ACCURIDS

FAIR Data Spaces Demonstrator: Persistent IDs in Pharma & Healthcare

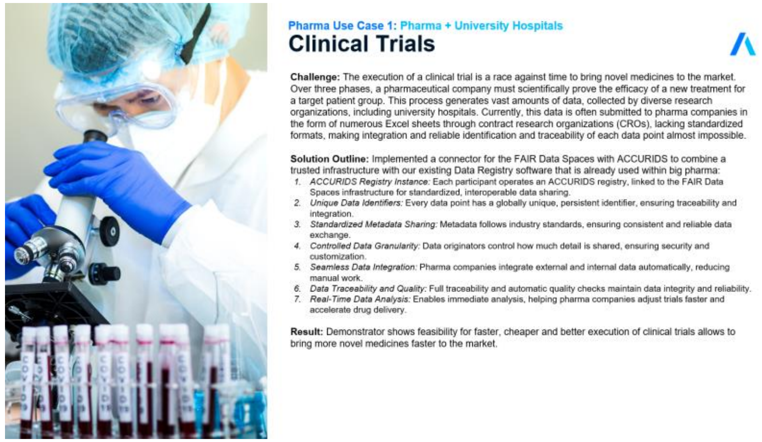
Connecting pharma with clinical healthcare and academic biochemical research to speed-up drug development

Heiner Oberkampf, Helge Krieg, Oleksandr Snizhko

2024-12-03 Final Presentation & Demo

Pharma Use Cases

1 Pharma + University Hospitals



Pharma Use Case 1: Pharma + University Hospitals

Clinical Trials

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.

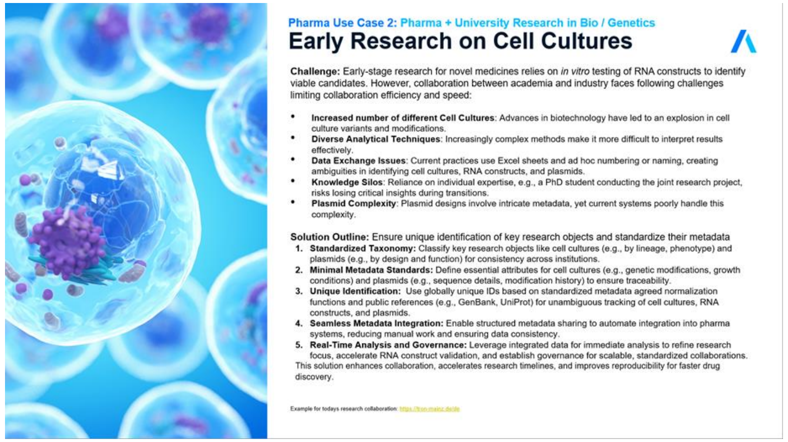
Solution Outline: Implemented a connector for the FAIR Data Spaces with ACCURIDS to combine a trusted infrastructure with our existing Data Registry software that is already used within big pharma:

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: Demonstrator shows feasibility for faster, cheaper and better execution of clinical trials allows to bring more novel medicines faster to the market.

→ ***Live demo today***

2 Pharma + University Bio Research



Pharma Use Case 2: Pharma + University Research in Bio / Genetics

Early Research on Cell Cultures

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 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: Ensure unique identification of key research objects and standardize their metadata

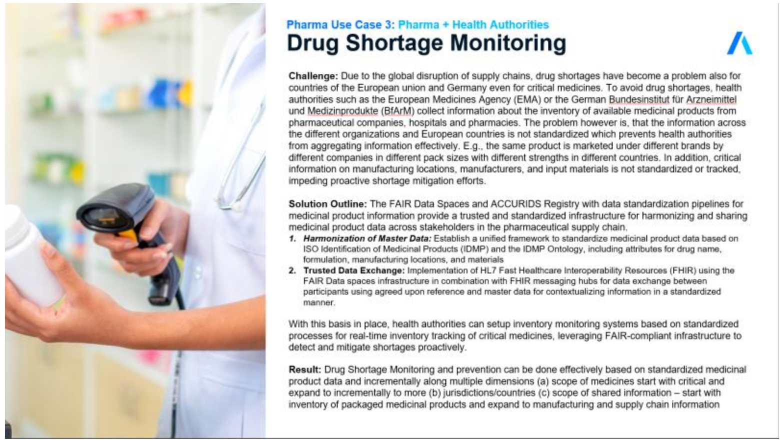
1. **Standardized Taxonomy:** Classify key research objects like cell cultures (e.g., by lineage, phenotype) and plasmids (e.g., by design and function) for consistency across institutions.
2. **Minimal Metadata Standards:** Define essential attributes for cell cultures (e.g., genetic modifications, 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.

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

Example for todays research collaboration: https://tron-mainz.de/de

→ *Small PoC screen recording*

3 Pharma + Health Authorities



Pharma Use Case 3: Pharma + Health Authorities

Drug Shortage Monitoring

Challenge: Due to the global disruption of supply chains, drug shortages have become a problem also for countries of the European union and Germany 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. E.g., 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 on 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.

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 inventory monitoring systems 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 prevention can be done effectively based on standardized medicinal product data and incrementally along multiple dimensions (a) scope of medicines start with critical and expand to incrementally to more (b) jurisdictions/countries (c) scope of shared information – start with inventory of packaged medicinal products and expand to manufacturing and supply chain information

→ *Implemented separately*

Pharma Use Case 1: Pharma + University Hospitals

Clinical Trials

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: 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:

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: Demonstrator shows feasibility for faster, cheaper and better execution of clinical trials allows to bring more novel medicines faster to the market.

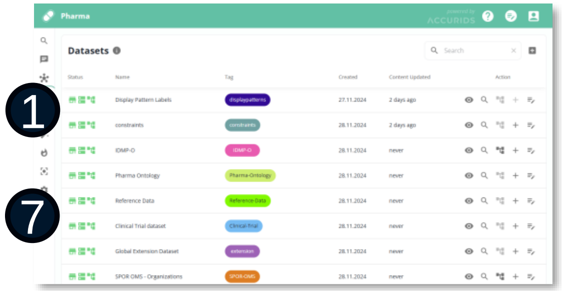
Demonstrator Setup for Use Case 1

Clinical Trials – connecting clinical trial sites (e.g., university hospitals) with pharmaceutical companies

Pharma FAIR Data Registry & Hub

ACCURIDS

Internal repository with
- Medicinal product data
- Experiments
- Clinical studies
- …



Datasets

Status	Name	Tag	Created	Content Updated	Action
	Display Pattern Labels	displaypatterns	27.11.2024	2 days ago	
	constraints	constraints	28.11.2024	2 days ago	
	IDMP-O	IDMP-O	28.11.2024	never	
	Pharma Ontology	Pharma-Ontology	28.11.2024	never	
	Reference Data	Reference-Data	28.11.2024	never	
	Clinical Trial dataset	Clinical-Trial	28.11.2024	never	
	Global Extension Dataset	extension	28.11.2024	never	
	SPOR OMS - Organizations	SPOR-OMS	28.11.2024	never	

2 Publish datasets through EDC

6 Consume datasets through EDC

nfdi
Nationale Forschungsdaten Infrastruktur
gaia-x

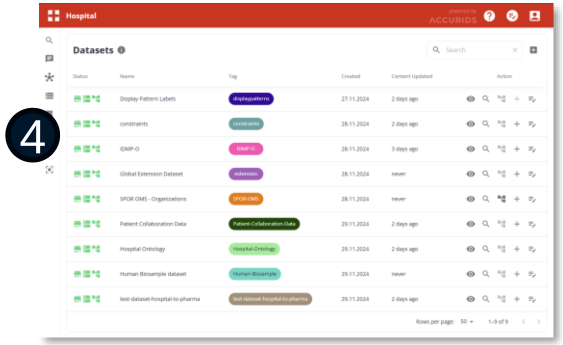
3 Consume datasets through EDC

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
- …



Datasets

Status	Name	Tag	Created	Content Updated	Action
	Display Pattern Labels	displaypatterns	27.11.2024	2 days ago	
	constraints	constraints	28.11.2024	2 days ago	
	IDMP-O	IDMP-O	28.11.2024	3 days ago	
	Global Extension Dataset	extension	28.11.2024	never	
	SPOR OMS - Organizations	SPOR-OMS	28.11.2024	never	
	Patient Collaboration Data	Patient-Collaboration-Data	29.11.2024	2 days ago	
	Hospital Ontology	Hospital-Ontology	29.11.2024	2 days ago	
	Human Biosample dataset	Human-Biosample	29.11.2024	never	
	test-dataset-hospital-to-pharma	test-dataset-hospital-to-pharma	29.11.2024	2 days ago	

Rows per page: 50 1–9 of 9

The demonstrator implements healthcare and pharma use cases showcasing how, a pharma company and a hospital can exchange data about human bio samples in context of a clinical trial sponsored by the 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.

Demo Steps:

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

Demo

EDC implementation

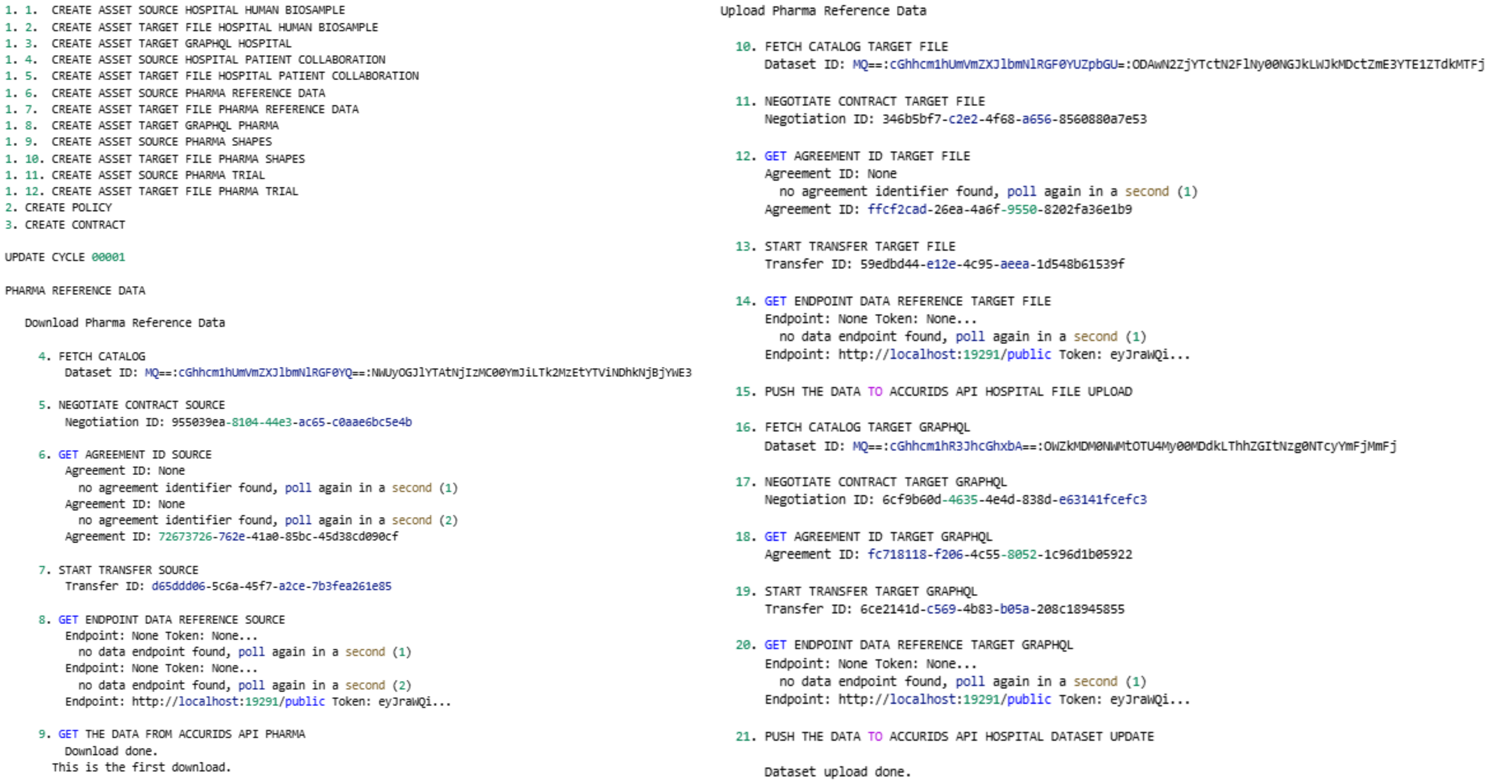
As part of the demonstrator implementation EDC connectors can communicate with ACCURIDS instances

Configured/Implemented 2 connectors

1. Data Provider EDC Connector
2. Data Consumer EDC Connector

Data Sharing through FAIR Data Spaces EDC connectors

pharma → hospital → pharma



```
1. 1.  CREATE ASSET SOURCE HOSPITAL HUMAN BIOSAMPLE
1. 2.  CREATE ASSET TARGET FILE HOSPITAL HUMAN BIOSAMPLE
1. 3.  CREATE ASSET TARGET GRAPHQL HOSPITAL
1. 4.  CREATE ASSET SOURCE HOSPITAL PATIENT COLLABORATION
1. 5.  CREATE ASSET TARGET FILE HOSPITAL PATIENT COLLABORATION
1. 6.  CREATE ASSET SOURCE PHARMA REFERENCE DATA
1. 7.  CREATE ASSET TARGET FILE PHARMA REFERENCE DATA
1. 8.  CREATE ASSET TARGET GRAPHQL PHARMA
1. 9.  CREATE ASSET SOURCE PHARMA SHAPES
1. 10. CREATE ASSET TARGET FILE PHARMA SHAPES
1. 11. CREATE ASSET SOURCE PHARMA TRIAL
1. 12. CREATE ASSET TARGET FILE PHARMA TRIAL
2. CREATE POLICY
3. CREATE CONTRACT

UPDATE CYCLE 00001

PHARMA REFERENCE DATA

   Download Pharma Reference Data

     4. FETCH CATALOG
        Dataset ID: MQ==:cGhhcm1hUmVmZXJlbmNlRGF0YQ==:NWUyOGJlYTAtNjIzMC00YmJiLTk2MzEtYTViNDhkNjBjYWE3

     5. NEGOTIATE CONTRACT SOURCE
        Negotiation ID: 955039ea-8104-44e3-ac65-c0aae6bc5e4b

     6. GET AGREEMENT ID SOURCE
        Agreement ID: None
          no agreement identifier found, poll again in a second (1)
        Agreement ID: None
          no agreement identifier found, poll again in a second (2)
        Agreement ID: 72673726-762e-41a0-85bc-45d38cd090cf

     7. START TRANSFER SOURCE
        Transfer ID: d65ddd06-5c6a-45f7-a2ce-7b3fea261e85

     8. GET ENDPOINT DATA REFERENCE SOURCE
        Endpoint: None Token: None...
          no data endpoint found, poll again in a second (1)
        Endpoint: None Token: None...
          no data endpoint found, poll again in a second (2)
        Endpoint: http://localhost:19291/public Token: eyJraWQi...

     9. GET THE DATA FROM ACCURIDS API PHARMA
        Download done.
      This is the first download.
```

```
Upload Pharma Reference Data

  10. FETCH CATALOG TARGET FILE
      Dataset ID: MQ==:cGhhcm1hUmVmZXJlbmNlRGF0YUZpbGU=:ODAwN2ZjYTctN2FlNy00NGJkLWJkMDctZmE3YTE1ZTdkMTFj

  11. NEGOTIATE CONTRACT TARGET FILE
      Negotiation ID: 346b5bf7-c2e2-4f68-a656-8560880a7e53

  12. GET AGREEMENT ID TARGET FILE
      Agreement ID: None
        no agreement identifier found, poll again in a second (1)
      Agreement ID: ffcf2cad-26ea-4a6f-9550-8202fa36e1b9

  13. START TRANSFER TARGET FILE
      Transfer ID: 59edbd44-e12e-4c95-aeea-1d548b61539f

  14. GET ENDPOINT DATA REFERENCE TARGET FILE
      Endpoint: None Token: None...
        no data endpoint found, poll again in a second (1)
      Endpoint: http://localhost:19291/public Token: eyJraWQi...

  15. PUSH THE DATA TO ACCURIDS API HOSPITAL FILE UPLOAD

  16. FETCH CATALOG TARGET GRAPHQL
      Dataset ID: MQ==:cGhhcm1hR3JhcGhxbA==:OWZkMDM0NWMtOTU4My00MDdkLThhZGItNzg0NTcyYmFjMmFj

  17. NEGOTIATE CONTRACT TARGET GRAPHQL
      Negotiation ID: 6cf9b60d-4635-4e4d-838d-e63141fcefc3

  18. GET AGREEMENT ID TARGET GRAPHQL
      Agreement ID: fc718118-f206-4c55-8052-1c96d1b05922

  19. START TRANSFER TARGET GRAPHQL
      Transfer ID: 6ce2141d-c569-4b83-b05a-208c18945855

  20. GET ENDPOINT DATA REFERENCE TARGET GRAPHQL
      Endpoint: None Token: None...
        no data endpoint found, poll again in a second (1)
      Endpoint: http://localhost:19291/public Token: eyJraWQi...

  21. PUSH THE DATA TO ACCURIDS API HOSPITAL DATASET UPDATE

      Dataset upload done.
```

Important Requirements for Collaboration with Pharma

Demonstrator scope is implemented

#	Requirement	Scope	Demonstrator Implementation	Status & *Extension Options*
1	**Trusted security framework**	In	Usage FAIR Data Spaces EDC and OAuth2.	**DONE**
2	**Globally Unique, Persistent and Resolvable Identifiers (GUPRIs)**	In	Shared through datasets and resolvable based on OAuth2.	**DONE** *Connect security config with FAIR Data Spaces*
3	**Standardized reference taxonomy**	In	Sharable RDFS class hierarchy based on agreed healthcare industry standards.	**DONE** *Provide shared reference concept library centrally*
4	**Standardized data models (schema)**	In	Sharable SHACL shapes linked to agreed healthcare industry standards.	**DONE** *Provide shared data model library centrally*
5	**Data quality reporting**	Add-on	Execution of data quality before sharing and after subscription/entry of data.	**ONGOING**, Dec/Jan product implementation *Collaborative data quality resolution*
6	**Dataset subscription and synchronization**	Add-on		**FUTURE** *Allow subscription to FAIR DS published datasets*
7	**Intellectual property management**	Out		**FUTURE** (use metadata for now)

FAIR Data Spaces Demo Scenario

Persistent IDs for plasmid vectors based on normalized DNA sequences

A **plasmid vector** is a small, circular piece of DNA that is used as a tool in molecular biology and genetic engineering to carry foreign genetic material into a host organism. It acts as a "vehicle" to transfer and replicate genes within a cell, typically bacterial or eukaryotic cells.

PoC implementation done with a large pharma that is specialized in bio tech engineering to develop personalized medicines
→ ***Screen recording***

Pharma Use Case 2: Pharma + University Research in Bio / Genetics

Early Research on Cell Cultures

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 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: Ensure unique identification of key research objects and standardize their metadata

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.

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

Example for todays research collaboration: https://tron-mainz.de/de

Pharma Use Case 3: Pharma + Health Authorities

Drug Shortage Monitoring

Challenge: Due to the global disruption of supply chains, drug shortages have become a problem also for countries of the European union and Germany 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. E.g., 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 on 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.

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 inventory monitoring systems 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 prevention can be done effectively based on standardized medicinal product data and incrementally along multiple dimensions (a) scope of medicines start with critical and expand to incrementally to more (b) jurisdictions/countries (c) scope of shared information – start with inventory of packaged medicinal products and expand to manufacturing and supply chain information

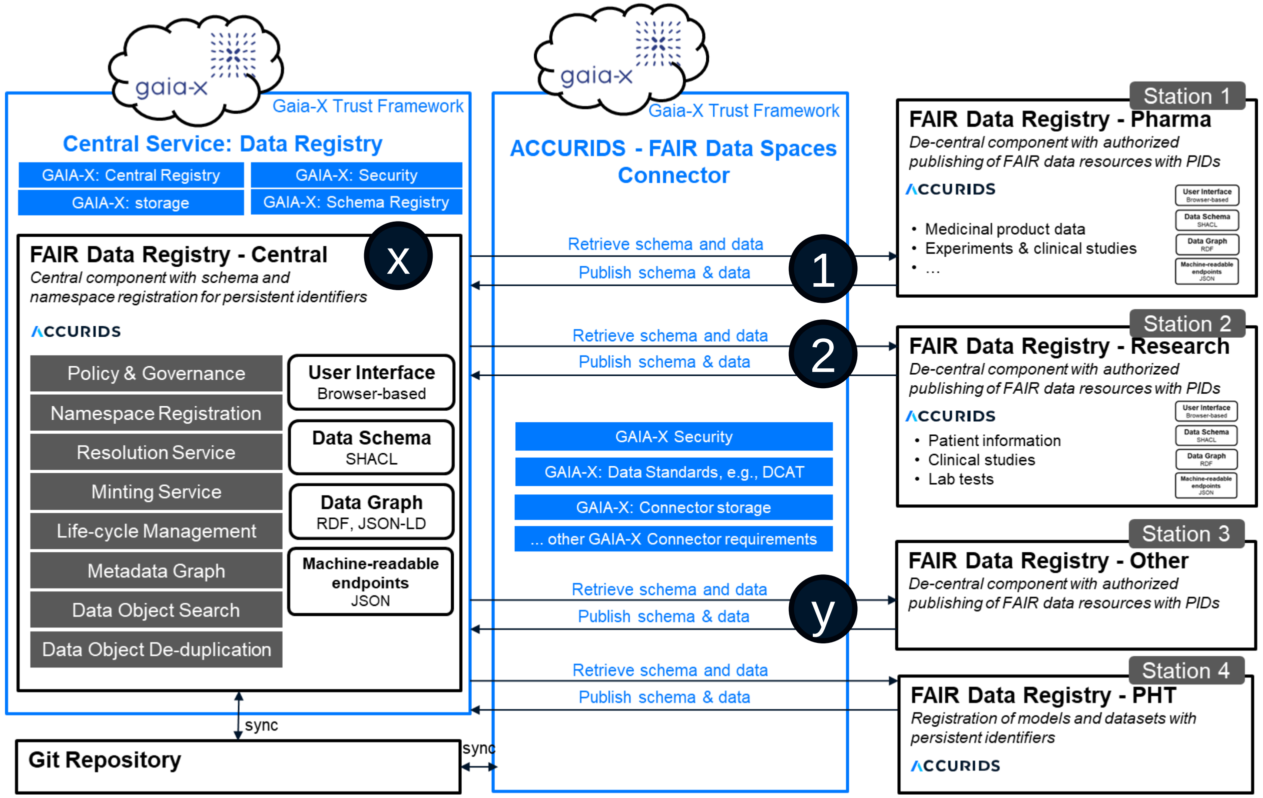
Potential Future Work

With the successful PoC we can…

1. **Augment ACCURIDS UI to share data through EDC-connector** (instead of PoC background script)
2. **Use a central place for sharing metadata** standards and PID uniqueness constraints (SHACL files) to align different organizations before exchanging data.
   → *Connect with NFDI terminology and PID services*
3. **Enhance use case implementations, e.g., on early research collaboration** in context of cell cultures
   → *check relevancy for NFDI4… -Health, -Immuno, -Microbiota*
4. **Implement a PoC with non-ACCURIDS sources** or targets of data exchange

Backup

Implemented Demonstrator Scenario



gaia-x

Gaia-X Trust Framework

Central Service: Data Registry

GAIA-X: Central Registry

GAIA-X: Security

GAIA-X: storage

GAIA-X: Schema Registry

FAIR Data Registry - Central

Central component with schema and namespace registration for persistent identifiers

ACCURIDS

x

- Policy & Governance
- Namespace Registration
- Resolution Service
- Minting Service
- Life-cycle Management
- Metadata Graph
- Data Object Search
- Data Object De-duplication

User Interface Browser-based

Data Schema SHACL

Data Graph RDF, JSON-LD

Machine-readable endpoints JSON

sync

Git Repository

sync

gaia-x

Gaia-X Trust Framework

ACCURIDS - FAIR Data Spaces Connector

GAIA-X Security

GAIA-X: Data Standards, e.g., DCAT

GAIA-X: Connector storage

... other GAIA-X Connector requirements

Retrieve schema and data

Publish schema & data

1

Retrieve schema and data

Publish schema & data

2

Retrieve schema and data

Publish schema & data

y

Retrieve schema and data

Publish schema & data

Station 1

FAIR Data Registry - Pharma

De-central component with authorized publishing of FAIR data resources with PIDs

ACCURIDS

- Medicinal product data
- Experiments & clinical studies
- …

User Interface Browser-based

Data Schema SHACL

Data Graph RDF

Machine-readable endpoints JSON

Station 2

FAIR Data Registry - Research

De-central component with authorized publishing of FAIR data resources with PIDs

ACCURIDS

- Patient information
- Clinical studies
- Lab tests

User Interface Browser-based

Data Schema SHACL

Data Graph RDF

Machine-readable endpoints JSON

Station 3

FAIR Data Registry - Other

De-central component with authorized publishing of FAIR data resources with PIDs

Station 4

FAIR Data Registry - PHT

Registration of models and datasets with persistent identifiers

ACCURIDS

1 The Pharma instance shares schema and data with the university hospital through EDC FAIR Data Space connector.

2 The university hospital shares data about collected human bio samples through EDC FAIR Data Space connector.

x To focus on the decentralized aspect of the FAIR Data Spaces the demonstrator was implemented without a central component. Depending on the FAIR Data Spaces roadmap a corresponding central registry instance can be setup to support multi-stakeholder collaboration.

y In the future, the data exchange with an organization that doesn’t have a ACCURIDS Data Registry instance can be explored.

Demo Steps Summary

1. Show two ACCURIDS instances: Pharma (PH) + University Hospital (UH) with configuration, customization and security setup
2. Both, PH and UH configure a standardize reference taxonomy based on ISO IDMP with classes for clinical trial and human bio samples that specify also the Persistent ID minting mechanisms based on defined characteristic metadata.
3. PH defines a controlled reference terminology for bio sample types
4. EDC FAIR Data Spaces connector is used to share metadata and clinical trial base data from PH with the UH instance
5. The UH sets up the registry for collecting human bio samples as part of a clinical trial with persistent identifiers
6. The UH collects human bio samples and links them to internal patient master data and the clinical trial of the sponsoring PH
7. UH shares bio sample dataset through EDC with PH
8. PH can view bio-sample data in PH instance automatically connected with other PH clinical trial data
9. PH can run analytics queries on interoperable connected trial information
10. Updates and synchronization:
    1. UH collects further bio samples and enriches data about bio samples already shared.
    2. This is automatically shared as a updated dataset with PH
    3. Data about individual samples can be accessed through PIDs that resolve to the mastering UH instance