Advancing Data Sovereignty for Sensitive Data and Patient Data in Africa

An EEPA-VODAN Group Report

Summary: This report serves as a strategic overview and practical reference for policy-makers, funders, and university leadership. It outlines what is already achievable on Data Sovereignty today, identifies available tools, training, and support, and highlights pathways for collaboration and scale-up. The report is intended to support sovereignty-respecting investment, institutional capacity-building, and policy alignment, ensuring that the work of the EEPA-VODAN Group and its partners contributes sustainably to African, European, and global data space developments.

December 2025

Table of content

Foreword: Data Sovereignty, FAIR Data, and Trusted Health & Humanitarian Data Spaces

Data sovereignty is a strategic global priority and a central concern for African and European decision-makers. Across Africa and Europe, governments, universities, and public institutions are increasingly concerned about who controls data, where data are stored, and how they are reused. These concerns are particularly acute for health, patient, and humanitarian data, which are often collected locally but analysed, stored, or monetised elsewhere. Ensuring that data remain under national and institutional jurisdiction, while still enabling responsible use for public good, is a shared challenge across Africa and Europe. It is also a foundational principle of emerging Common Data Spaces, including the European Health Data Space (EHDS), and the Africa Health Data Space.

VODAN, the Value-driven and Ownership of Data and Accessibility Network, was established in 2020 by the Africa University Network on FAIR Open Science (AUN-FOS) to respond directly to these concerns. VODAN applies the FAIR Data Principles, Findable, Accessible (under well-defined conditions), Interoperable and Reusable, as a practical means to operationalise data sovereignty. Its approach enables data to remain with African data holders, while allowing federated, governed, and purpose-driven access for research, policy, and service delivery. This model closely aligns with data space architectures and offers a concrete pathway towards trusted EU and African health and humanitarian data spaces.

The strategic relevance of FAIR Data is well recognised. In 2020, the European Commission and major research funders formally adopted the FAIR Data Principles, embedding them in Europe's data policy framework. FAIR now underpins interoperability, reuse, and innovation across sectors, while respecting legal, ethical, and sovereignty constraints. These developments are reflected in the book FAIR Data – FAIR Africa – FAIR World, which positions FAIR data infrastructures as a cornerstone for equitable, globally connected, and regionally governed data spaces.

VODAN's work is linked to policy and operational practice through its partnership with the Europe External Programme with Africa (EEPA), a centre of expertise on refugee protection and resilience. Together with partners, they form the EEPA-VODAN Group. Through the EU Digital Governance in Africa initiative, EEPA and VODAN collaborate with multiple organisations to establish structured, interoperable, and reusable data infrastructures for sensitive humanitarian and health data. This work contributes directly to the emergence of Health and Humanitarian Data Spaces that respect African and European governance priorities while remaining interoperable with global systems.

What is FAIR Data?

History

Leiden University is the original home of FAIR Data, which was conceptualised at a workshop held in 2014 in the Leiden Lorentz Centre. FAIR data stands for data that is Findable, Accessible (under well-defined conditions), Interoperable and Reusable. It builds on data as Linked Data resources, defining digital data instances in semantically rich terms and transforming them in human and machine-readable formats. The aim is to recognise the value of digital data, including research data, and to ensure that these assets are available for future use. It recognises the rapid growth of volume of data, and that computers are needed to create additional computationally based insights, and increase insights from diverse data sources.

Semantic data or metadata increase the value of the data as archived resources, and allows preservation of the origins of the data. FAIR also stands for Federated AI Ready, speaking to the character of FAIR Data that allows data to speak to each other while in different repositories, controlled by different researchers or data owners. FAIR data is increasingly relevant, since Machine Learning and AI benefit from high quality data; application of the FAIR data principles increase the quality of the data.

Definitions

FAIRification is the process of curating digital data assets with semantic description and machine-readable formats. Semantic enrichment of the data can be defined as data that is expanded with data about the data. This can include different types of metadata, such as the metadata about the research, the researcher and the datasets (provenance and archiving metadata); machine-actonability metadata concerning how and under what conditions the data can be visited by algorithmic queries, metadata about the meaning of data instances itself and linking these to existing standards and definitions of the meaning of the data, as well as machine-readable links on the internet that prescribe the definitions and relations expressed in semantic format.

FAIR-OLR all data are handled by assumption as FAIR with Ownership by data producers, Localisation of repositories and under Regulatory compliance.

FAIR Data Point is a tool to standardise the findability and accessibility to data repositories of data generated and curated by researchers.

Federated data means that data can be controlled in closed repositories, and ensures that the data can be fully protected by the person or group responsible for the data. This includes the computations that are carried out over the data. The strict control over data in FAIR format allows data access to be conditioned on regulatory compliance and ethical considerations, particularly relevant for sensitive data and personal data. The concerns that are managed in a FAIRification process of the data relate to risks such as potential re-identification, security concerns pertaining to the integrity of the data and the ability to protect the data, respecting the highest standards of legal and ethical frameworks.

Federated repositories are places where data is stored and from where data is handled. These can be FAIR-repositories, such as Linked data repositories. Another iteration of FAIR Data is Solid Data, which is particularly designed to protect personal data. The combination of FAIR and Solid gives increasing understanding of how data can be computed across different repositories.

Interdisciplinary FAIRification refers to the various needs that are associated with FAIRification in different disciplines as well as the different skills sets needed for any FAIRification process, including data science, legislation of digital data handling, medicine, computer science, humanities, mathematics, AI, social science and any subject matter that pertains to a domain of data that is handled.

Linked Data is the methodology of creating semantic machine-actionable assets of digital data on the internet that offers an alternative to the concepts of Open Data or unstructured Big Data.

Query: A query is a request for information from a database. It allows users to retrieve specific data by specifying criteria or conditions, making it possible to search, filter, and manipulate information stored within the database. In simpler terms, it's like asking a question to get the answers you need from a collection of data.

SOLID is a setup of Linked Data that is used for personal data repositories.

The Africa University Network on FAIR Open Science (AUN-FOS) hosted by Tangaza University in Kenya, provides a strong institutional anchor for this approach. Tangaza University is committed to evidence-based decision-making and supports the application of FAIR data science across health, social, economic, and humanitarian domains.

Data Space: data remain under the control of the data owners, while enabling secure computation, analytics, and cross-border collaboration, reflecting the governance logic of both the Africa Health Data Space and European data spaces.

Applications

Creation of a FAIR Data Point for Reseach Data

What: The creation of a FAIR Data Point allows you as a researcher to make your data catalogues and datasets findable. The data resides in a secure container, in which the data can only be 'visited' for computation, if permission is explicitly granted.

Why: The findability of your data will enhance the possibility of your academic work to be scrutinised by others, and to understand the credibility of your research and strength of you data.

Examples: There are several examples of FAIR Data Points:

- FAIR Data Point Meta Roestenberg
- A generic LUMC FAIR Data Point is available here
- Other FAIR Data Points can be visited here

How: You can follow instructions from the VODAN GitHub page or ask our advise on how to proceed.

Future possibilities: In the future this will allow researchers and consortia to publish their data in secure and transparent ways, and carry out Machine Learning and AI queries over the data if access is permitted. Other researchers (third parties) can find the data and carry out analytics over the data if permission is granted.

Challenges: The possibility of algorithmic specific data visiting is possible, but this needs to be made more user friendly. Other challenges pertain particularly to automation of data visiting permission, taking into account re-identification risks and cybersecurity concerns and other challenges associated with the data access.

Support: The EEPA-VODAN Group together with LUMC provides FAIR data support which has carried out work to set up the FAIR Data Point. The Roestenberg FAIR Data Point is advanced in that it includes all aspects of a FAIR Data Point: namely all archival metadata, a dashboard of the data resources and substantive metadata on the datasets. The conditions for data access are also identified. The FDP is visitable.

Figure 1. Example of a dashboard.

FAIR Data for epidemiological surveillance

What: The curation of data instances through a FAIRification format, which allows secure running of routine algorithms over the data, while the data remains in single or multiple secure repositories.

Why: This increases the control over data while enabling real-time surveillance on your data, using models through which particular data patterns can be picked up. The work allows interoperability with data in different repositories and of different origin. This also allows participation in consortia which are advancing understanding of the strength of federated queries, research on real-time accessibility and data security.

Examples: VODAN-Africa is a research-project which has carried out a FAIRification programme to deploy real-time Antenatal Care, Delivery and Postnatal Care data data in FAIR-format, allowing real-time insights in the data to emerge.

How: The work involves the creation of common data-models that are mapped to common standards, and is carried out in collaboration with LUMC Global and the colleagues from gynecology.

Challenges: There is still research required to develop these into mature systems.

Support: The researchers of the Chair FAIR Data Science focus on the maturisation of the federated FAIR data systems and are interested to provide assistance as part of the research programme.

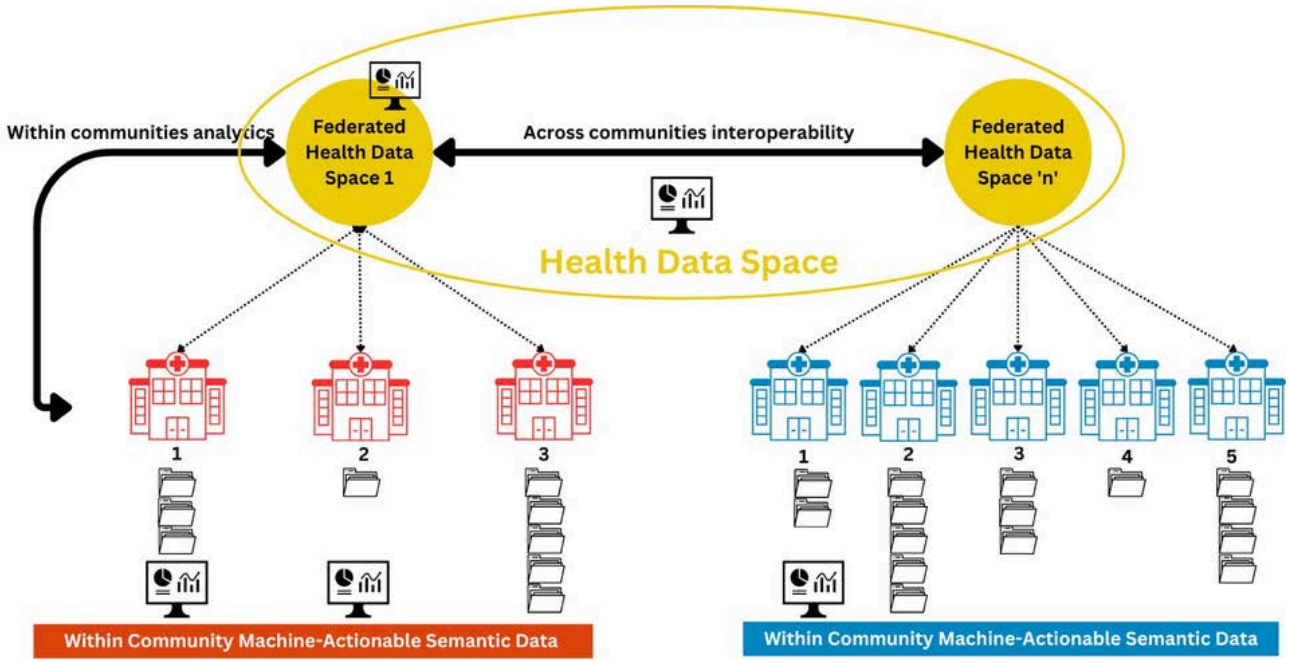


Figure 2. FAIR-OLR Federated Analysis and Learning.

Understanding interoperability needs on patient data for epidemic preparedness

What: Study the possibility of creating common data models to use patient data for surveillance for different purposes.

Why: This research specifically advanced federated analytics on different types of surveillance for epidemic control with data that are under control of hospitals.

Examples: Together with RIVM (PRAISE programme) and masterstudents in the Leiden Institute for Advanced Computer Science (LIACS), and the Dutch Government Federated Architectures, we study the conditions for application of federated data modes for patient data for early epidemic detection in patient data, while allowing the patient data also to be re-used for analysis for other purposes, using similar models across two continents (Europe and Africa).

How: The work involves the creation of workflows that can be tested in different places.

Challenges: There is still research required to develop these into mature systems.

Support: The researchers of the Chair FAIR Data Science focus on the use of patient data in the federated FAIR data systems and are interested to provide assistance as part of the research programme. It is implemented with the VODAN research group.

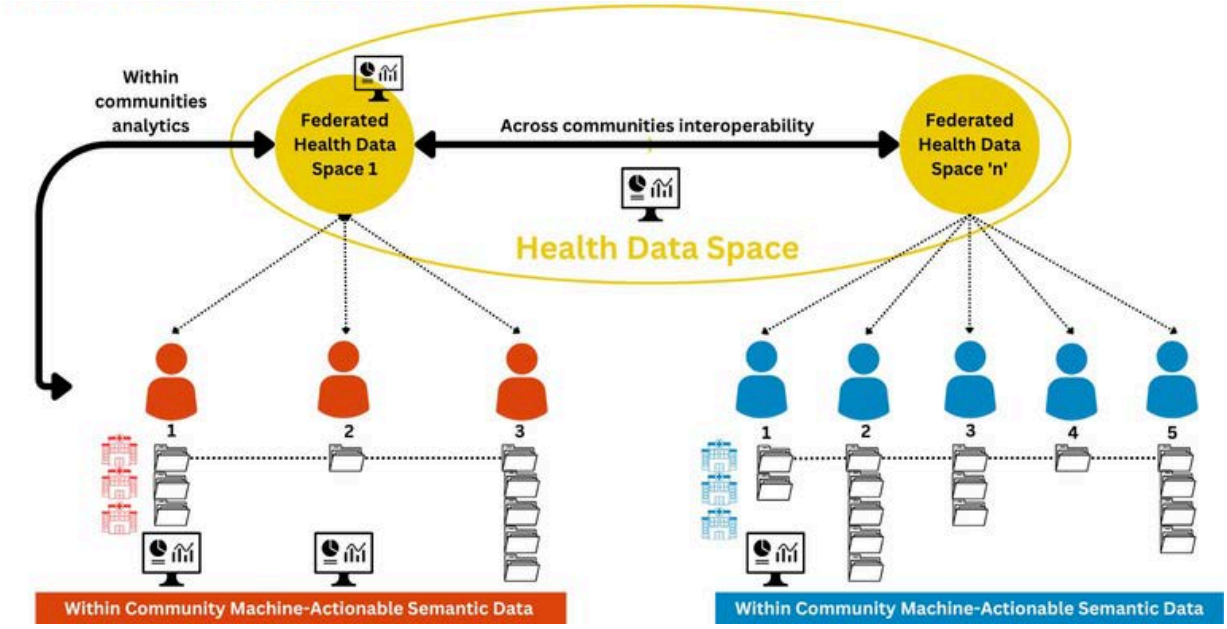


Figure 3. FAIR-OLR Federated Personal Pods.

Investigating modalities of inclusion of highly sensitive data of vulnerable populations in federated analytical models based on FAIR

What: Study the possibilities of targeted computations in data-visiting models using highly secure repositories in a federated data architecture. It is funded by a Leiden University Seed Money programme as well EEPA's 'Digital Governance in Africa' programme.

Why: This research focuses on inclusion of highly sensitive data for federated analytics in secure ways, increasing the inclusion of sensitive data. More inclusive datasets will increase the power of the AI applications developed to run over federated data architectures.

Examples: The EEPA-VODAN Group focuses on understanding of data bias and exclusion of vulnerable populations, creating population bias in the data. The populations include pregnant women in conflict situations and victims of sexual violence. A collaboration with the Africa University Network for FAIR Open Science and the Africa Health Data Space offer a collaboration for data analytics with data stored in separate repositories under the control of the health facility or university. FDPs have been set up in Ethiopia, Kenya, Nigeria and Uganda. The link to the programme is here.

How: The work involves interdisciplinary training to build trust with representatives working with populations in vulnerable circumstances, and training to create of common data-models that are mapped to common standards as well as strong conditions for access control on the data, with clear regulatory and ethical permission conditions specified, understood and translated in operational terms. The focus is on maximising operational strategies to ensure personal data protection, including the challenge of reidentification in large datasets.

Challenges: There is still research required on how to strengthen the availability of sensitive data and vulnerable populations in large federated AI models and ensure protection from reidientification.

Support: The researchers of the Chair FAIR Data Science focus on the use of patient data in the federated FAIR data systems and are interested to provide assistance as part of the research programme. This is implemented with the EEPA-VODAN Group.

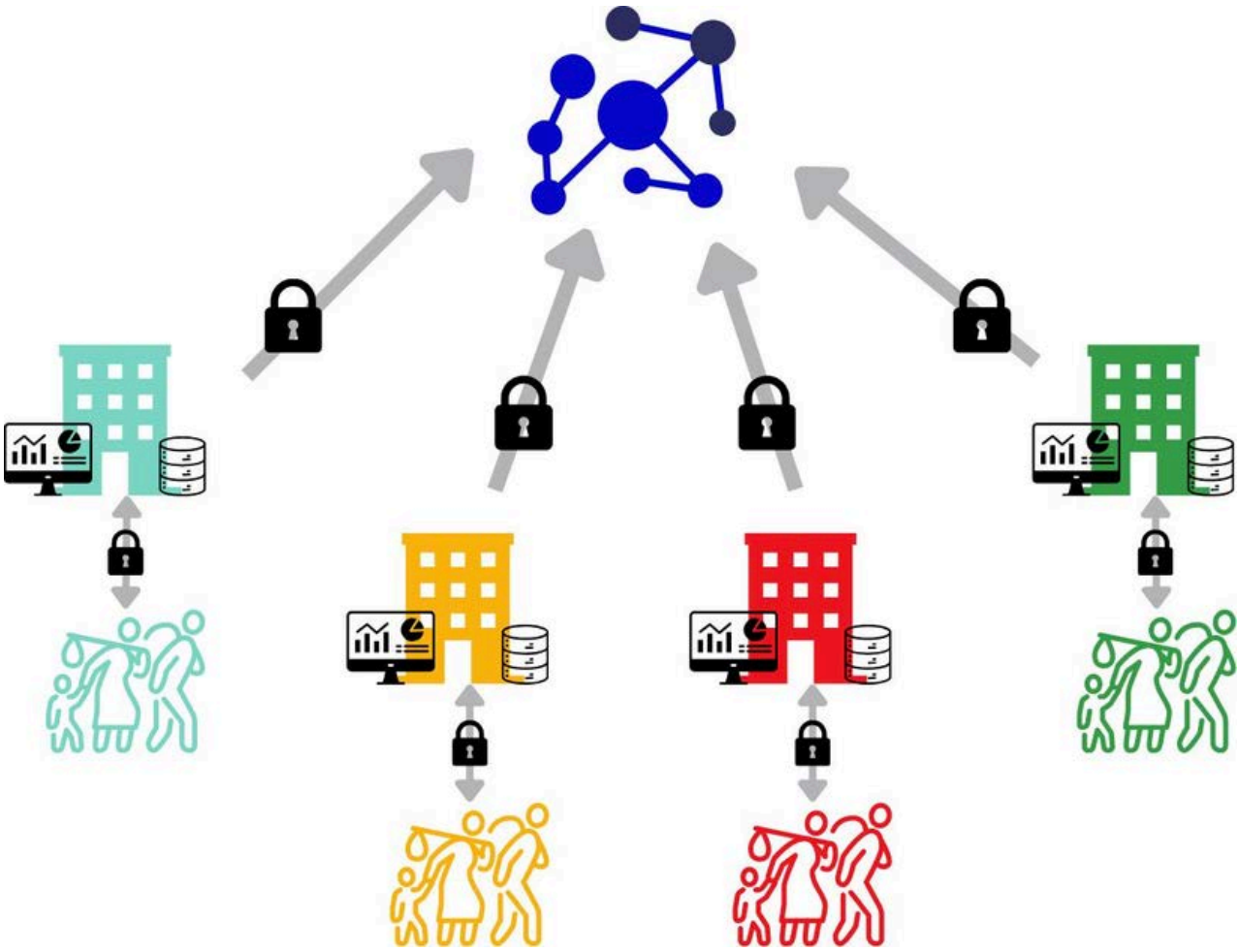


Figure 4. Federated analytics model for sensitive data from vulnerable populations.

FAIR data stewardship for humanitarian data

Architecture developed

An architecture was developed that facilitates the capture of humanitarian data in a FAIR-compliant manner. FAIR stands for Findable, Accessible, Interoperable, and Reusable, emphasising the importance of ensuring that data is reusable and machine-actionable. The strategy followed the ideal scenario of creating FAIRification by design. This means that the process begins at the moment of data creation. The architecture allows for the secure capture of sensitive data, promoting its reuse and sharing while maintaining strict security protocols. The data remains at its point of origin, guaranteeing clear ownership and localisation. Access to the data or its metadata is provided through a data visiting mechanism, which requires approval from the data owners.

Our architecture has been deployed and tested in humanitarian organisations and clinics in Ethiopia, Uganda, Kenya, and Nigeria. Data stewards from South Africa, Namibia, Ghana, Liberia, Rwanda, Nigeria, and Guinea-Bissau have received training.

The core architecture consists of the following elements:

Common Data Model (CDM): Common Data Models have been established all based on an ontology developed for the Humanitarian Data Space. This vocabulary maps out important concepts and their relationships to one another, as well as their relationship to existing foundational ontologies to ensure interoperability. Our ontology supports a common data space for linked humanitarian and sensitive data.

Extract, Transform, Load (ETL) pipeline: The principal component of the architecture is the ETL pipeline, which starts with a data input model, followed by an ontology mapper linking the data to the CDM. The data is then transformed into Resource Description Format (RDF), a language that formats data instances through their relationships with one another. Once the data is in RDF, its syntax is validated before being loaded into a triple store.

Monitoring dashboard: The aim of the architecture and the data space is for queries to be run on the different datasets. Through these queries, users can obtain information from the datasets, without directly getting access to the data itself. The monitoring dashboard allows for certain information to be extracted from the datasets and viewed by users – usually in an aggregated format. This would for instance include information on the whereabouts of refugees, occurrence of incidents, etc.

Access Control: Since in these applications, we are working with very sensitive data, one of the fundamental aspects is access control mechanisms. Whilst the aim of the data space is to enable more information to be extracted from datasets, it is essential that privacy and security of data subjects is preserved, as well as data ownership. To ensure this, access control mechanisms are in place in the architecture to ensure that the owners of the data can control who has access to their data, and that access can be revoked at any time. The consent by the data-owner determines what metadata of source data can be queried and on which metadatasets the queries can be conducted. The aggregates are computed over the queries implemented in federated format over annotated datasets for which the query has been consented to.

FAIR Data Points (FDPs): The architecture contains machine-actionable information on all repositories information about datasets available in HDS, contained in catalogues published in FDPs, pointing towards the conditions for access to datasets that are available in HDS FAIR formats. The FDPs are hosted in local servers, and ensure that all datasets are findable, even if the contents thereof may not be available to all.

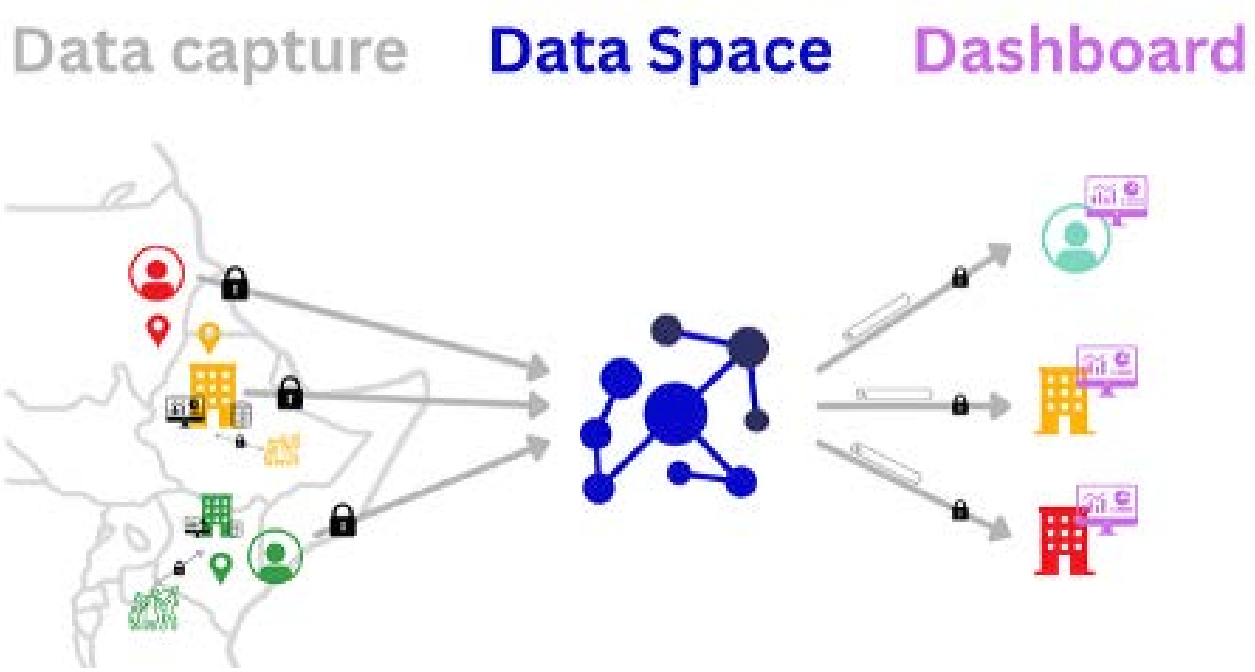


Figure 5. Process of data capture into the data space, monitored through a dashboard.

Use cases

The EEPA-VODAN Group facilitated the co-creation and development of several use cases. These were co-designed with humanitarian organisations. The organisations has requested specific needs to be addressed. The specifications and requirements relate to the architecture set out for the Humanitarian Data Space.

The co-design resulted in different applications that resolve various data problems in different communities. The applications are all interoperable as the data is (a) published in the same syntax, namely RDF, (b) use the same vocabulary and CDM developed for the Humanitarian Data Space, (c) published in triple stores.

Application 1. **Reporting of incidents of Gender-Based Violence (GBV)**
This use case relates to reporting of GBV incidents that take place in sensitive contexts, such as in conflict areas. An application was developed that allows organisations and persons that work with survivors of GBV to input information in a FAIR format.

Application 2. **Reporting of information for Situation Reports**
EEPA has been publishing regular situation reports since 2020. However, the information in these reports, whilst valuable, is not published in a FAIR format, limiting its usability and interoperability. To address this, we created an application to enable the reporting of relevant information from sources on the ground. This data is FAIR by design, and stored in a triple store within our server. Reports in pdf format are created based on this data, but data instances can also be queried upon to, for instance, better see how trends progress over time.

Application 3. **Reporting of information related to refugee needs and protection**
An application was developed to facilitate humanitarian organisations on the ground to respond to needs of refugees and IDPs. The platform enables workers within camps to input information related to scarcity, needs, and events that are taking place within the camps and surrounding areas. This information can be shared with relevant organisations to ensure better and more efficient responses.

Application 4. **Reporting of human-trafficking related information**
The aim of this application is to enable source-persons who have access to information about human trafficking events and networks to input this information in a FAIR-compliant way. This kind of information is often spread through WhatsApp and similar communication platforms by survivors to source-persons. By making sure that the information can be stored in RDF format and linked to other similar data, better understanding can be acquired from the information.

Application 5. **Tools for the FAIRificartion of human trafficking research data**
This application is similar to the one above, but related specifically to research data, such as interview data. These are often kept in excel files and similar formats. The analysis of this data requires a lot of manual labour. This application focuses on the input of such excel files, enabling the contents to be extracted and linked to vocabulary of the Humanitarian Data Space. As such, the information can be linked to other relevant datasets, but can also be analysed and visualised in a graph format in of itself.

Application 6. **Personal pods for refugees and survivors of human trafficking**
This use case is targeted for refugees and survivors of human trafficking themselves, consisting of personal pods, based on the Solid project, which allows users to control their data and information, and potentially share this with relevant organisations. The first part of the pods consists of the vault – it is meant for users to keep all their documentation in a digital format in one place. This is relevant for those that are on the move, sometimes migrating from country to country. The pods ensure that important documents, such as IDs, passports, but also those relating to legal and medical history, can move with them. The second part of the pods consist of a communication system. This allows users to share information or documents with relevant organisations and instances, including their location, situation, condition, needs, etc. They can revoke this access to any organisation at any time, thereby maintain control and ownership of their data.

Application 7. **Deriving Antenatal Care (ANC) from Afya.ke EMR**
The final use case relates to information collected by clinics on ANC patients. The patient data is immediately collected in a FAIR format and stored in a triple store within the clinic itself. The data never leaves the clinic, but the datasets can be visited through the mechanisms described above. By including this use case in the Humanitarian Data Space, linkages between medical and humanitarian data can be made, which is relevant for survivors of human trafficking, GBV, and people that find themselves in vulnerable situations.

Investigating modalities of inclusion of highly sensitive research data in federated repositories governed by researchers

What: Study the possibilities of triangulation with different types of research data. this programme was supported by UNA Europa FAIR Data for PhD Students.

Why: This research focuses on interoperability across different research sources, for instance in consortia with multiple research modalities.

Examples: The EEPA-VODAN Group and partners carry out interdisciplinary research using sources from different PhD research projects to carry out triangulation. FDPs have been set up in different European and African countries, to test cross country and cross continental collaboration on federated repositories connected with FDPs.

How: The work involves interdisciplinary collaboration across researchers to understand the value of across source triangulation and engage in training to understand what steps are involved in FAIRification to strengthen interoperability.

Challenges: There is still research required to strengthen the availability of sensitive data and vulnerable populations in large, federated AI models and ensure protection from reidentification.

Support: Training for PhD students has been developed and is available on request from the researchers of the FAIR Data Science programme. This is implemented with the EEPA 'Digital Governance in Africa' programme through its collaboration with the VODAN research group.

Expertise: Centres of Expertise have been established in: Ethiopia, Kenya, Nigeria

Challenges: There is still research required to strengthen the availability of sensitive data and vulnerable populations in large, federated AI models and ensure protection from re-idientification.

Support: Training for PhD students has been developed and is available on request from the researchers of the FAIR Data Science programme. The training has been developed and tested with the EEPA 'Digital Governance in Africa' programme in its collaboration with the VODAN research group.

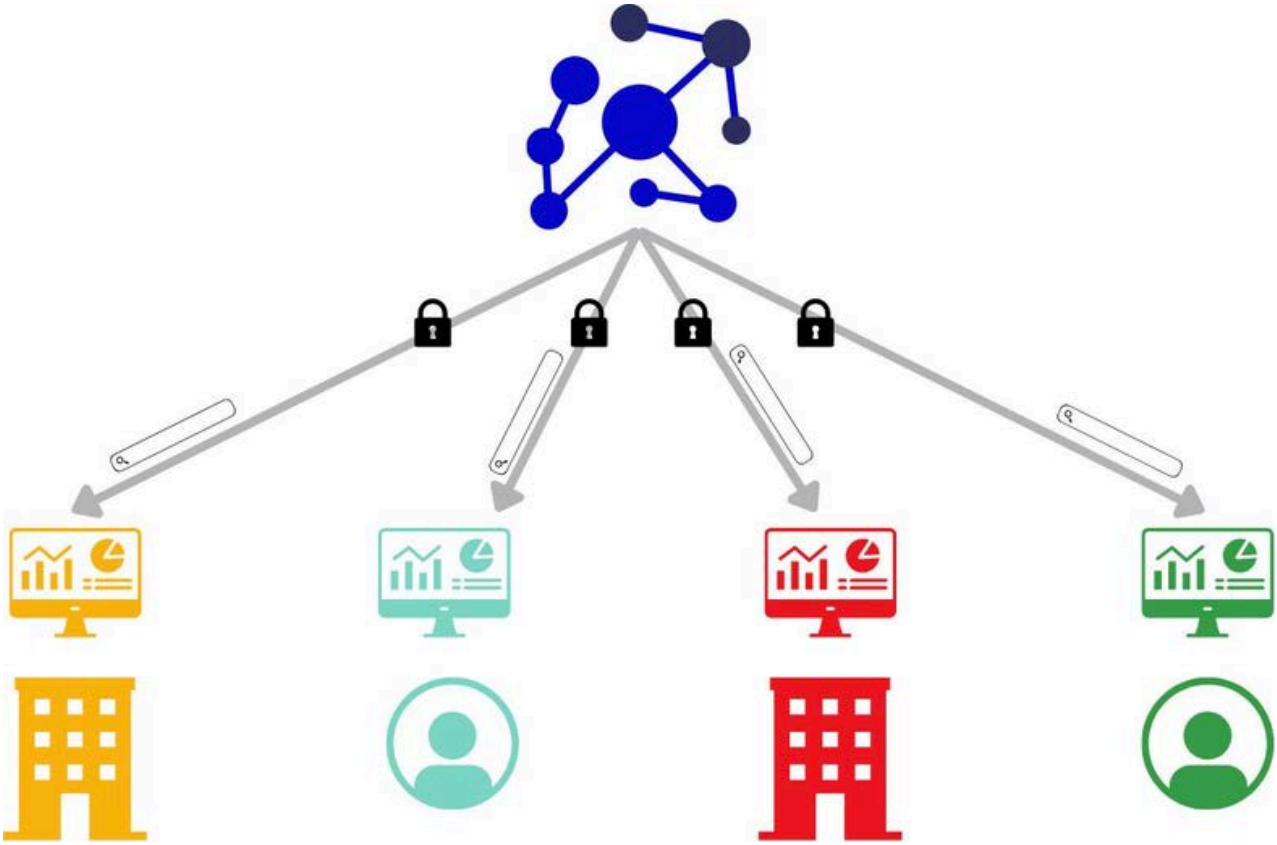


Figure 6. Framework for data visiting from different sources.

Development of a Health Data Space

What: Study the possibilities of the creation of a Health Data Space.

Why: FAIR Data allows the creation of Health Data Space, such as the European Health Data Space and the Africa Health Data Space.

Examples: VODAN research has focused on data sovereignty, keeping data with full ownership in localised residence under regulatory compliance in jurisdiction. The leapfrogging of an African Health Data Space tests the understanding of options for the creation of the European Health Data Space.

How: The work involves collaboration across countries in different jurisdictions with clearly set principles. These are called FAIR with ownership in Local under Regulatory Compliance (FAIR-OLR). The work involves collaboration with regulators such as the European Commission and the African Union. See the following links:

- AHDS conference
- Pan-African Parliament champions Africa's Quest for Data Sovereignty and Ethical AI

Challenges: The technical challenges are mostly resolved, but the challenges remain in the needs for awareness of the potential of such systems, training and regulatory arrangements supporting such architectures. The research focuses on access permission control.

Support: The FAIR Data Science group is interested in collaboration with researchers who want to test how their data can be exposed in an Health Data Space.

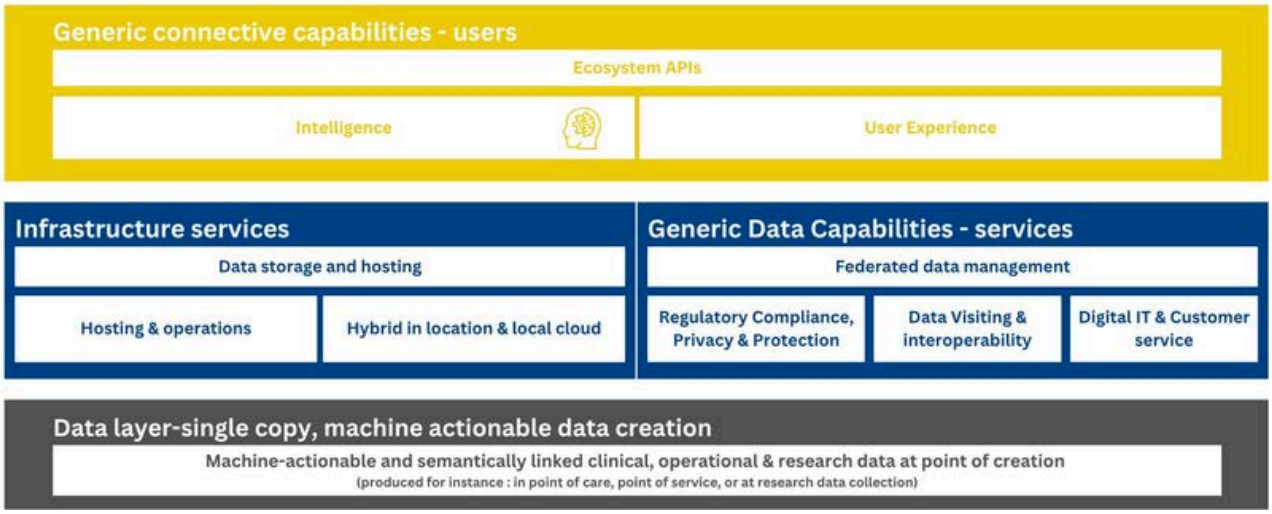


Figure 7. Federated Health Data Space - a proposed landscape.

Investigating Access Control Permission

What: Study the scenarios for granular access control permission in federated data architectures that form a data space that develops and grows in real time, assisting the multidirectional orchestration and reuse of data analytics.

Why: FAIR-data allows for handling of granular permission of access to parts of the data for the generation of insights. A expansion of access-control options will increase the intelligence layer of the Data Space as a secure safe and. privacy preservation layer to implement controlled computations.

Examples: Implementing different types of clinical research over routine patient data.

How: Create tools to respond to multi-variate access control permissions under strict regulatory and ethical considerations on mature data-sets developed in real-time with granular FAIRification at instance level, integrated in a number of different FDPs to test different scenarios.

Challenges: The concerns relating to risks such as potential reidentification, security concerns pertaining to the integrity of the data and the ability to protect the data, respecting the highest standards of legal and ethical frameworks.

Support: Collaborating with the Gent group of Decentralized Web Technology to develop personal patient pods.

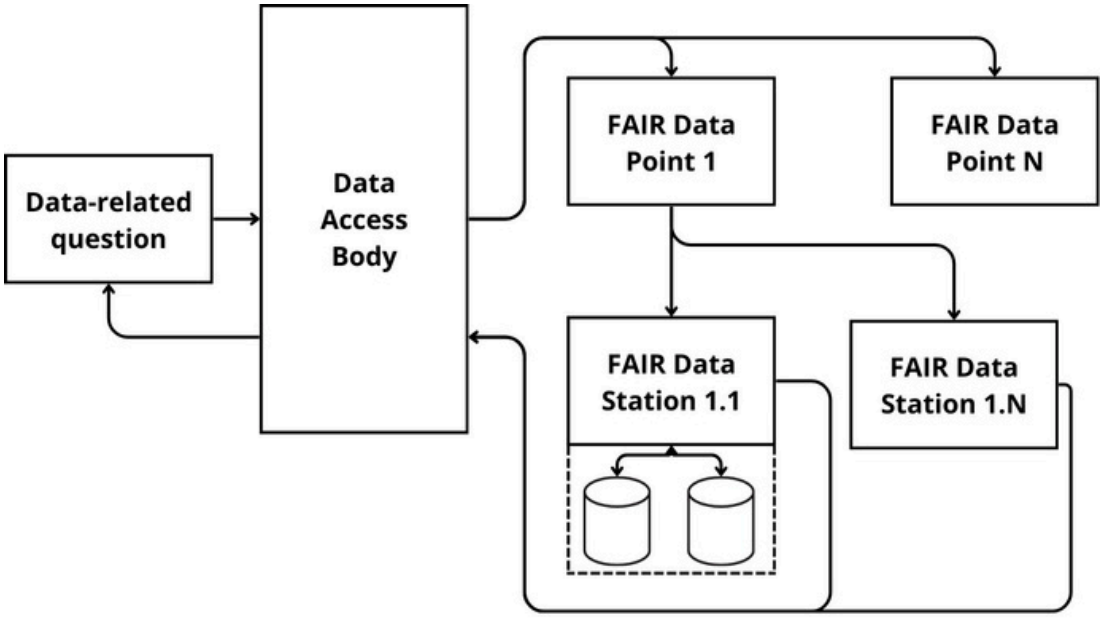


Figure 8. Investigating of Access-Control in Federated multidirectional reuse scenarios of FAIR data

FAIR-OLR Access Control as a Process
A field overview in the context of data visiting

R. Kievit*, R. Plug, S. Y. Amare and M. Van Reisen

Problem Statement

Since the publication of the FAIR guiding principles in 2016, a lot of FAIR research has focused on enhancing the Findability, Interoperability and Reusability of open data with research on Accessibility mostly aimed at equitable access. However FAIR does not necessitate open data [1]. The importance of this notion has been realized, however a proper integration of access control in the FAIR architecture is still lacking.

By extending the FAIR guiding principles with Ownership, Localization and Regulatory compliance to FAIR-OLR [2] we further highlight the importance of both privacy preserving data handling and keeping data at its source. This leads us to the concept of **data visiting** where instead of sharing data between data providers in a federated ecosystem. We send the algorithms to the data providers and only return verified results. A current implementation of this concept is VODAN Africa.

In the data visiting paradigm we can think of access control at all points of the access pipeline. At each stage checks and balances can be included as policies which influence how the query is executed at the data and which results are returned to the user.

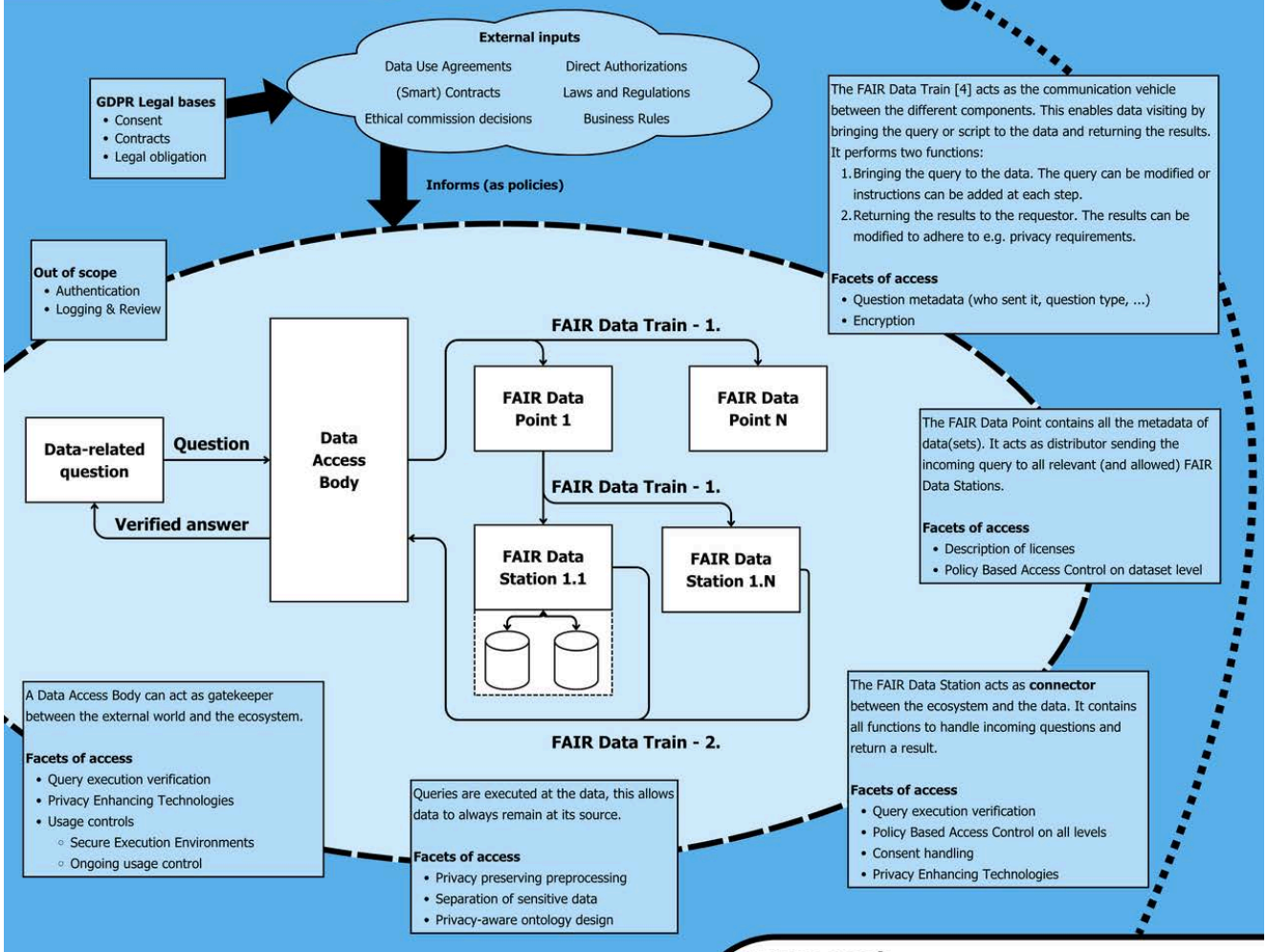
Technological Context

The most viable current federated ecosystem protocol are the EU Data Spaces (IDSA, iSHARE, GAIA-X, etc.). While control of data is often kept close to the data owner, analytics are still typically enabled by sharing data and enforcing usage control policies afterwards [3]. Furthermore authorization in current Data Space implementation is often quite limited - being restricted to simple XACML policies or direct authorizations granted to a data consumer.

In Semantic Web research a lot of facets of access control have been developed such as access control frameworks, machine interpretable law and semantic contracts. However a comprehensive view on access control that incorporates all of these aspects is still missing. Especially in the data visiting paradigm this has not yet been sufficiently considered.

This leads to the main research question of this poster:

Where do the various facets of access control manifest in the FAIR architecture?



The FAIR Data Train [4] acts as the communication vehicle between the different components. This enables data visiting by bringing the query or script to the data and returning the results. It performs two functions:

1. Bringing the query to the data. The query can be modified or instructions can be added at each step.
2. Returning the results to the requestor. The results can be modified to adhere to e.g. privacy requirements.

Facets of access
- Question metadata (who sent it, question type, ...)
- Encryption

The FAIR Data Point contains all the metadata of data(sets). It acts as distributor sending the incoming query to all relevant (and allowed) FAIR Data Stations.

Facets of access
- Description of licenses
- Policy Based Access Control on dataset level

A Data Access Body can act as gatekeeper between the external world and the ecosystem.

Facets of access
- Query execution verification
- Privacy Enhancing Technologies
- Usage controls
  - Secure Execution Environments
  - Ongoing usage control

Queries are executed at the data, this allows data to always remain at its source.

Facets of access
- Privacy preserving preprocessing
- Separation of sensitive data
- Privacy-aware ontology design

The FAIR Data Station acts as **connector** between the ecosystem and the data. It contains all functions to handle incoming questions and return a result.

Facets of access
- Query execution verification
- Policy Based Access Control on all levels
- Consent handling
- Privacy Enhancing Technologies

Future work

We aim to expand this framework to capture all relevant facets of access and connect these facets to tools and methods used by academia and industry that implement (parts of) them. With this we aim to build a safe data access pipeline based on the FAIR-OLR principles using existing tools where possible, or developing new tools where necessary.

*Corresponding Author
Rens.Kievit@rijksoverheid.nl

[1] Jati, P.H.P. (2022). Data Intelligence, 4.
[2] Van Reisen, M. (2023). FAIR Connect, 1.
[3] Munoz-Arcentales, A. (2019). Procedia Computer Science, 160.
[4] Bonino da Silva Santos, L.O. (2023). SN Computer Science, 4.

Federated Data Analytics

What: Studying the possibilities of the creation of a federated Health Data Space for federated data analytics.

Why: FAIR Data allows for federated data analytics in an data space, such as the European Health Data Space and the Africa Health Data Space.

Examples: VODAN research has focused on data sovereignty, keeping data with full ownership in localised residence under regulatory compliance in jurisdiction. VODAN has proposed a structure for a federated Health Data Space which includes multiple layers, including a data layer, a service layer and a user layer.

How: The work involves collaboration across countries in different jurisdictions with clearly set principles. These are called FAIR with ownership in Local under Regulatory Compliance (FAIR-OLR).

Challenges: The technical challenges are mostly resolved, but the challenges remain in the needs for awareness of the potential of such systems, training and regulatory arrangements supporting such architectures. The research focuses on access permission control.

Support: The FAIR Data Science group is interested in researchers who want to test how their data can be exposed in a federated Health Data Space.

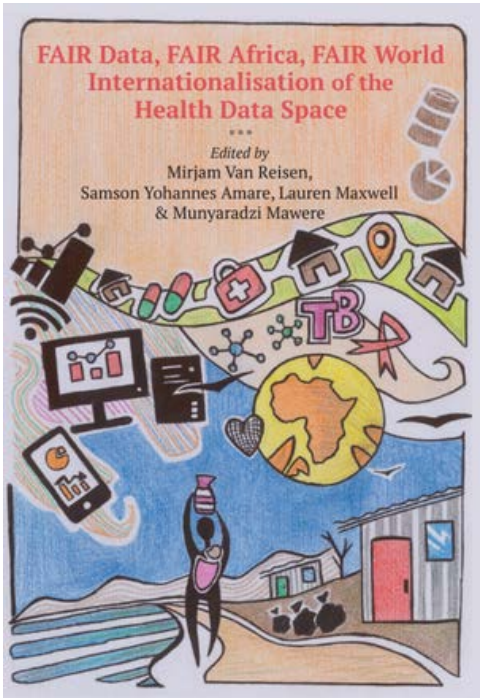


Figure 9. Book FAIR Data, FAIR Africa, FAIR World.

Internationalisation of the Health Data Space.

https://africanbookscollective.com/books/fair-data-fair-africa-fair-world/

https://raee.eu/fair-data-fair-africa-fair-world-internationalisation-of-the-health-data-space/

Contextual adaptation of tools for FAIRification

What: Study the possibilities of the creation of a Health Data Space

Why: FAIR Data allows the creation of Health Data Space, such as the European Health Data Space and the Africa Health Data Space.

Examples: VODAN research has focused on data sovereignty, keeping data with full ownership in localised residence under regulatory compliance in jurisdiction. The leapfrogging of an African Health Data Space tests the understanding of options for the creation of the European Health Data Space.

How: The work involves collaboration across countries in different jurisdictions with clearly set principles. These are called FAIR with ownership in Local under Regulatory Compliance (FAIR-OLR).

Challenges: The technical challenges are mostly resolved, but the challenges remain in the needs for awareness of the potential of such systems, training and regulatory arrangements supporting such architectures. The research focuses on access permission control.

Support: The FAIR Data Science group is interested in researchers who want to test how their data can be exposed in an Health Data Space.

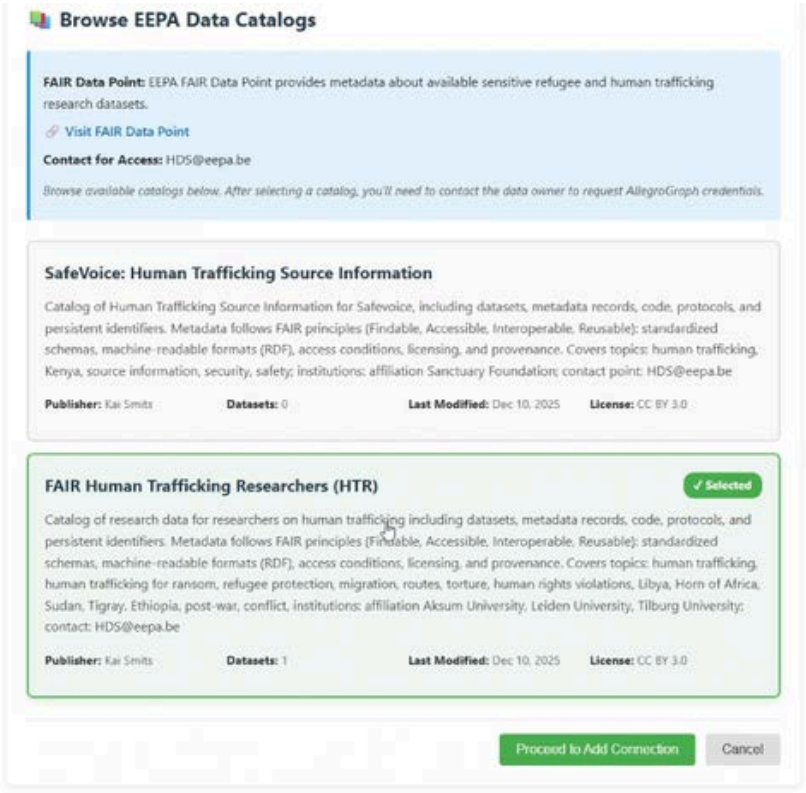


Figure 10. A screenshot of a tool for FAIRification of human trafficking research data.

Curriculum and Training Development

2024

2-day FAIR Data stewardship training

2025

One day intensive introduction to FAIR Data stewardship

One week FAIR Data stewardship training for PhD students with UNA Europa

6 months FAIR data curation, training Seed money project

3 months Training for Senior FAIR Datastewards with VODAN Team

3 months Training for FAIR Engineers in Specialised areas with VODAN Team

Data Science in Practice (master students)

Regulatory Governance in Data science (master students)

Leadership for Sustainable Development (master students)

Mobility, Migration and Globalisation (Bachelor students)

Curricula links:

AHDS FAIR senior data stewardship training

EEPA-VODAN Group Organisation

Centres of Excellence

- Tangaza University - Nairobi - Kenya
- Mekelle University - Mekelle - Ethiopia
- IBBUL - University - Lapai - Nigeria

Countries:

- Kenya
- Ethiopia
- Uganda
- Rwanda
- Ghana
- Guinea Bissau
- Liberia
- Namibia
- South Africa
- Zimbabwe

Experts FAIR Engineering

- FDP establishment
- Common Data Models
- Access Control
- Application development (ETL)
- Documentation

Members of the Board

Prof. Dr. Mouhamad Mpezamihigo (Chair)
Prof. Dr. Francisca Oladipo (Vice-Chair)
Prof. Dr. Janice Busingye (Administration)
James Wilderspin (Audit)
Han Baartmans (Remuneration)

Memoranda of Understanding on exchange of expertise

- AU Pan African Parliament: https://pap.au.int/en/news/press-releases/2025-07-25/pan-african-parliament-champions-africas-quest-data-sovereignty-and
- Africa University Network on FAIR Open Science: https://aun.mu.edu.et/ with Tangaza University, Kenya, University of Nairobi, Kenya, Equator University of Science & Technology, Uganda, Great Zimbabwe University, Zimbabwe, Mekelle University, Ethiopia, University of Grand Bassa, Liberia.

References (selection)

Plug, R. (2025). Secure distributed machine learning in healthcare: A study on FAIR, compliance and cybersecurity for federated learning (Doctoral dissertation, Leiden University). Leiden University Repository.

Kievit, R., Amare, S. Y., van Reisen, M., Stocker, J. & Gotlieb, P. (2026). Wireless technologies for disaster management and protection. *Handbook of Sustainable Wireless Communication Systems: Innovations and Application*.

Stocker, J. (2026). Assessing Ecosystem Resilience through Passive Acoustic Monitoring: Current Challenges and Potential Applications. *Handbook of Sustainable Wireless Communication Systems: Innovations and Application*.

Plug, R. B. F., Amare, S. Y., Jati, P. H. P., Kievit, R., Kawu, A. A., Liang, Y., van Reisen, M. (2025). Secure distributed machine learning for edge computing. In A. L. Imoize, F. Xhafa, M. S. Obaidat, & H. H. Song (Eds.), *Cybersecurity Defensive Walls in Edge Computing*, Ch.10. Elsevier. ISBN: 978-0-443-34109-0.

Plug, R. B. F., Jati, P. H. P., Amare, S. Y., van Reisen, M. (2025). Benchmarking TinyML Encrypted Federated Learning with Secret Sharing in Medical Computer Vision. In A. L. Imoize, D. Dinh-Thuan, & H. H. Song (Eds.), *Tiny Machine Learning: Design Principles and Applications*, Ch.19. Wiley-IEEE Press. ISBN: 978-1-394-29454-1.

Jati, P. H. P., van Reisen, M., Amare, S. Y., Kawu, A. A., Nandwa, W. M., Taye, G. T. (2025 - forthcoming). Promotion of Data Autonomy to Support Insights From Patient Health Data Records in Africa.

Van Reisen, M., Amare, S. Y., Maxwell, L. & Mawere, M. (2025). *FAIR data, FAIR Africa, FAIR world: Internationalisation of the Health Data Space*. Bamenda: Langaa . https://raee.eu/fair-data-fair-africa-fair-world-internationalisation-of-the-health-data-space/

Van Reisen, M., Oladipo, F., Mpezamihigo, M., Plug, R., Basajja, M., Aktau, A., Purnama Jati, P.H., Nalugala, R., Folorunso, S., Amare, Y.S., Abdulahi, I., Afolabi, O.O., Mwesigwa, E., Taye, G.T., Kawu, A., Ghardallou, M., Liang, Y., Osigwe, O., Medhanyie, A.A., Mawere, M. (2022). Incomplete COVID-19 data: The curation of medical health data by the Virus Outbreak Data Network-Africa. *Data Intelligence* 4(4). doi: 10.1162/dint_e_00166

Kievit, R., Kawu, A. A., van Reisen, M., O'Sullivan, D., & Hederman, L. (2024). Semantic integration of heterogeneous PGHD sources in a digital health system. In *press*.

Nalugala, R. & Van Reisen, M. (2024). Ownership of health data as the new frontier for the future of VODAN-Africa. *JHJA*, 2(2), 1–13. https://repository.tangaza.ac.ke/items/c85b58ca-7319-4e4e-a215-80c1b4ecda51/full

Kawu, A. A., Kievit, R., Abubakar, A., Van Reisen, M., O'Sullivan, D. & Hederman, L. (2024). Exploring the Integration of a Patient Generated Health Data in a FAIR Digital Health System in Low-Resourced Settings: A User-Centered Approach. *Proceedings of the 4th African Human Computer Interaction Conference* (AfriCHI 2023), 215–220. https://doi.org/10.1145/3628096.3629059

Van Reisen, M., Amare, S.Y., Nalugala, R., Taye, G.T., Gebreselassie, T.G., Medhanyie, A.A., Schultes, E. & Mpezamihigo, M. (2023). Federated FAIR principles: Ownership, localisation and regulatory compliance (OLR). In: *FAIR Connect*, 1(1), 1-7. IOS Press. DOI: https://doi.org/10.3233/FC-230506

Amare, S. Y., Taye, G. T., Gebreselassie, T. G., Plug, R., & van Reisen, M. (2023). Realizing health data interoperability in low connectivity settings: The case of VODAN-Africa. In: *FAIR Connect*, 1(1), 55-61. IOS Press. DOI: https://doi.org/10.3233/FC-221510

Cardoso Silva Ferreira, A., & van Reisen, M. (2023). Hourglass-based interoperability through nanopublications in VODAN-A. In: *Fair Connect*, 1(1), 5-11. IOS Press. DOI: https://doi.org/10.3233/FC-221512

Gebreslassie, T. G., van Reisen, M., Amare, S. Y., Taye, G. T., & Plug, R. (2023). FHIR4FAIR: Leveraging FHIR in health data FAIRfication process: In the case of VODAN-A. In: *FAIR Connect*, 1(1), 49-54. IOS Press. DOI: https://doi.org/10.3233/FC-230504

Van Reisen, M., Oladipo, F., Stokmans, M., Mpezamihigo, M., Folorunso, S., Schultes, E., Basajja, M., Aktau, A., Yohannes Amare, S., Tadele Taye, G., Hadi Purnama Jati, P., Chindoza, K., Wirtz, M., Ghardallou, M., van Stam, G., Ayele, W., Nalugala, R., Abdullahi, I., Osigwe, O., Graybeal, J., Abrha Medhanyie, A., Abubakar Kawu, A., Liu, F., Wolstencroft, K., Flikkenschild, E., Lin, Y., Stocker, J., Musen, M.A. (2021). Design of a FAIR digital data health infrastructure in Africa for COVID-19 reporting and research. In: *Advanced Genetics*, Volume 2, Issue 2. https://doi.org/10.1002/ggn2.10050

Acknowledgements

- **Prof. Dr. Anna Fensel,** Wageningen University: Interoperability between health and food
- **Prof. Dr. Dympna O Sullivan**, Dublin University of Technology: FAIR adoption
- **Dr. Lauren Maxwell**, Heidelberg University: Health Data Space for down- and upward data use
- **Prof. Dr. Meta Roestenberg**, LUMC, LU, FDP for scientific data of INNNO4VAC
- **Prof. Dr. Thomas van den Akker**, LUMC, LU, Seed money for ANC data integration
- **Dr. Steffie Hemelaar**, Integration sensitive patient data in four university hospitals
- **Dr. Stephanie van Rooden,** RIVM, PRAISE, Common Data Model for epidemiological surveillance
- **Prof. Dr. Lu Cao**, LIACS, LU, FAIRification of data for improved AI applications,
- **Dr. Louise Bezuidenhout,** Social Sciences, LU, Open Science
- **Prof. Dr. Ronald Cornet,** Amsterdam UMC, FAIR integration health data
- **Prof; dr. Katy Wolstencroft,** Amsterdam UMC, FAIR integration health data
- **Prof. Dr. Simcha Jong,** London University College, disruptive business development
- **Prof. Dr. Ruben Verborgh**, Universiteit van Gent, Federation of personal data
- **Dr. Beatriz Gonçalves Crisóstomo Esteves**, Universiteit van Gent, semantic FAIR Access control
- **Dr. Helen Oliver,** Postdoctoral Research Fellow, The ESPRESSO Project, Birkbeck, University of London

Memoranda of Understanding on exchange of expertise

- AU Pan African Parliament
- Africa University Network on FAIR Open Science with Tangaza University, Kenya, University of Nairobi, Kenya, Equator University of Science & Technology, Uganda, Great Zimbabwe University, Zimbabwe, Mekelle University, Ethiopia, University of Grand Bassa, Liberia.

Research group (status December 2025) (selected)

Prof. Dr. Mirjam van Reisen, Chair, Research Group FAIR Data Science

Ruduan Plug - Security of sensitive health data in FAIR architectures - PhD finalised
Putu Hadi Purnama Jati - Adoption of FAIR architectures for patient data - PhD manuscript ready
Samson Yohannes Amare - Deployment of an integrated FAIR Data infrastructure for ANC patient data in a Data Space - PhD manuscript ready
Rens Kievit - Data Space Access Control in FAIR sensitive data architectures - Dual PhD with the Dutch government
Abdulahi Kawu - Integration of mobile applications and adoption of FAIR data systems - PhD Researcher
Kai Smits - Sensitive research data analytics across federated stores - PhD candidate - manuscript ready
Joëlle Stocker - Sensitive data structuring and information development - PhD researcher
Bereket Godifay - Ontologies for culturally specific health practices in maternal health - PhD candidate
Tesfit Gebremeskel Gebreslassie - FHIR and FAIR integration for reuse of patient data - PhD candidate
Danial Zemchal - Common Data models for sensitive research data analytics in ANC across federated stores - PhD candidate
Liya Mamo - Common Data modeling health - Researcher VODAN Team
Philip Opiyo - Common Data Modeling - Researcher VODAN Team
James Ngoge - Common Data Models - Researcher VODAN Team
Kokob Gebru - Sexual violence and resilience for mental health - PhD candidate
Jimmy Ben Forry - Sexual violence and resilience for mental health - PhD candidate
Zhengyu Lin - Data Space FAIR Data architecture patient data integration - rcandidate
James Jacob - Federated Health informatics in a patient data Health Data Space - PhD candidate
Liam van Dreumel - Patient data visiting models for an RIVM health data space for infectious diseases - researcher
Mirte Balke - Inno4Vac Data Expert
Abdulahi Ibrahim - Computer Science engineer - VODAN Team
Steffie Hemelaar - Integration of sensitive patient data in four academic hospitals in The Netherlands - PostDoc researcher

Beatrix Callard - Integration of sensitive patient data across registrars - PostDoc researcher
Rita Mazocchi - sensitive data handling - PhD candidate - VODAN Team
I**lham Ibrahim** - sensitive data handling - VODAN Team
Onesmus Wanjiku - Installation FAIR Data Points
Abdulnazif Abdulkadir - Installation FAIR Data Points
Kbrom Kidu - Dashboard creation - VODAN Team
Maxwell Omare - Installation Health facilities - VODAN Team
Sakinat Folrunso - Curriculum development - VODAN Team
Julia Duncan-Cassel - Gender expertise - Senior researcher, VODAN Team
Reginald Nalugala - Health Data autonomy - Senior Researcher, VODAN Team
Obinna Osigwe - Communcations and internal communication VODAN Team
Yemane Seged - webmaster AHDS

Important links

AHDS: https://aun.mu.edu.et/ahds/

Manual;: https://aun.mu.edu.et/ahds/ahds-fair-senior-datastewardship-training-2025/

AUN-FOS: https://aun.mu.edu.et/

FAIR in Africa: https://aun.mu.edu.et/fair-in-africa/

VODAN: https://aun.mu.edu.et/vodan/