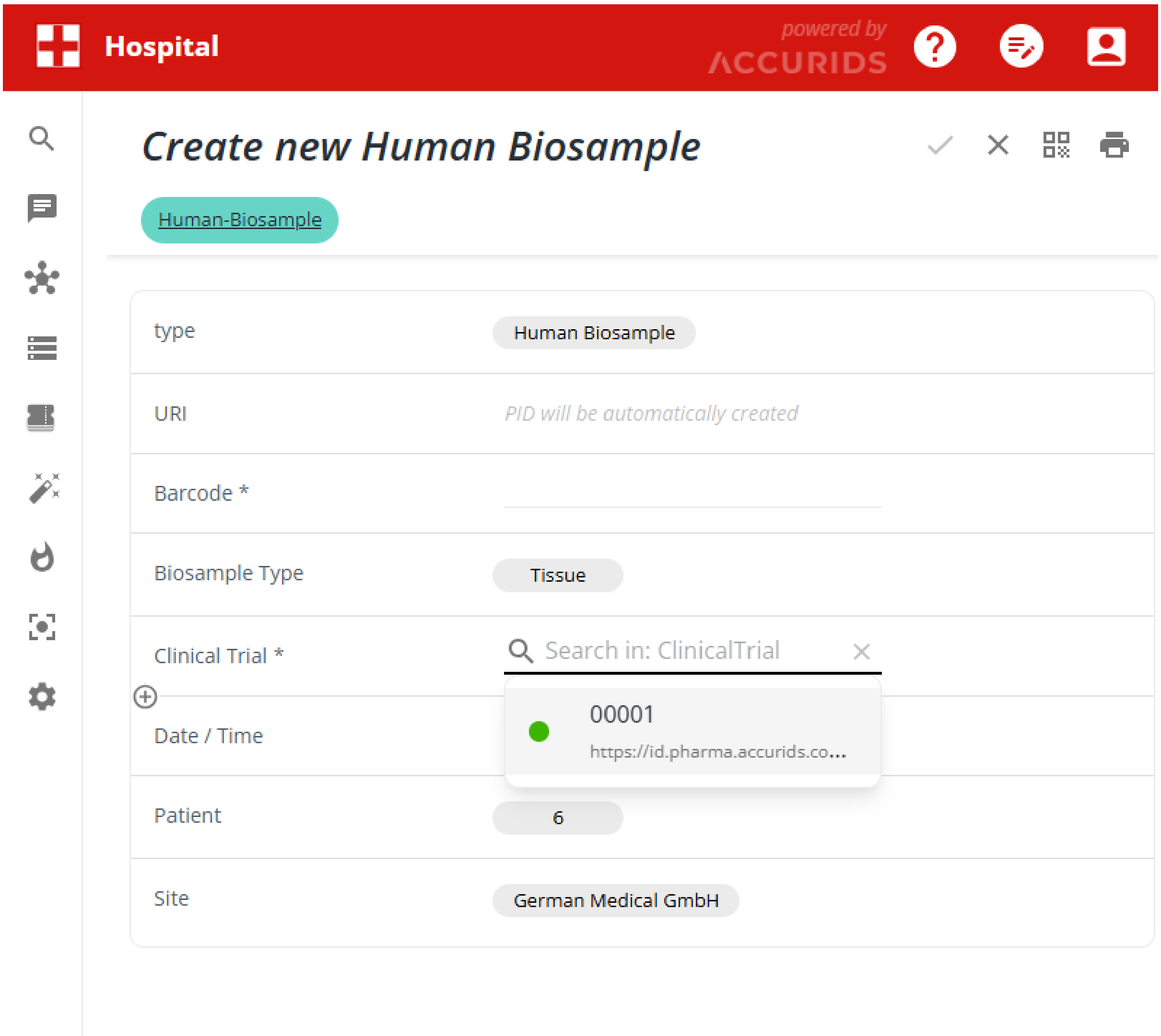
Datasets

Status	Name	Tag	Created	Content Updated	Action
	Clinical Trial dataset	clinical-trial	12/03/2024	never	
	constraints	constraints	11/28/2024	17 days ago	
	Display Pattern Labels	displaypatterns	11/27/2024	17 days ago	
	Global Extension Dataset	extension	11/28/2024	never	
	Hospital Ontology	Hospital-Ontology	11/29/2024	18 days ago	
	Human Biosample dataset	Human-Biosample	12/02/2024	17 days ago	
	IDMP-O	IDMP-O	11/28/2024	22 days ago	
	Patient Collaboration Data	Patient-Collaboration-Data	11/29/2024	18 days ago	
	Patient Masterdata	Patient-Masterdata	11/29/2024	18 days ago	
	Reference Data	reference-data	12/03/2024	never	
	SPOR OMS - Organizations	SPOR-OMS	11/28/2024	18 days ago	

Rows per page: 50 1–11 of 11

- Hospital Ontology
    - It is the vocabulary of the Hospital to describe information about patients and human bio samples.
- Patient Master Data
    - It is the personal information of patients in the hospital. Only anonymized information can be exchanged with external parties.
- Patient Collaboration Data
    - It is anonymized patient data that can be exchanged with external parties like the Pharma company.
- Constraints
    - At the beginning of the demonstration, this dataset is empty. Once we start the synchronization between the Accurids platforms from Pharma and Hospital, the Hospital receives the information model from the Pharma company as SHACL shapes.

Dynamic User Interface



Hospital

powered by ACCURIDS

Create new Human Biosample

Human-Biosample

type	Human Biosample
URI	*PID will be automatically created*
Barcode *	
Biosample Type	Tissue
Clinical Trial *	Search in: ClinicalTrial ● 00001 https://id.pharma.accurids.co...
Date / Time	
Patient	6
Site	German Medical GmbH

The User Interface to enter information about human bio samples is generated dynamically in the Hospital instance from the information model of the Pharma company.
We send the information model as SHACL shapes from the Pharma instance to the Hospital instance through the Eclipse Dataspace Connector.
In the same way, we send the Clinical Trial information from the Pharma instance to the Hospital instance.
As a result, the laboratory person doing the data entry can provide the data in the required structure of the Pharma company, and can also connect the information to the related clinical trial (e.g., “00001”, see screenshot).
Once the data has been entered, our synchronization makes sure that the information becomes immediately available in the Pharma instance.

Technical Details

Overview

Two ACCURIDS platforms were deployed for the FAIR Dataspace demonstrator. They were added as instances to the Accurids cluster in the Accurids cloud. Each instance is a Java Spring Boot application that is running inside of a Docker container. We use Postgres as the storage backend, the database is available as part of the same cluster. In addition to the relational database, we apply an Elasticsearch database to provide a search index with high performance.

We prepared an EDC launcher similar to one of the examples of the Eclipse Dataspace components (https://github.com/eclipse-edc/Connector/tree/main/launchers/sts-server). The launcher is executed as two separate services: a consumer and a provider. For the demonstrator, both services were executed on the same Windows machine.

We implemented a Python script that manages the communication with the International Dataspace (IDS) protocol. The script has basically two steps of execution. First, we setup the FAIR Dataspace i.e., we create data assets, and add the required roles and policies for the data exchange. Second, we run a continuous job, that looks for updates in datasets of the two Accurids platforms and synchronizes the data content if needed. The data exchange happens entirely through the HTTP data plane of the underlying Eclipse Connector. We use REST calls and GraphQL queries to execute data updates between the Accurids instances.

The detailed architecture is documented in a separate document that is available at https://github.com/FAIR-DS4NFDI/accurids-demonstrator/blob/main/doc/FAIR%20DS%20Architecture%20Document%20ACCURIDS.pdf.

Technical Challenges

We faced a couple of technical challenges during implementation of the demonstrator setup.

Challenging Documentation

There is documentation about the Eclipse Connector. The challenge was to find the proper documentation and to decide if it is applicable to the current state of development of the EDC Connector. There is an MVP example and some Samples like a technical tutorial. Both help to get the basics, but the knowledge barrier for a developer to reuse the Connector, or to extend it, is still very high. There seems to be the perspective of the Eclipse Connector team that it would not make sense to provide concrete examples of implementation because everyone would have a different use case, so such an example would not help.
From our perspective, such a working example would help developers to see the technical requirements and understand the principles. Such an example could be like a template for reuse. It would have saved us 10 days of effort.

Modularized source code

The Eclipse Connector source code follows best practices of software development such as modularization, clean code, abstractions, applying software design patterns. This helps for software development in an efficient way and allows developers to better maintain and reuse code fragments. On the other hand, such a code repository can require a lot of training for new software developers. They need to understand all the patterns and abstractions before they can start working on their actual goal. The fact that the Eclipse Connector only consists of modules requires new developers to prepare a runnable piece of code before they can actually apply the Connector to their use case. Another option for new developers would be to look through unit tests. Executing and debugging unit tests can help to understand basic building blocks, yet the approach can be cumbersome and time consuming.

EDC launcher modules

This topic is closely related to the previous one. How can a new developer create a launcher without knowing the required Connector modules? In our case, the task required hours of reading through source code, examples, issue tickets, and a lot of trial and error. It would have saved us at least 1 day of effort if there were provided a selection of launchers for a couple of the common use cases.

Communication with HTTP

We reused the HTTP data plane of the Eclipse Connector for the data transfer. It turned out that the underlying library (okhttp3), or the way how it is used in the Eclipse

Connector results in HTTP multipart messages that deviate from the specification. At least in our applications and tests, we found issues with system separators and line breaks that resulted in communication errors. Our workaround is to wrap the communication with okhttp3 with a layer of Spring Webflux. As soon as we used the Webflux library, the data exchange of multipart messages worked as expected. We suggest to do thorough testing of the HTTP data plane of the Eclipse Connector for different kinds of small and large messages, various kinds of HTTP methods, authentication methods, and including single as well as multipart messages.

IDS protocol

The Eclipse Connector makes use of the IDS protocol. It is a challenge for a new developer to understand the schema of IDS messages as well as the overarching ideas and specifications. On the one hand, there a very abstract specification documents about many (theoretical) aspects of data spaces, on the other hand there is a concrete implementation that makes use of parts of the specification. Both sides have different versions, and evolve in parallel. How can a developer easily find out what the currently required technical protocol looks like? In our case, we followed several Sample implementations, tutorials and other documentation in order to extract the technical workflow of message exchange that is required today. A concrete example would have saved us at least 5 days of effort.

Front-End & Back-End

Back-End Technologies

- Programming Languages: We utilize Java and Python. Java for building RESTful APIs aligning with W3C Recommendations and, and Python for the data synchronization service following the Gaia-X Trust Framework.
- API Development: Our APIs have been designed using RESTful principles and GraphQL to maximize efficiency and flexibility in data handling, aligning with W3C Recommendations.
- Data Handling and Persistence: We utilize PostgreSQL for complex transactions, ElasticSearch for fast indexing and searching, Apache TDB as triple store for persistence.
- Semantic Technologies: We implement the master data management platform with Apache Jena, OWL, and SPARQL to manage RDF data models, enhancing data interoperability within the FAIR Data Spaces.

Front-End Technologies

- We follow a minimalist approach, focusing on functionality for system monitoring and configuration management.

- React is used for building user interfaces, and providing efficient, dynamic interaction capabilities.

Programming and Development Standards

- Enforcement Tools: Automated tools such as Prettier for consistent code formatting and SonarQube for static code analysis are integrated into our GitLab pipeline. This ensures adherence to coding standards and help identify security vulnerabilities and code smells.
- Code Reviews: All code changes undergo thorough peer reviews in GitLab Merge Requests, ensuring adherence to our coding standards and overall quality.
- System Change Control Procedures: Documented change control procedures are in place to manage system modifications, ensuring all changes meet stringent standards for security and efficiency.
- Continuous Testing: Integration of continuous security and acceptance testing within our development lifecycle guarantees that no code is deployed to production without passing rigorous tests.

Software Development Process: CI/CD, Documentation, Testing, Security

- Continuous Integration/Continuous Deployment: We have CI/CD pipelines for the ACCURIDS product using GitLab CI, including linting and automated deployment in Kubernetes. This setup ensures automated testing and deployment processes that align with open-source community standards. We apply GitHub CI/CD to the building and deployment of a Docker container that runs the Demonstrator Manager Service as well as the consumer and provider instances of the EDS Connector.
- Documentation: We maintain a comprehensive documentation that is available online at https://docs.accurids.com/current/. The documentation includes detailed API documentation created with Swagger.
- Testing: We employ a robust testing strategy that includes automated unit tests, and integration tests.
- Code Analysis
  - SAST (Static Application Security Testing): We use SonarQube for static code analysis to detect security vulnerabilities and enforce coding standards. This tool helps to ensure that all developed software adheres to the highest security standards before deployment.
  - SCA (Software Composition Analysis): Our development process will include the use of tools like OWASP Dependency-Check to perform SCA, assessing our codebase for known vulnerabilities, and checking for available updates. This ensures that our software components are secure and up-to-date with the latest security patches.

Development Environment and Tools

- Common Development Environment: We authorize the use of IDEs such as IntelliJ IDEA, Visual Studio Code, and Eclipse to standardize our development environment.
- Source Code Reviews: We have established requirements for source code reviews to ensure high standards of code quality. Reviews are conducted before any code is merged into our main codebase, supported by automated tools and peer assessments within our GitLab Merge Requests.

Security

- Encryption and Data Security: We make use of SSL for robust encryption in transit and at rest, safeguarding data integrity and confidentiality.
- Access and Identity Management: We utilize open-source solutions and integrate OAuth2 for robust user authentication and authorization, aligning with common web standards.
- Compliance and Privacy: The connector adheres to open data compliance standards like GDPR, leveraging community-driven solutions for data governance and security. This ensures that all processes and technologies meet the necessary regulatory requirements.

Further contributions

We can report the following additional outcomes of the project:

- NDFI podcast with Heiner Oberkampf elaborating the use cases for clinical trials and the need for more automated data exchange in healthcare.
- Contribution to the research paper "From Theory to Practice: Demonstrators of FAIR Data Spaces Across Different Sectors"
- Inclusion of the FAIR Data Spaces topic in the ACCURIDS WHODrug Knowledge Graph hackathon to explore pharma and healthcare use cases in the context of global medicinal product standardization with the Uppsala Monitoring Center, pharma and hospitals.