# FAIR Federated Data Space - A proposed Landscape

**Prof. Mirjam van Reisen**

21 December 2023

# The FAIR DataLandscape

| | | |
|---|---|---|
| 1 | Opening by Prof. Dr. Mpezamihigo, VC KIU & Chair VODAN-Africa | *VODAN-Africa proudly presenting the results of 2023* |
| 2 | Keynote address by Jeroen Maas, director Access to Care, Technology and Partnerships | *Towards the Africa Health Data Space* |
| 3 | Misha Stocker, Research Advisors & Experts Europe | *Towards the European Health Data Space* |
| 4 | Putu Hadi Purnama Jati, LUMC & VODAN-Africa | *GDPR & Privacy Preservation in the FAIR Patient Data production Pipeline* |
| 5 | Zhengyu Lin, LIACS | *The strength of De Novo FAIRification* |
| 6 | Ruduan Plug, TNO, LUMC & VODAN-Africa | *Federated surveillance* |
| 7 | Discussion | *Q &A* |

**Coffee break**

# VODAN-Africa 2023 Report

| | | |
|---|---|---|
| 8 | Samson Yohannes, Mekelle University, PhD LUMC | *Update 2023* |
| 9 | Julia Duncan-Cassell, Chair Global Data Stewarship Competence Centre | *Special Address -*<br>*Opening of the Ceremony of Certification* |
| 10 | Certification of Health Facilities | *Pronouncement of Certification of Health Facilities* |
| 11 | Certification of Data Clerks | *Pronouncement of the certified Data Clerks* |
| 12 | Certification of Data Stewards Level 4 | *Pronouncement of the certified Data Stewards* |
| 13 | Vote of Appreciation to all sponsors | *Thank you and close* |

**Lunch and discussion**

# Data Science in Practice - Fieldlabs

**DSIP 1** with GetuTadele & Samson Yohannes

*Development of a regional administration patient data dashboard for antenatal care monitoring based on federated algorithms*

**DSIP2** with Mildred Akandinda, Bwaga Ibrahim & Mariam Basajja

*Comparing DHIS and FAIR-Data for privacy preserving health analytics of patient data in Uganda*

**DSIP 3** with Samson Yohannes

*Pregnancy risk analytics using multicountry maternal health data*

**DISP 4** with Abdullahi Kawu & Rens Kievit

*Vascular disease risk analytics through interoperability of FAIR patient records with wearables*

**DSIP 5** with Lars Schrijver & Ria Landa

*Data interoperability of relevant information pertaining to elderly with dementia in The Netherlands*

**DSIP 6** with Kai Smits

*GraphGuard: Safeguarding Refugees by Analyzing Perpetrator Networks*

**DSIP 7** with Kai Smits

*SecuRePod: Personal data safety for refugees on the move created through permission control with a FAIR ontology*

**END**

# ADD PDFs

# List of certified clinics

- Mekelle General Hospital
- Ayder comprehensive specialized Hospital
- Olabisi Onabanjo University (OOU) Health Services Ago Iwoye Ogun State Nigeria
- General Hospital, Lapai
- Nigerian Railway Clinic
- Pumwani Maternity Hospital
- KEMRI
- Hoima referral hospital
- KIU Teaching Hospital
- Lira referral hospital
- Beacon of Hope Hospital

# List of certified data clerks

| Name | Institution |
|---|---|
| Kibrom Berhanu | Mekelle University,CHS,DHC |
| Kibrom G/Selassie | Chief Executive Director, MU, CHS |
| Liya Mamo | MU, Digital Health, CHS |
| Mphamedawel Mohamednigus | MU, Digital Health, CHS |
| Sara Bahta | Ayder Referral Hospital, MCH |
| Dr Filmon Mesfin | CEO, Mekelle Hospital |
| Simret Niguse | MD, Ayder Referral Hospital |
| Abreha G/her | CCD, Ayder Referral Hospital |
| Gebreamlak Gidey | MU CHS /Axum University |
| Ezekiel Adebayo Ogundepo | OOU and FUL |
| David Oziama Priscilla | FUL |

| Name | Institution |
|---|---|
| Adebanjo Kazeem Oluwatoyin | Olabisi Onabanjo University (OOU) Health Services Ago Iwoye Ogun State Nigeria |
| Sakinat Folorunso | Olabisi Onabanjo University (OOU) Health Services Ago Iwoye Ogun State Nigeria |
| Adamu Abubakar | |
| Ibrahim Bilyamin Ahmed | General Hospital, Lapai |
| Tajudeen Saka | Nigerian Railway Clinic Minna |
| Dr Ofudu Orobosa | Nigerian Railway Clinic Minna |
| Hauwa Idris | Nigerian Railway Clinic |
| Hafsat Zubairu | Nigerian Railway Clinic Minna |
| Zubairu Salamatu | Nigerian Railway Clinic Minna |
| Hussaina Muhammed Idris | Nigerian Railway Clinic Minna |
| Maimuna Saidu | Nigerian Railway Clinic Minna |

# List of certified data clerks

| Name | Institution |
|---|---|
| Awwal Zubairu | Nigerian Railway Clinic Minna |
| Muazu Yusuf | Nigerian Railway Clinic Minna |
| William Nandwa | Pumwani Maternity Hospital |
| David Aluodo | Pumwani Maternity Hospital |
| Jacinta Wairimu | Beacon of Hope Hospital |
| Dennis Murerwa Kinoti | Beacon of Hope Hospital |
| Lawrence Njoroge | Tangaza University ICT Manager |
| Seth Okeyo | KEMRI |
| Idrisa Muyingo | Hoima referral hospital |
| Enjamin Atugumye | KIU Teaching Hospital |
| Ongol Dennis | Lira referral hospital |

| Name | Institution |
|---|---|
| Rhoda Bello | Nigerian Railway Clinic Minna |
| Usman Musa | General Hospital, Lapai |
| ELIJAH, Joseph | General Hospital, Lapai |

# List of certified FAIR data stewards

## Level IV

Joëlle Stocker

Lars Schrijver

Rens Kievit

Misha Stocker

Kai Smits

Nathalia Morales Rojas

# The European Health Data Space

**Misha Stocker**

**Research Advisors & Experts Europe**

VODAN-Africa

21 December 2023


RAEE

# European Health Data Space (EHDS)

Is a key pillar of the **European Health Union**

Passed in the Council on 6 December 2023 & 13 December in the Parliament Plenary

Next Step: Negotiations between Parliament, Council, and Commission

# What does the European Health Data Space (EHDS) aim to do?

Primary use: Ensuring individuals control over electronic health data

Secondary use: Encourage the access to and use of health data for research

Creating a uniform legal framework for the Development and use of electronic health record systems

RAEE

VODAN AFRICA

# Primary Use of Data in de EHDS:

Ensure easy access to your own medical data free of charge

Enable the transfer and interoperability of health data to other member states

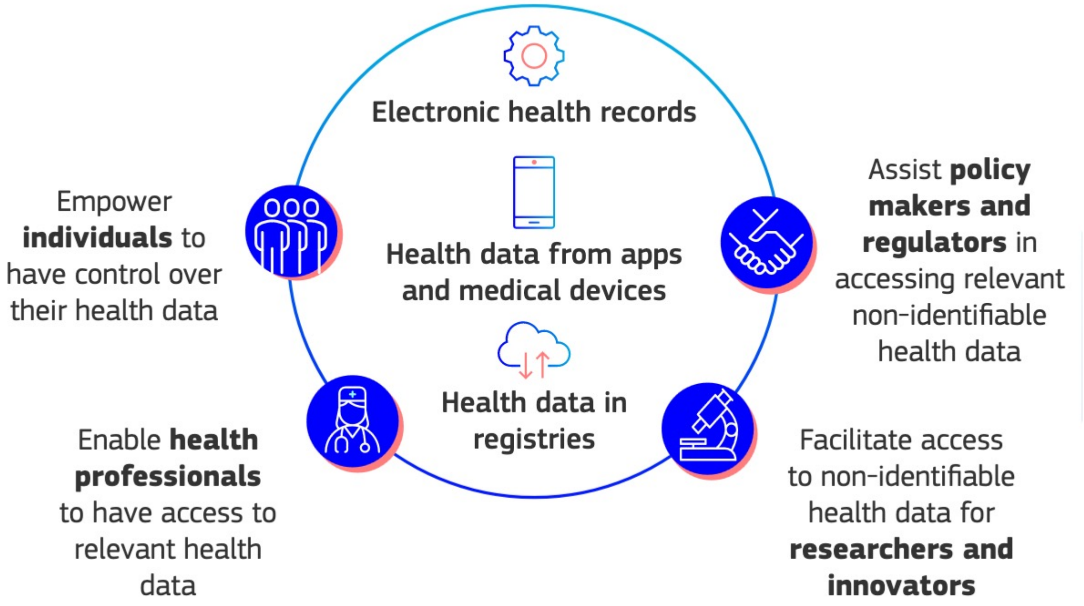
Establish Rights of Access for professionals

Establish a Digital Health Authority responsible as a contact poiint

Create uniform technical specifications

# Secondary use and interoperability - the EHDS:

- Establishes under which conditions, duties, and activities data can be used
- Provides the mechanisms for third countries to participate
- Establishes requirements EHR system operators need to adhere too
- Instructs the European Commission to establish requirements, technical specifications, the IT architecture
- Directs the European Commission to establish a EU Dataset Catalogue




RAEE

VODAN AFRICA

# Mentions of FAIR principles in the EU Council Position on the EHDS:

*"It also highlighted the need for sharing electronic health data that are findable, accessible, interoperable and reusable ('FAIR principles'), and ensuring that electronic health data are as open as possible and as closed as necessary. Synergies between the EHDS, the European Open Science Cloud4 and the European Research Infrastructures should be ensured" (p. 3, 16048/1/23)*

*"Improving the quality and utility of datasets through informed customer choice and harmonising related requirements at Union level, taking into account existing Union and international standards, guidelines, recommendations for data collection and data exchange (i.e. FAIR principles: Findable, Accessible, Interoperable and Reusable)" (p. 41, 16048/1/23)*

RAEE

VODAN AFRICA

# De Novo FAIRification

Zhengyu Lin
Supervisor: Mirjam van Reisen

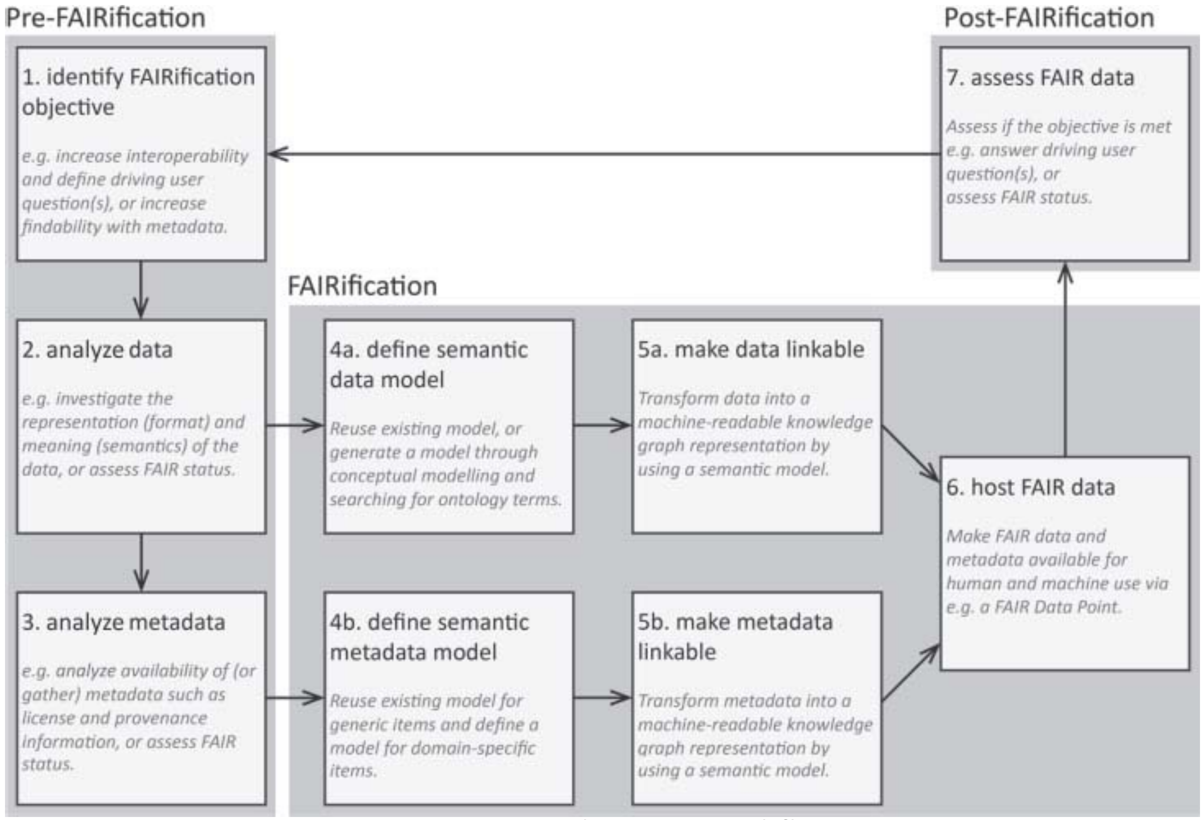
Universiteit Leiden
The Netherlands

# Generic FAIRification workflow

We make the already collected data FAIR.

And most workflows to FAIRify data are performed after data collection.

This way is post-hoc-FAIRification.



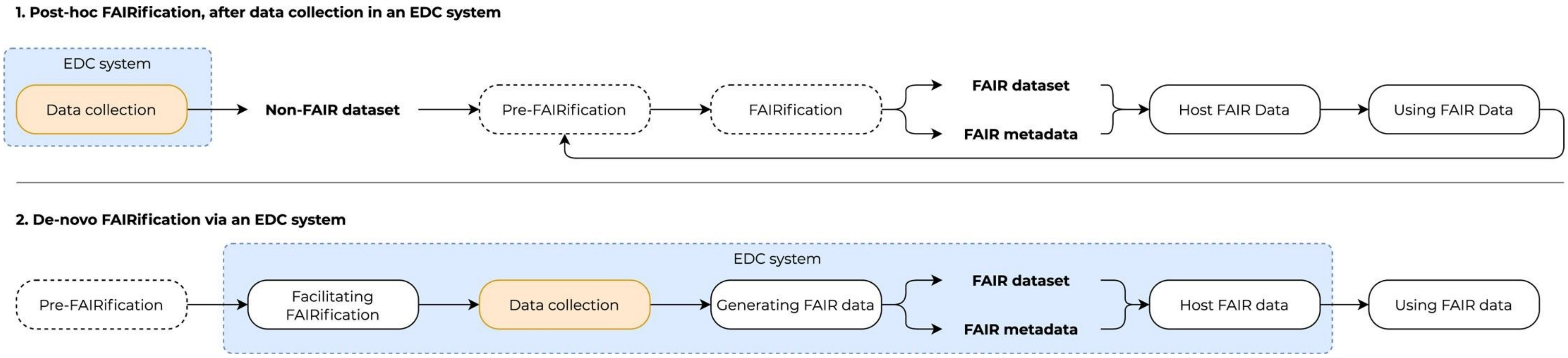

A generic step-by-step workflow

Think further, what about we do works before collecting data

*https://direct.mit.edu/dint/article/2/1-2/56/9988/A-Generic-Workflow-for-the-Data-FAIRification*
*https://www.sciencedirect.com/science/article/pii/S1532046421002264*

# What is De Novo FAIRification?

De novo FAIRification involves automating the process of making data FAIR in real-time during data collection.




The difference between post-hoc FAIRification and De novo FAIRification

Post-hoc FAIRification: No works before collecting data.

De novo FAIRification: By doing all the hands-on work for the FAIRification before data collection, data is made FAIR through entering them into an Electronic Data Capture (EDC) system.

*https://www.sciencedirect.com/science/article/pii/S1532046421002264*
*https://link.springer.com/article/10.1186/s13023-021-02004-y*

# What are the advantages?

“This mitigates the need for post-hoc FAIRification operations, which include repeated, semi-manual conversions of the data collected into machine-readable data that is performed after data collection.

The de novo approach saves time and budget for the actual FAIRification of the data in the VASCA registry. ”

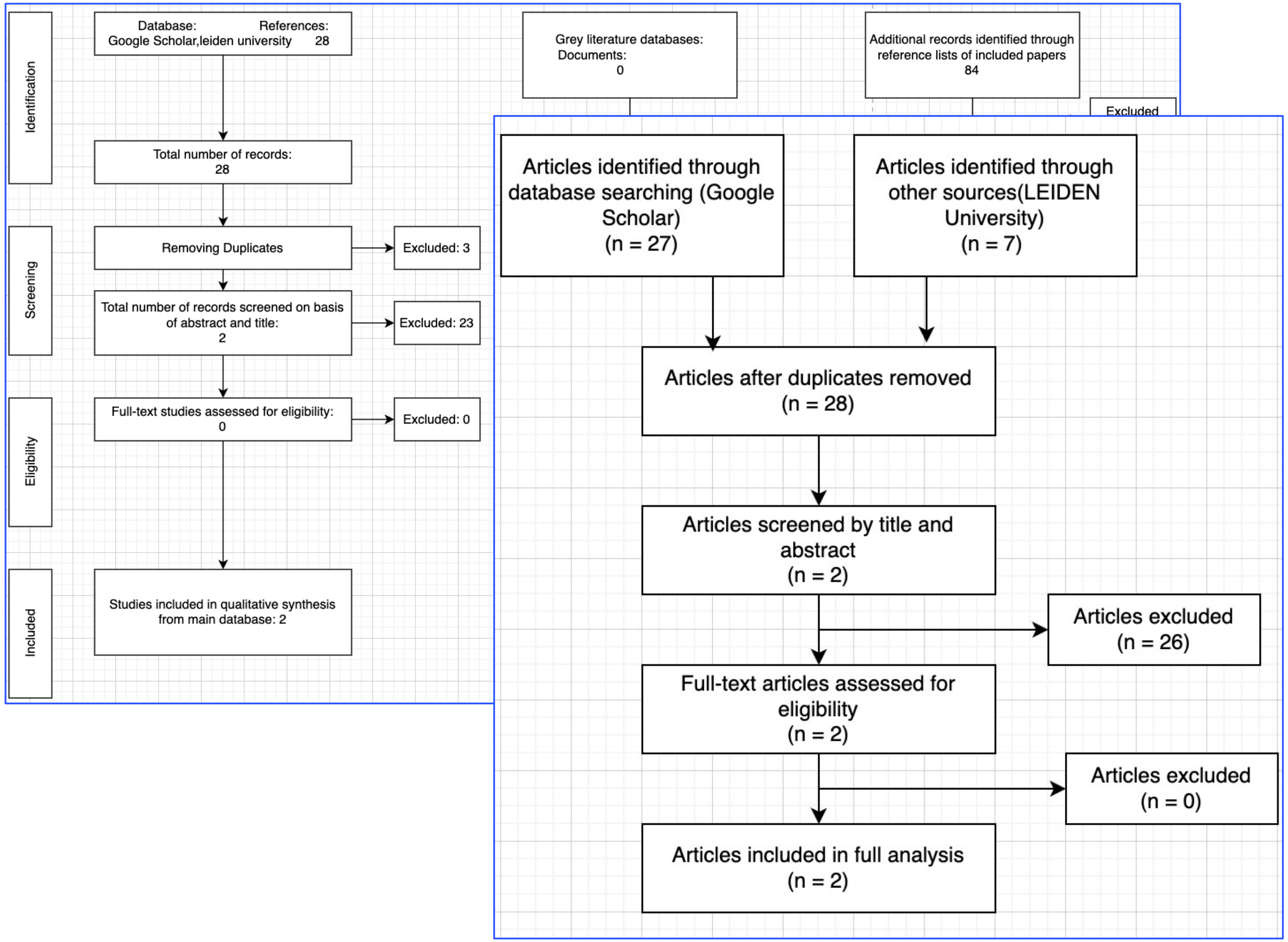
*https://link.springer.com/article/10.1186/s13023-021-02004-y*

# What is the research gap

The first project that successfully adopted De novo FAIRification was in 2021.

After that, there have been no subsequent projects utilizing De novo FAIRification.

Therefore an investigation is required to understand barriers to De novo FAIRification in order to define the key steps towards the widespread implementation of De novo FAIRification in research practices.



VODAN AFRICA

# VODAN-AFRICA PERSONAL DATA PROTECTION AND DATA PRODUCTION OPPORTUNITIES AND CHALLENGES

Putu Hadi Purnama Jati
December 2023

# Capacity Building Feedback

**2020**

Data, access, and control concept
https://doi.org/10.1162/dint_a_00180

**2021**

- Preparation survey
- Agreement on template
- Data Processing Agreement

**2022**

- Privacy Assessment
- GDPR Report

**2023**

- Challenge and Opportunities
  Data Pipeline
- The changes of domestic regulation concerning provisions for patient data and personal data protection

**2024**

Understanding data privacy preserving and data protection provisions among the key stakeholders in VODAN-Africa

# Opportunities and Challenges to Data Production

## Phase

- a survey by the data stewards to determine the extent of the technical problems experienced. T
- determine the different dimensions and nature of the problems experienced.

## Participants

Twenty (20) Country Coordinators and Data Stewards from Nigeria, Kenya, Ethiopia, Somalia, and Uganda, with each session lasting between 45 to 60 minutes


## Analysis

Employed thematic analysis and Critical Incident Theory in interviews to enhance discussions and uncover detailed insights into reported technical issues.

## Sample

Purposeful sampling was used to select health facilities, to enable us to obtain responsive data

VODAN AFRICA

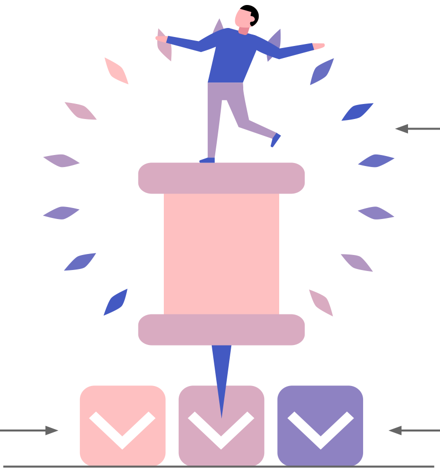
# Capacity Building Feedback

## Appreciation

Participants value capacity building in using the VODAN-Africa data production system.

## Rating

Participants rate support effectiveness ranges is from adequate to good



## Challenge

Challenges in funding for training and data production have been overcome by the own contributions of the health facilities

## Needs

A call for increased support for training

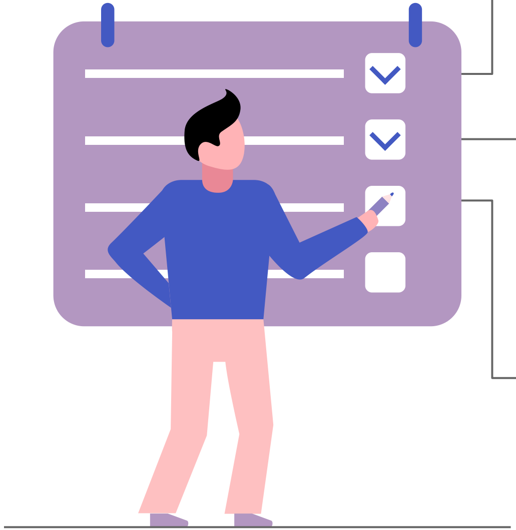
# Infrastructure

| | Challenge | Solution |
| --- | --- | --- |
| 01 | **Internet**<br>Challenges in Internet Connectivity: Participants highlighted major issues in accessing the internet, particularly in remote areas. | Local health facility contribution, but remote area challenges necessitate national infrastructure solutions because it is related to electricity. |
| 02 | **Electricity**<br>Widespread electricity shortages in various regions posed significant obstacles. The difficulty in accessing electricity led to solutions such as initially collecting data on paper before entering it into the platform. | Some facilities have resolved the issue of electricity even though it is not a permanent resolution such as use electricity backup (UPS) |

# System

| | Challenge | Solution |
| --- | --- | --- |
| 03 | **System's Memory Limitation**<br>Participants highlighted concerns regarding data loss when facing memory space issues. The system's lacks of temporary local memory causing difficulties in data entry when the internet is down or after prolonged usage. | Regularly clear computer memory to optimize data entry. |
| 04 | **Documentation**<br>The participant think the system needs proper documentation to provide support for the users. | Participant system proficiency increases with practice but needs improved documentation for quicker understanding |
| 05 | **Speed**<br>Many participants mentioned that the system is sometimes slow when it comes to booting to impute data. | No interim solution yet for system loading delays. |

# Suggestions for Future Improvements



| | | |
|---|---|---|
| 01 | **User Interface** | Participants emphasize the need for enhancing the user interface |
| 02 | **Customization** | Participants recommend adding customization features to the system |
| 03 | **Ease of use** | Participants recommend improvement in user experience |

Learning from

# Distributed Data

Ruduan Plug

21/12/2023

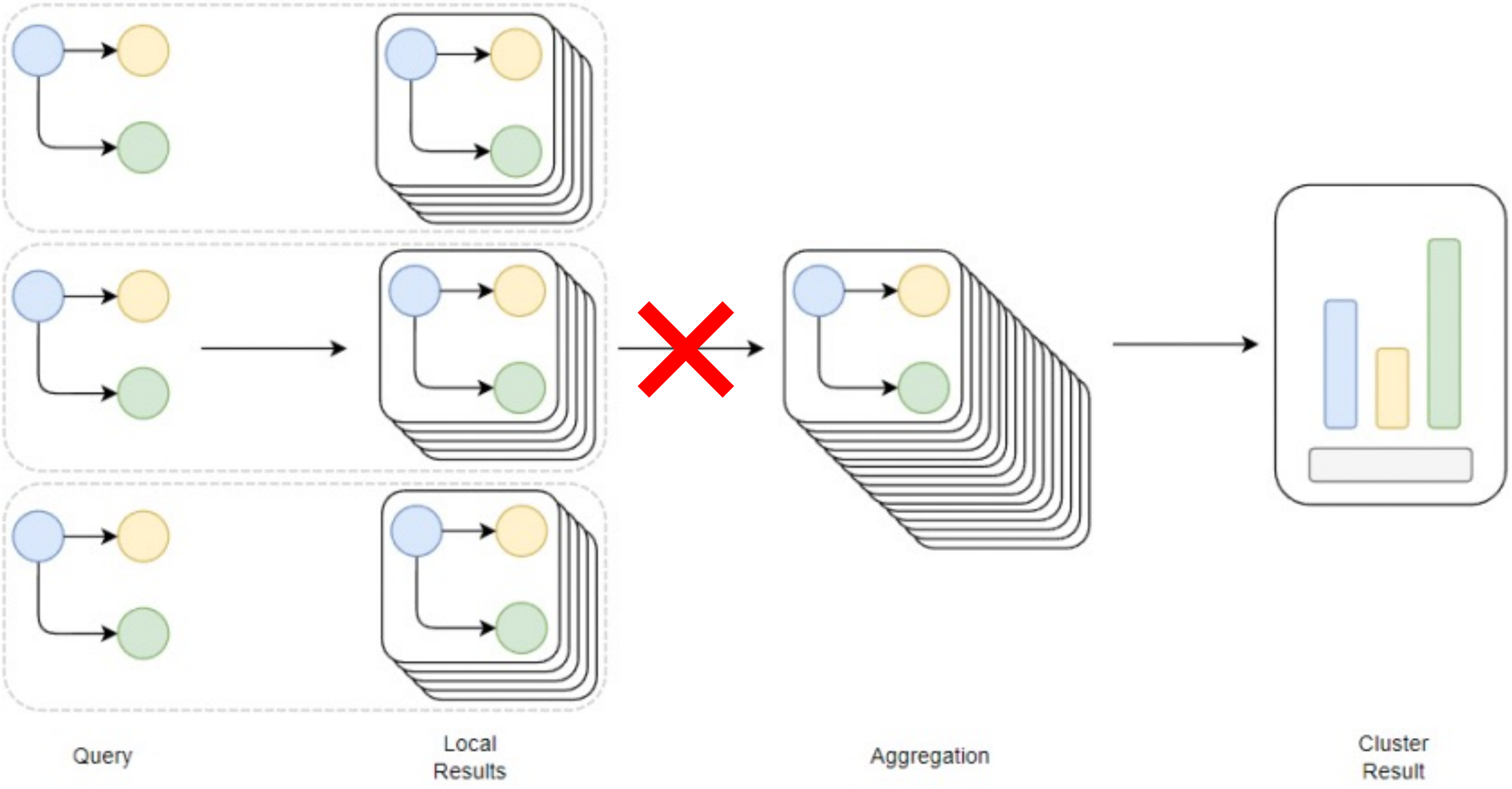
Leiden University

Advancements in
**Evidence Based Science**

Utilization of
**Mixed Data**

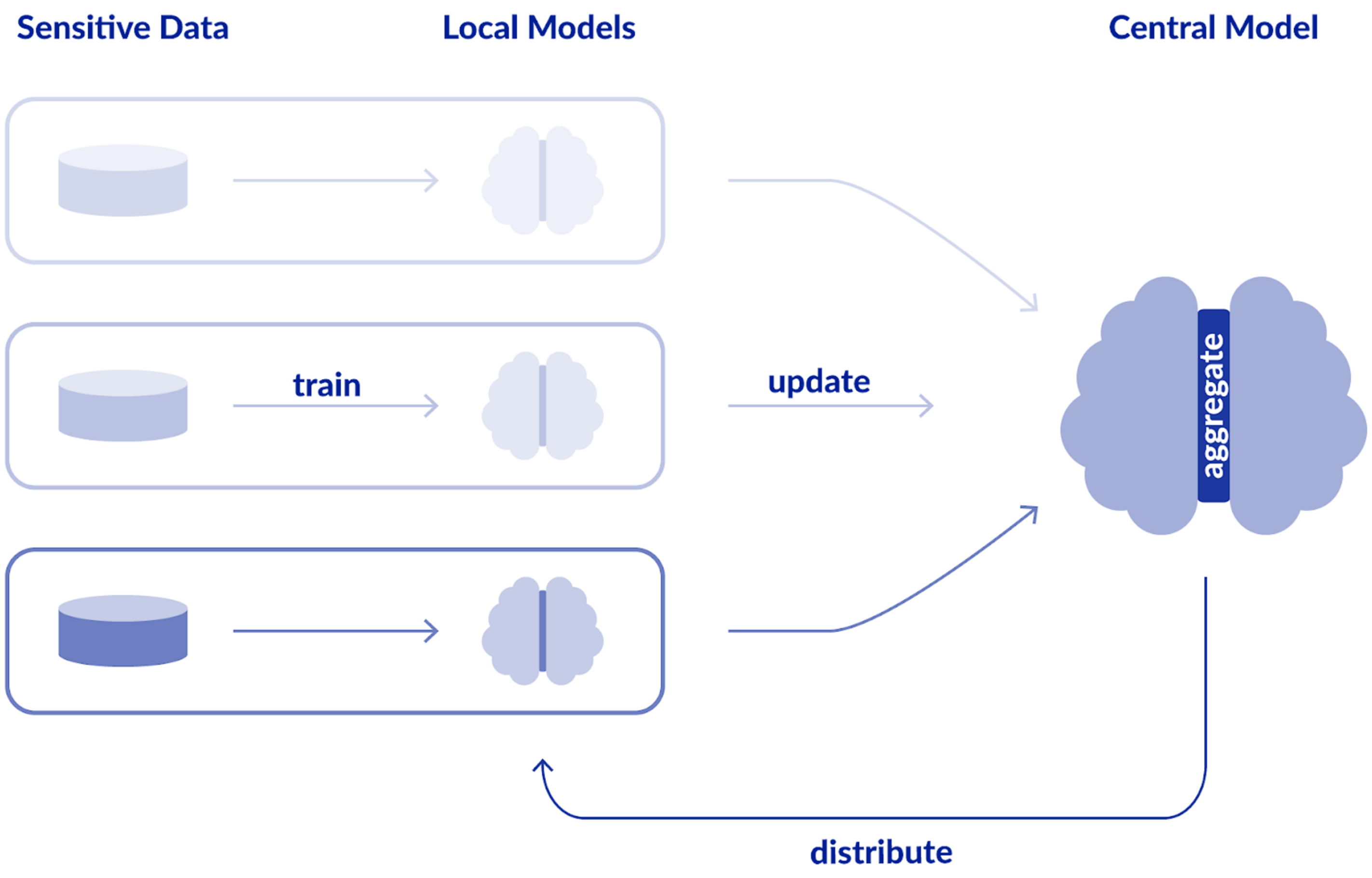
Evolving
**Regulatory Framework**

# Distributed Analysis



# Federated Analysis

# Federated Learning



# Flower: A Friendly Federated Learning Framework

Daniel J. Beutel [1 2] Taner Topal [1 2] Akhil Mathur [3] Xinchi Qiu [1] Javier Fernandez-Marques [4] Yan Gao [1]
Lorenzo Sani [5] Kwing Hei Li [1] Titouan Parcollet [6] Pedro Porto Buarque de Gusmão [1] Nicholas D. Lane [1]



## Abstract

Federated Learning (FL) has emerged as a promising technique for edge devices to collaboratively learn a shared prediction model, while keeping their training data on the device, thereby decoupling the ability to do machine learning from the need to store the data in the cloud. However, FL is difficult to implement realistically, both in terms of scale and systems heterogeneity. Although there are a number of research frameworks available to simulate FL algorithms, they do not support the study of scalable FL workloads on heterogeneous edge devices.

In this paper, we present Flower – a comprehensive FL framework that distinguishes itself from existing platforms by offering new facilities to execute large-scale FL experiments, and consider richly heterogeneous FL device scenarios. Our experiments show Flower can perform FL experiments up to *15M in client size* using only a pair of high-end GPUs. Researchers can then seamlessly migrate experiments to real devices to examine other parts of the design space. We believe Flower provides the community a critical new tool for FL study and development.

## 1 Introduction

There has been tremendous progress in enabling the execution of deep learning models on mobile and embedded devices to infer user contexts and behaviors (Fromm et al., 2018; Chowdhery et al., 2019; Malekzadeh et al., 2019; Lee et al., 2019; Yao et al., 2019; LiKamWa et al., 2016; Georgiev et al., 2017). This has been powered by the increasing computational abilities of mobile devices as well as novel algorithms which apply software optimizations to enable pre-trained cloud-scale models to run on resource-constrained devices. However, when it comes to the training of these mobile-focused models, a working assumption has been that the models will be trained centrally in the cloud, using training data aggregated from several users.

Federated Learning (FL) (McMahan et al., 2017) is an emerging area of research in the machine learning community which aims to enable distributed edge devices (or users) to collaboratively *train* a shared prediction model while keeping their personal data private. At a high level, this is achieved by repeating three basic steps: i) local parameters update to a shared prediction model on each edge device, ii) sending the local parameter updates to a central server for aggregation, and iii) receiving the aggregated model back for the next round of local updates.

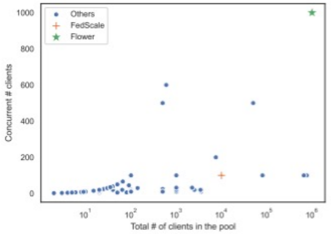
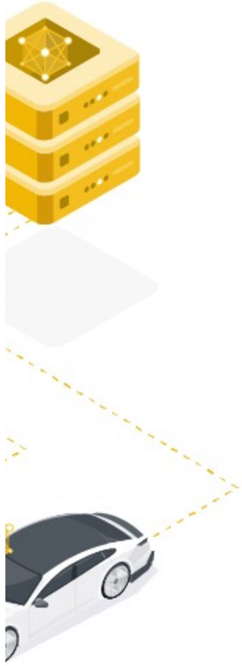
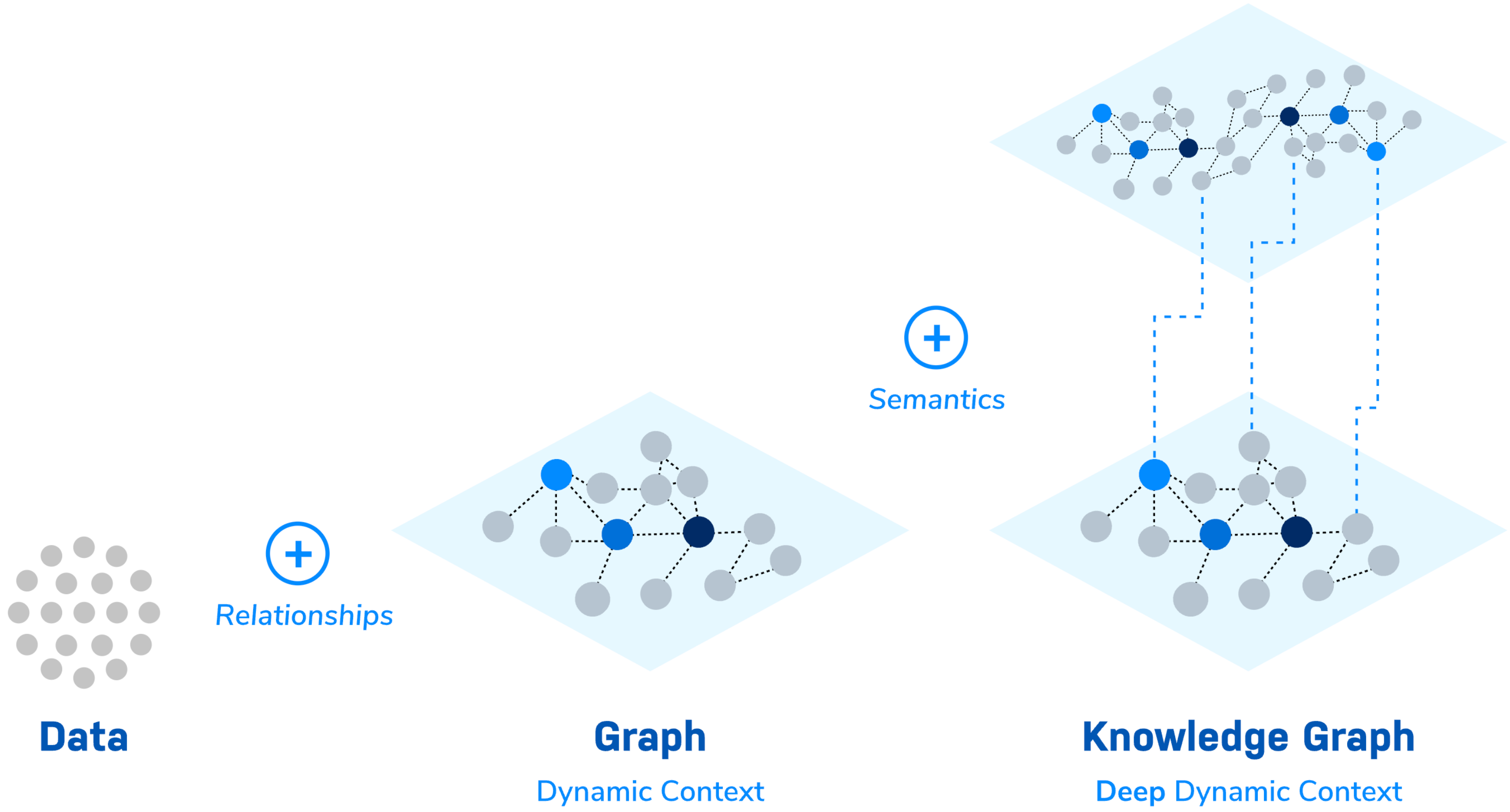



*Figure 1.* Survey of the number of FL clients used in FL research papers in the last two years. Scatter plot of number of concurrent clients participated in each communication round (y-axis) and total number of clients in the client pool (x-axis). The x-axis is converted to log scale to reflect the data points more clearly. FedScale can achieve 100 concurrent clients participated in each round out of 10000 total clients (orange point), while Flower framework can achieve 1000 concurrent clients out of a total 1 million clients (green point). The plot shows that Flower can achieve both higher concurrent participated client and larger client pool compared with other experiments existing the the recent research papers. Appendix A.1 gives details of the papers considered.

From a systems perspective, a major bottleneck to FL research is the paucity of frameworks that support scalable execution of FL methods on mobile and edge devices. While several frameworks including Tensorflow Federated (Google, 2020; Abadi et al., 2016a) (TFF) and LEAF (Caldas et al., 2018) enable experimentation on FL algorithms, they do not provide support for running FL on

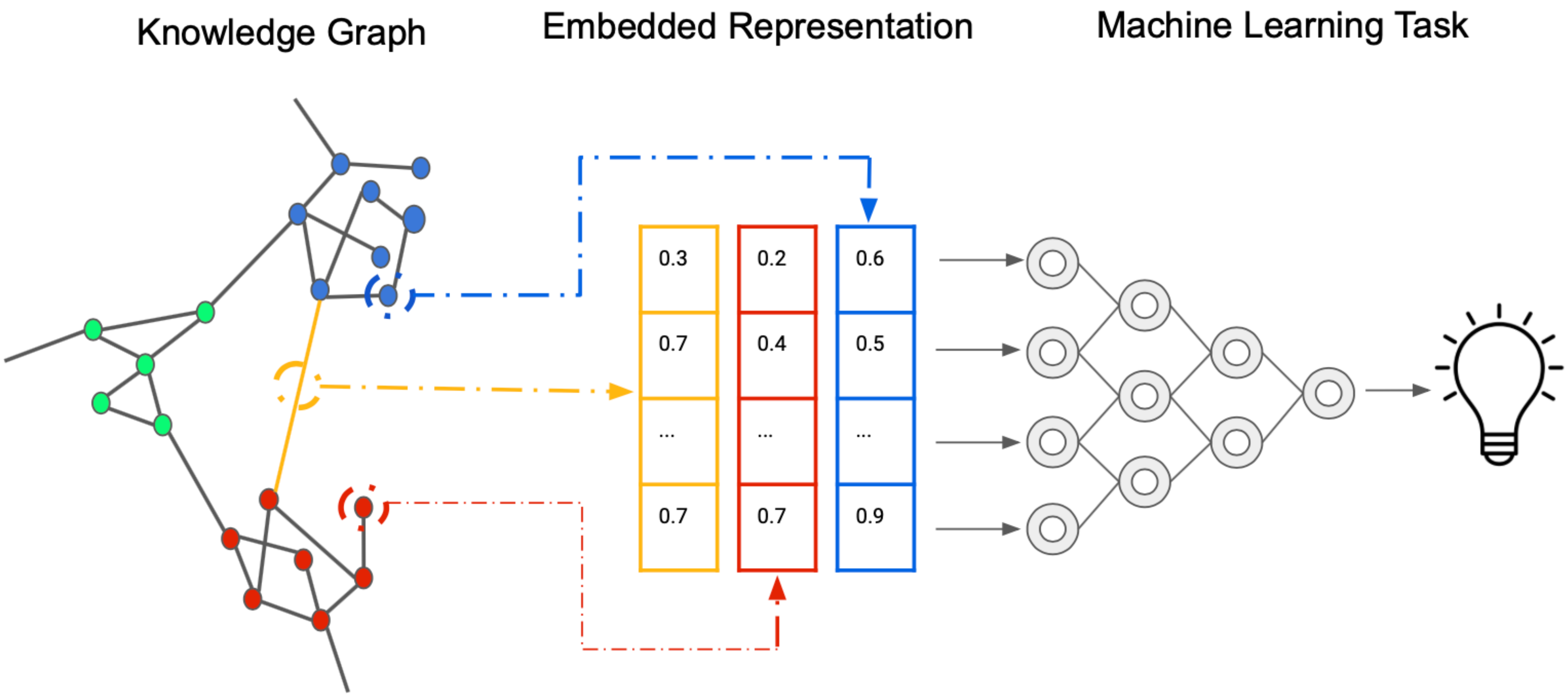
[1]Department of Computer Science and Technology, University of Cambridge, UK [2]Adap, Hamburg, Hamburg, Germany [3]Nokia Bell Labs, Cambridge, UK [4]Department of Computer Science, University of Oxford, UK [5]Department of Physics and Astronomy, University of Bologna, Italy [6]Laboratoire Informatique d'Avignon, Avignon Université, France. Correspondence to: Daniel J. Beutel <daniel@adap.com>.

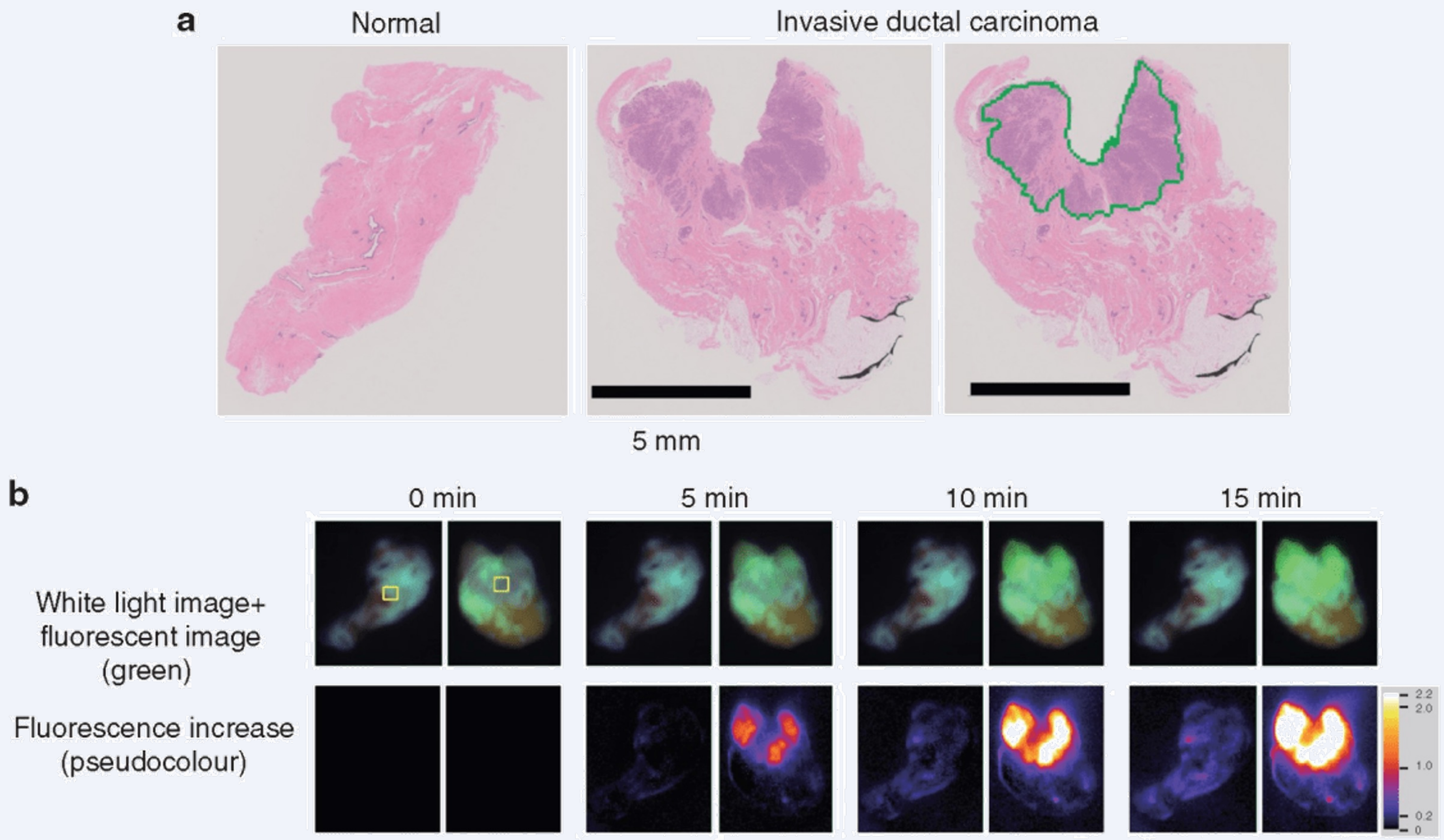





Source: Neo4J

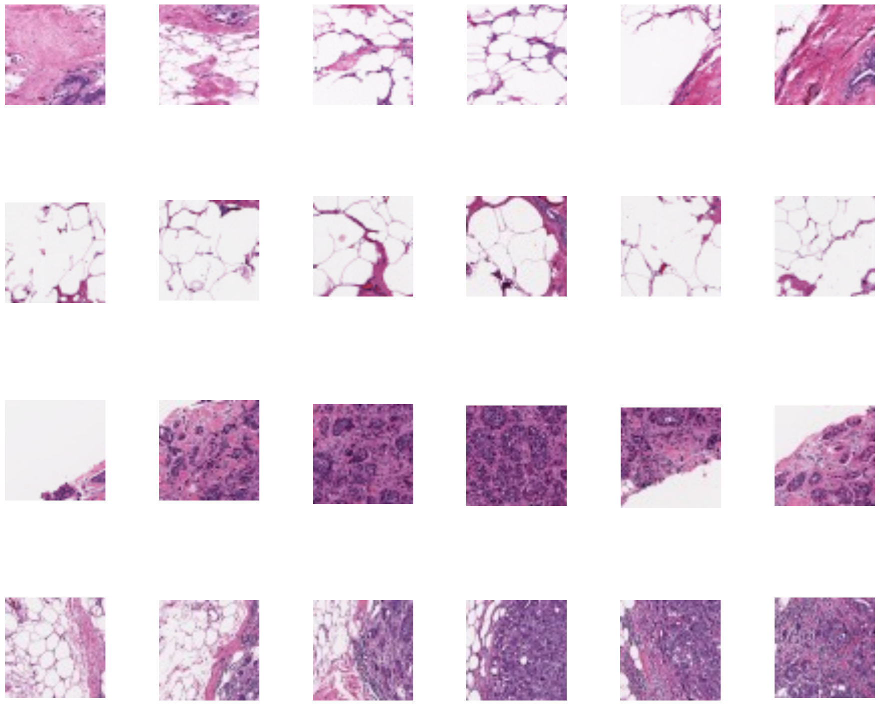
# Graph Machine Learning



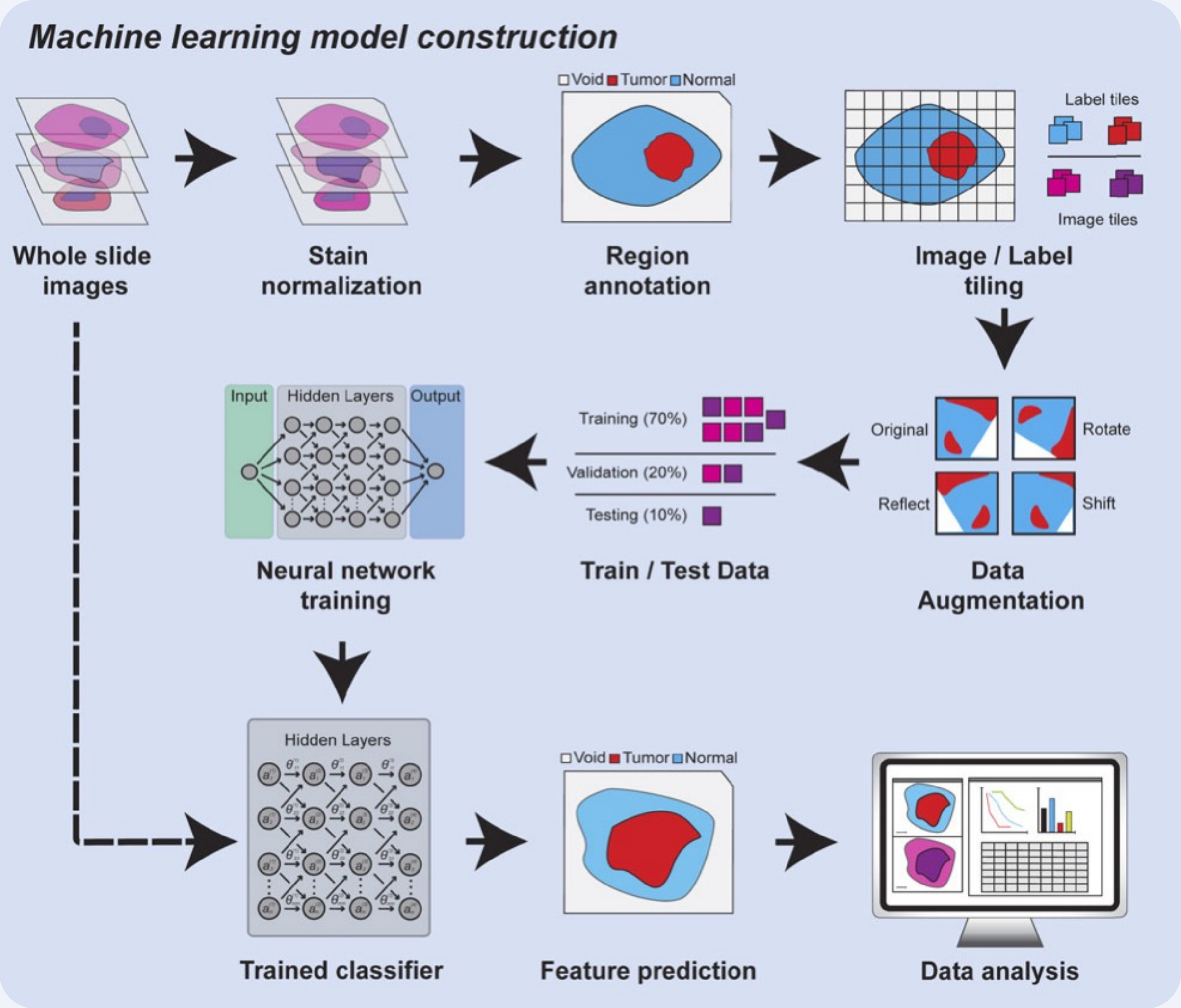



Source: Duke Pathology

EMORY
WINSHIP CANCER INSTITUTE
National Cancer Institute-Designated Comprehensive Cancer Center

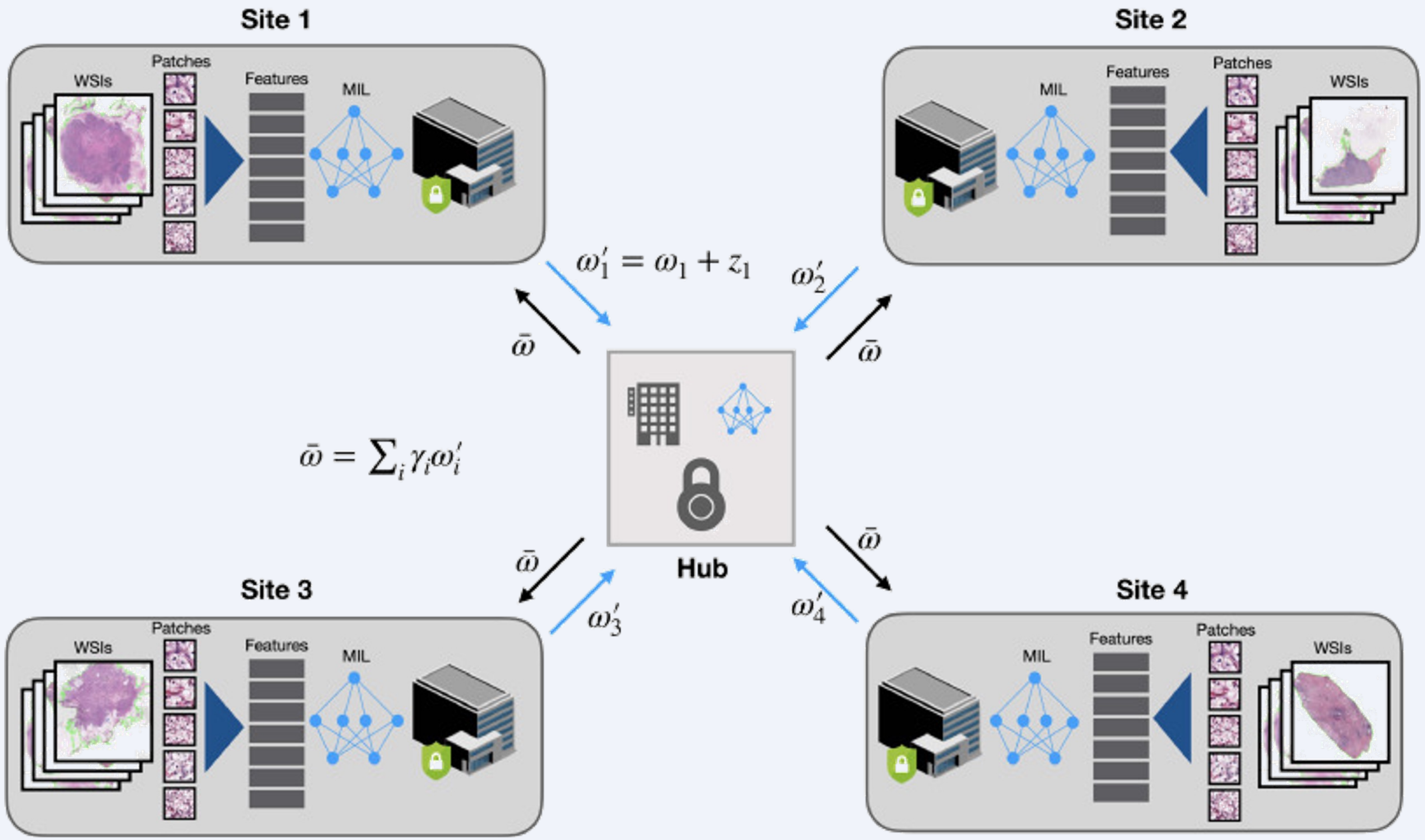


# Methodology



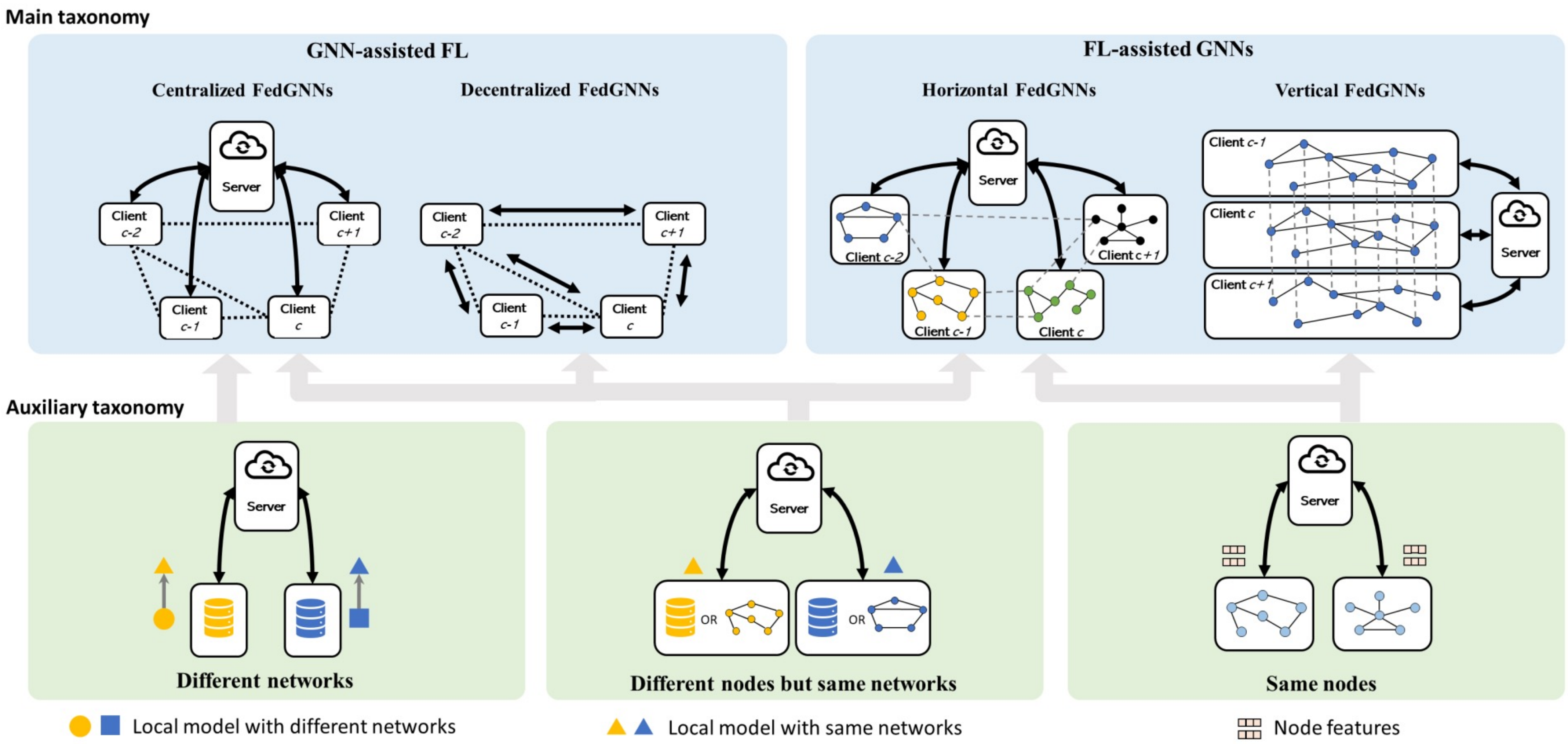

Lee, K., Lockhart, J.H., Xie, M., Chaudhary, R., Slebos, R.C., Flores, E.R., Chung, C.H., & Tan, A.C. (2021). Deep Learning of Histopathology Images at the Single Cell Level. Frontiers in Artificial Intelligence, 4.

# Federated Machine Learning



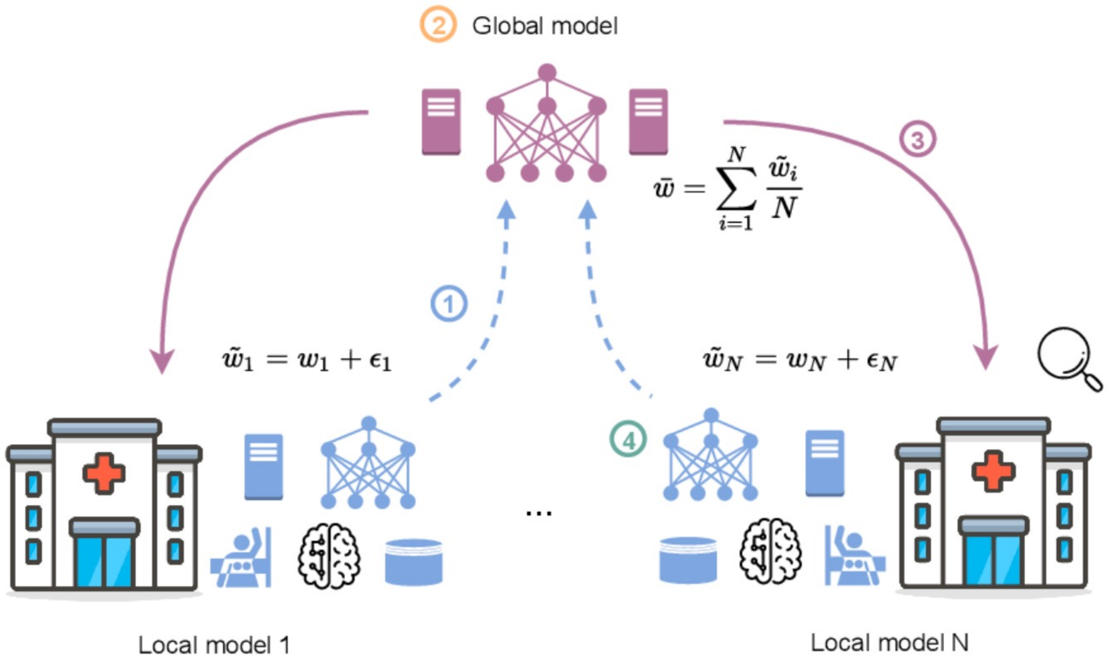

Lu, M.Y., Kong, D., Lipková, J., Chen, R.J., Singh, R., Williamson, D.F., Chen, T.Y., & Mahmood, F. (2020).
Federated learning for computational pathology on gigapixel whole slide images. Medical image analysis, 76, 102298 - 102298.

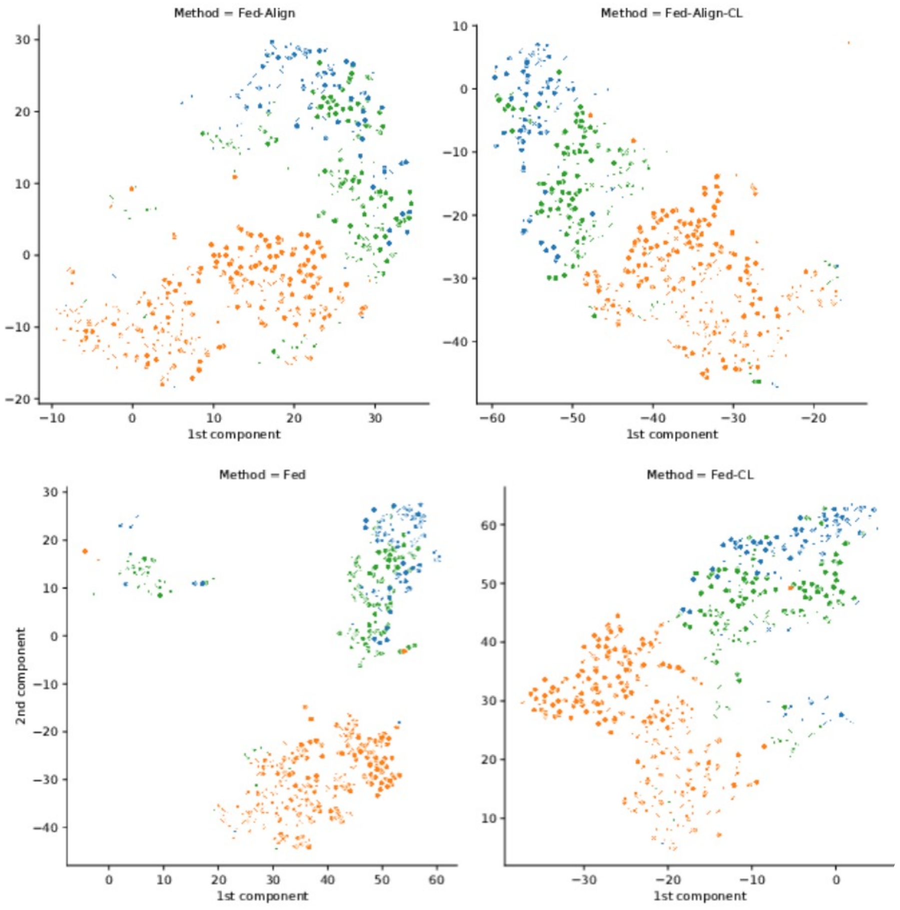
# Federated Graph ML




Liu, R., & Yu, H. (2022). Federated Graph Neural Networks: Overview, Techniques and Challenges. ArXiv, abs/2202.07256.

# Latent Space



# Concluding

# Annual Report VODAN-Africa 2023

VODAN-Africa Team

# OUTLINE

**01**
Architecture

**02**
Data Production

**03**
Business Case

**04**
Data stewardship

**05**
Governance

**06**
Way forward

01

# Architecture

*Samson Yohannes*

# Federated Health Data Space – a proposed landscape

## Generic connective capabilities - users

Ecosystem APIs

Intelligence | User Experience

## Infrastructure services

Data storage and hosting

Hosting & operations | Hybrid in location & local cloud

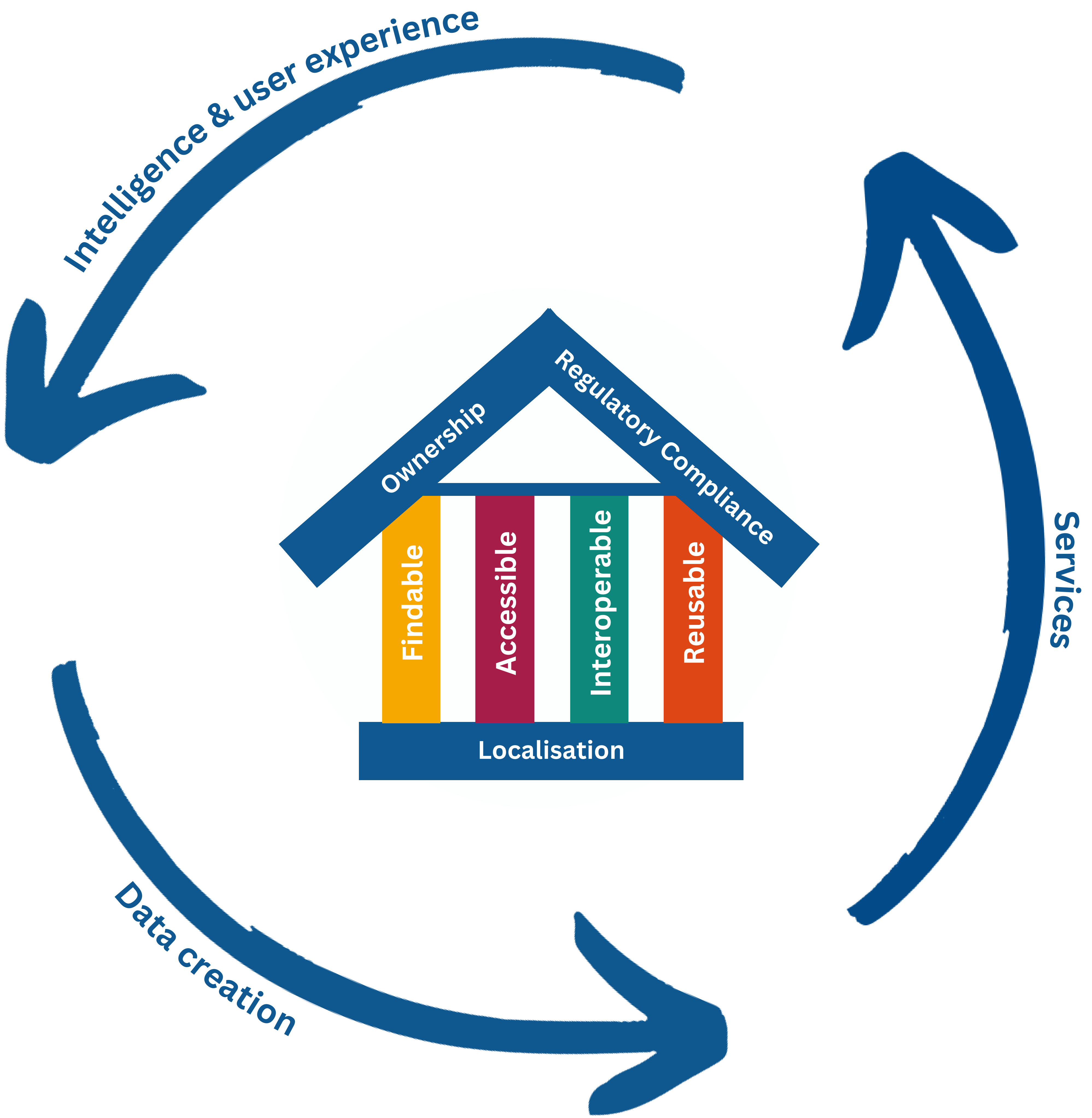
## Generic Data Capabilities - services

Federated data management

Regulatory Compliance, Privacy & Protection | Data Visiting & interoperability | Digital IT & Customer service

## Data layer-single copy, machine actionable data creation

Machine-actionable and semantically linked clinical, operational & research data at point of creation
(produced for instance : in point of care, point of service, or at research data collection)




# Realizations

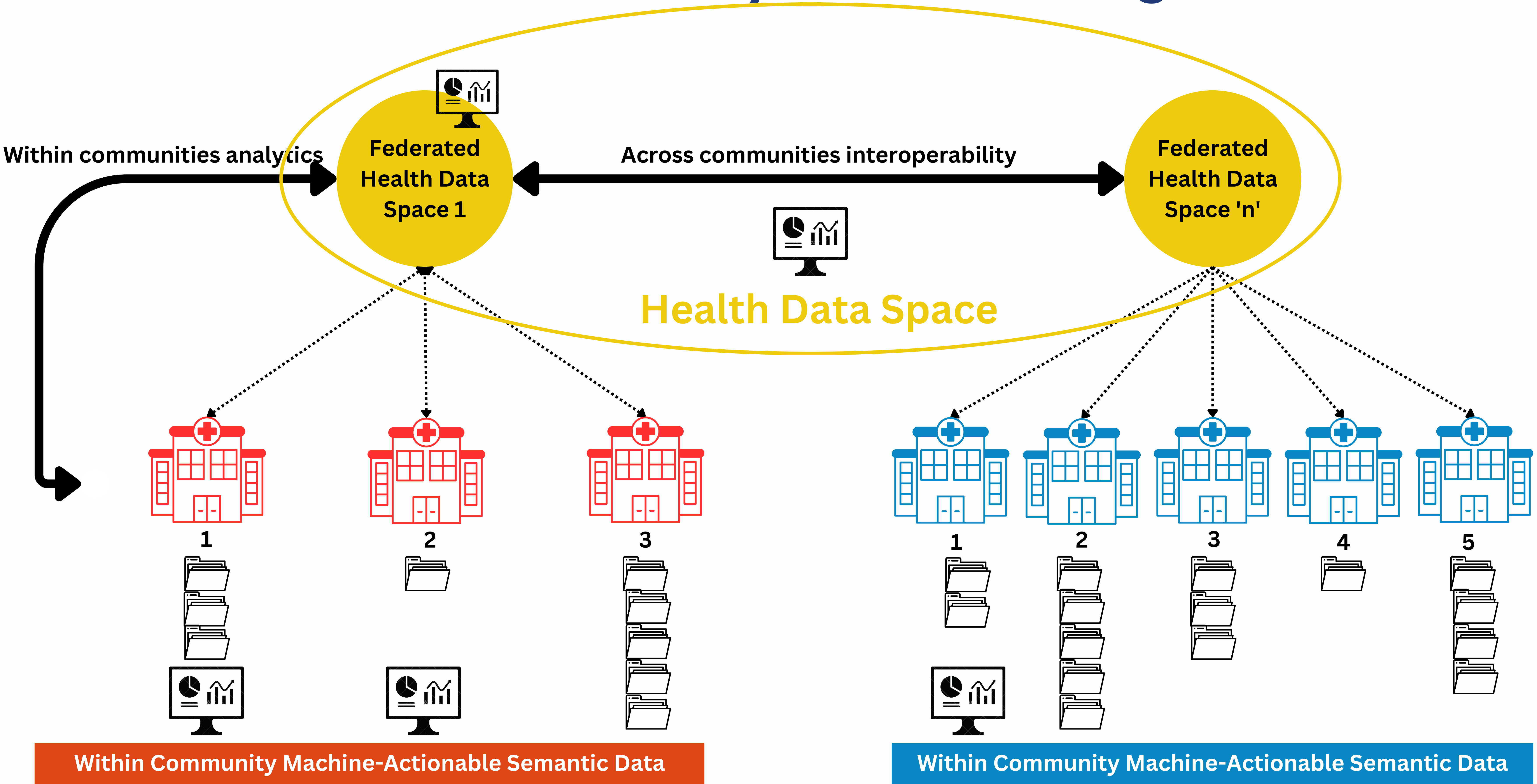
Data Production

Federated query

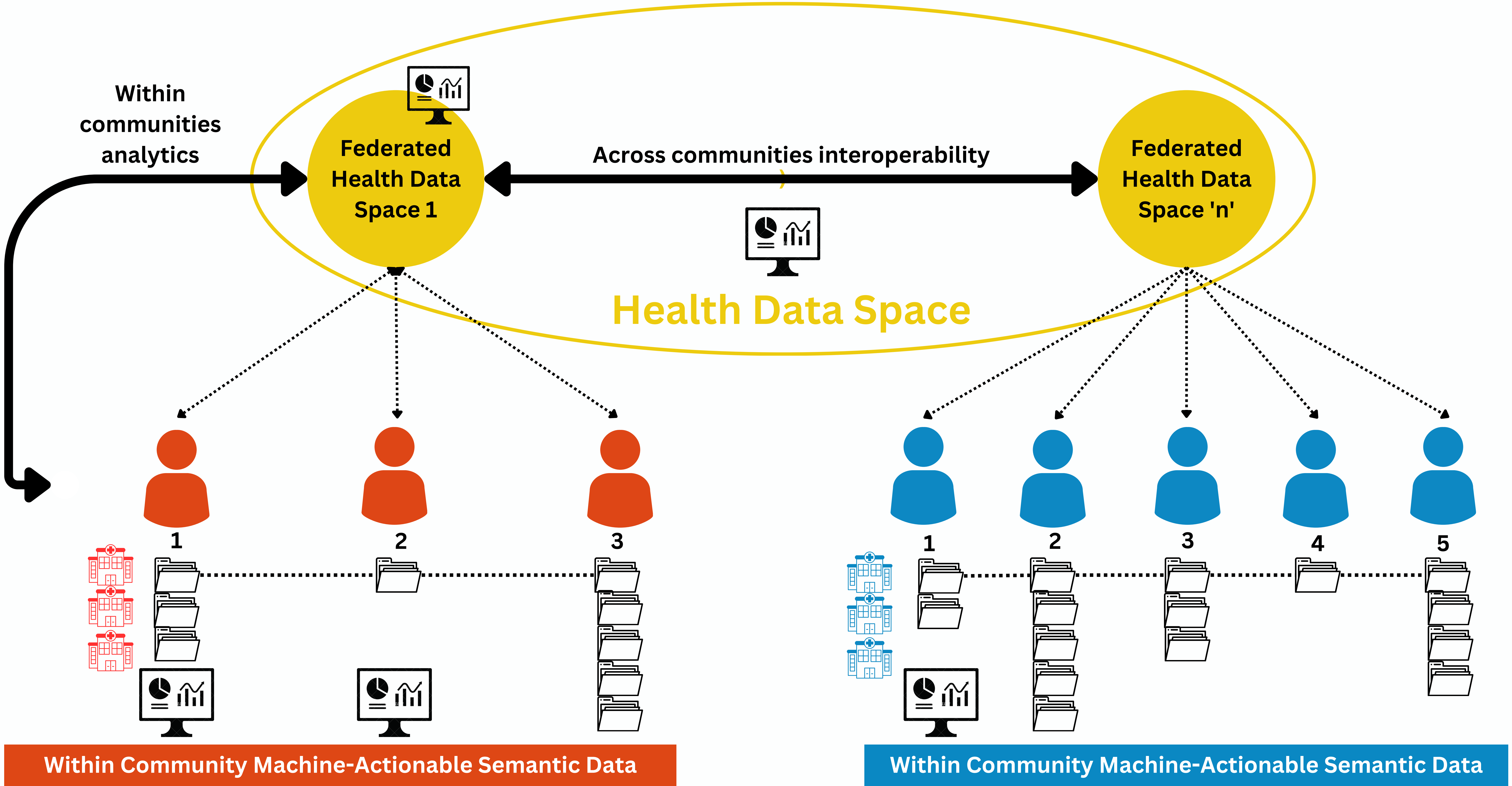
Regulatory compliance

Data stewardship

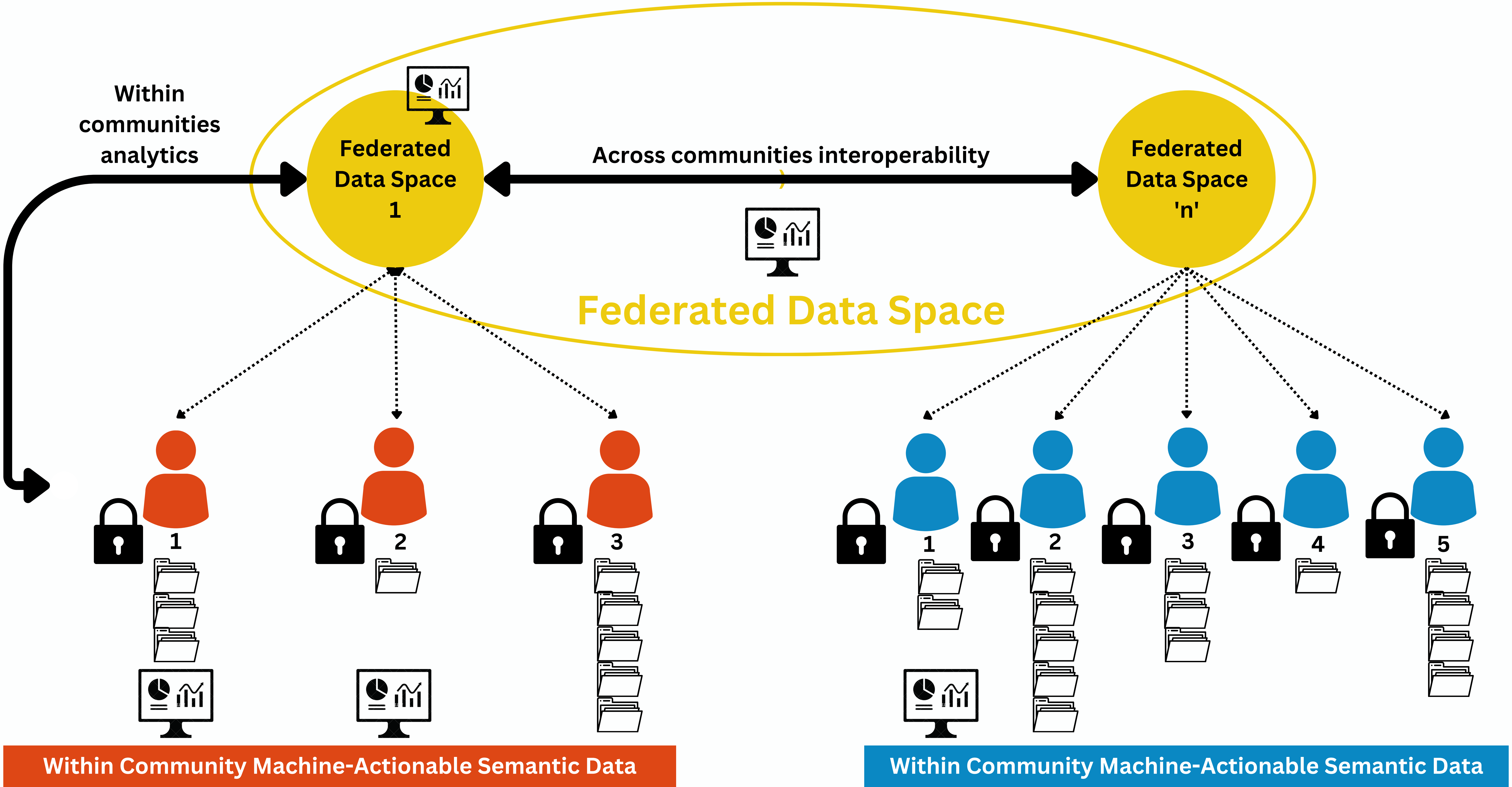
# FAIR-OLR Federated analysis & learning



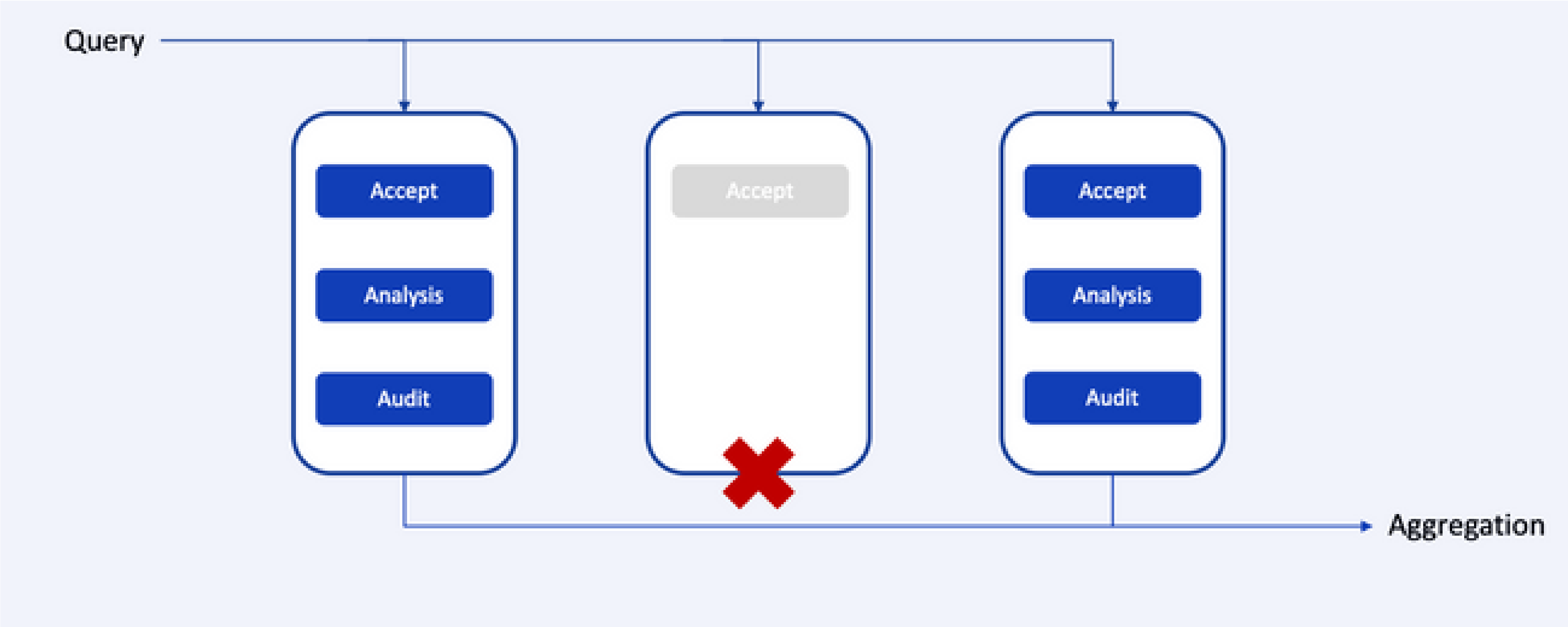
# FAIR-OLR Federated Personal Pods



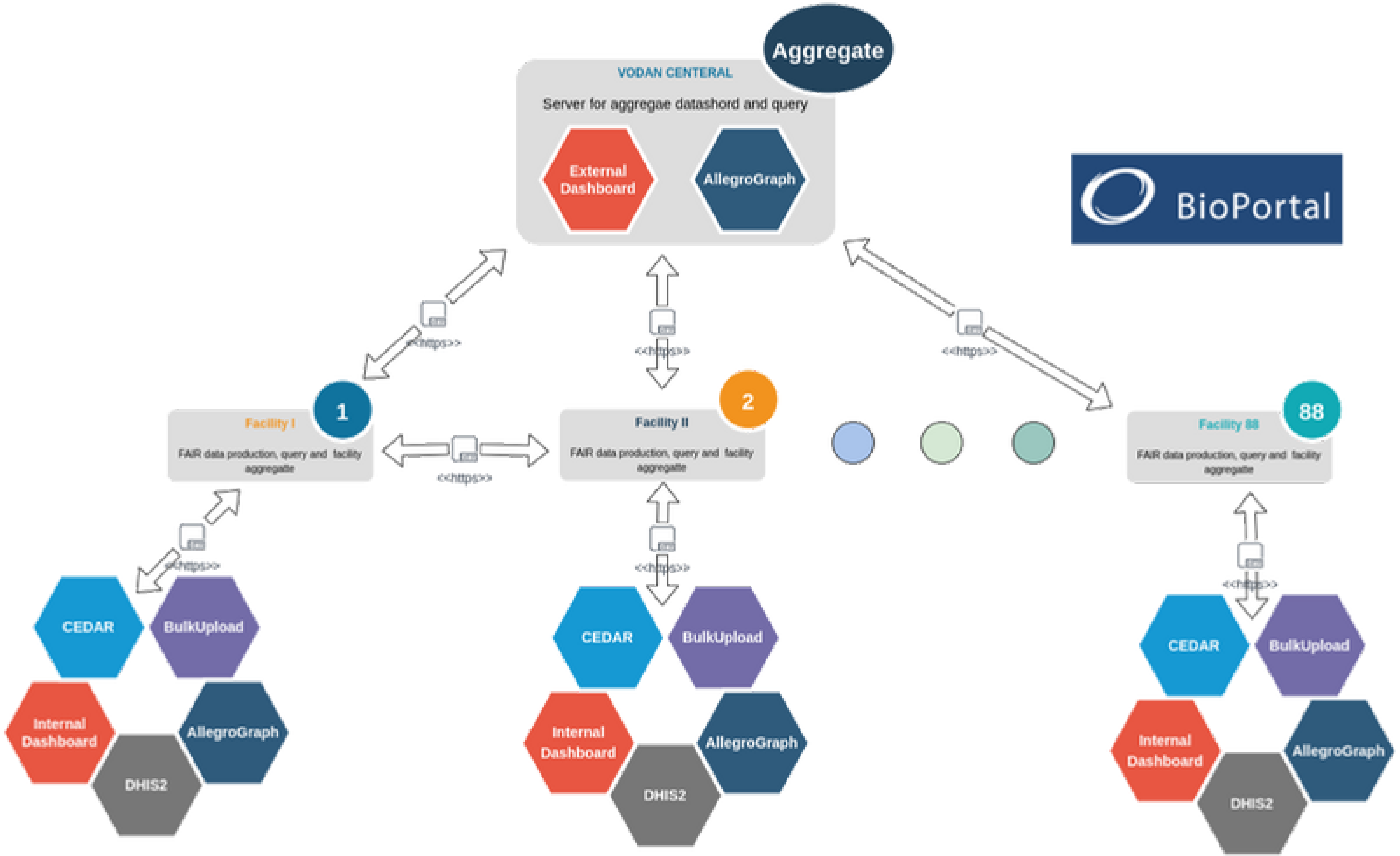
# FAIR-OLR Federated Personal Pods



# FAIR Federated Analysis



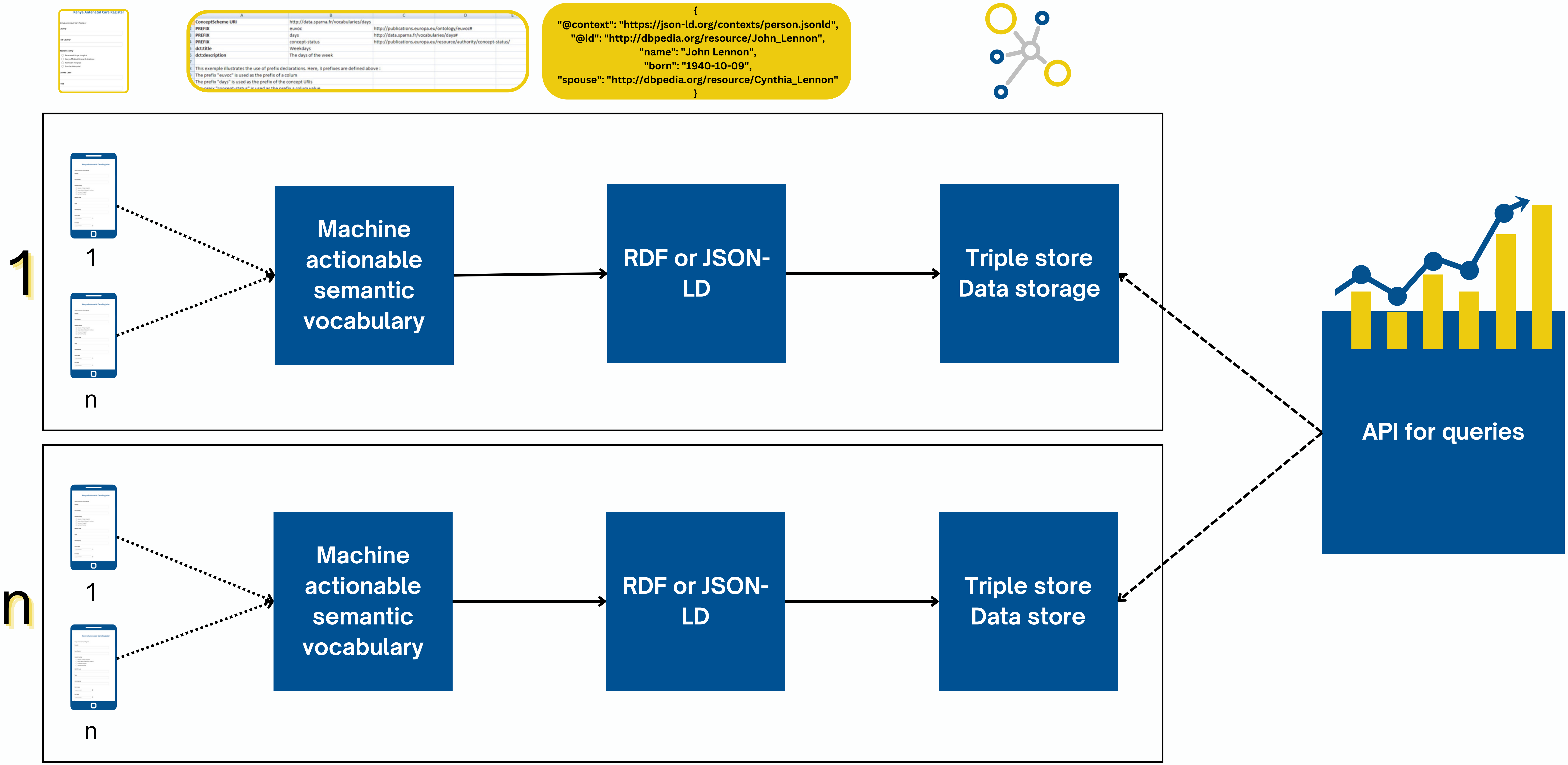
# FAIR Data Health Facilities: FAIR Software Infrastructure



# 02

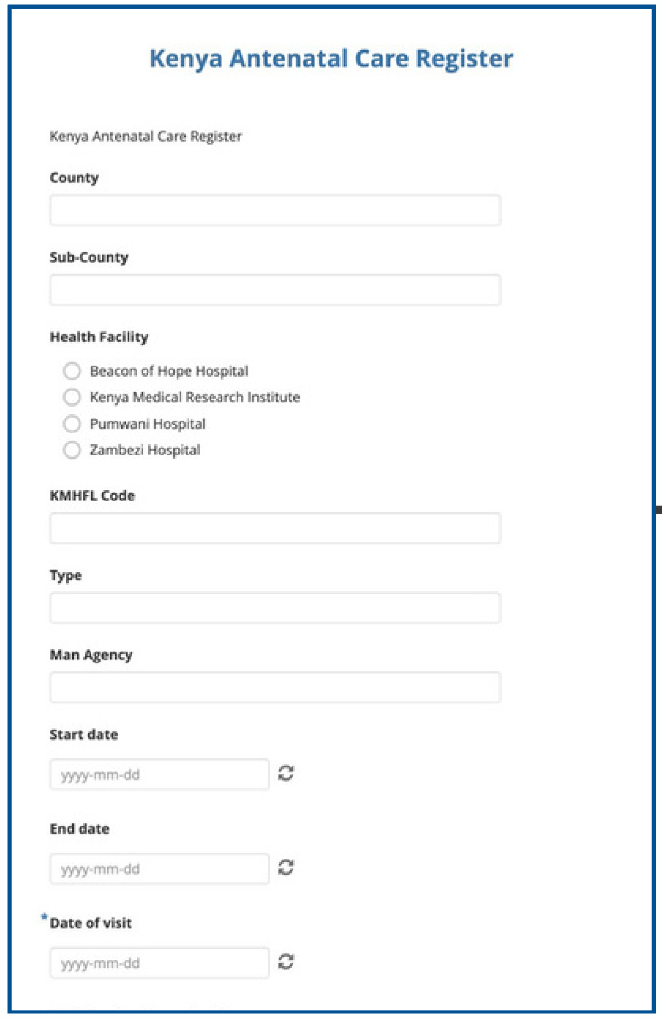
# Federated Data Production for knowledge and learning

Samson Yohannes




# FAIR Data production and processing

# Data creation process




Vocabulary does not exist

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | **ConceptScheme URI** | http://data.sparna.fr/vocabularies/days | | | |
| 2 | **PREFIX** | euvoc | http://publications.europa.eu/ontology/euvoc# | | |
| 3 | **PREFIX** | days | http://data.sparna.fr/vocabularies/days# | | |
| 4 | **PREFIX** | concept-status | http://publications.europa.eu/resource/authority/concept-status/ | | |
| 5 | **dct:title** | Weekdays | | | |
| 6 | **dct:description** | The days of the week | | | |
| 7 | | | | | |
| 8 | This exemple illustrates the use of prefix declarations. Here, 3 prefixes are defined above : | | | | |
| 9 | The prefix "euvoc" is used as the prefix of a colum | | | | |
| 10 | The prefix "days" is used as the prefix of the concept URIs | | | | |
| 11 | The preix "concept-status" is used as the prefix a colum value | | | | |

Vocabulary exists

## SKOS

**SKOS PLAY:**
to test and verify a vocabulary during the conception phase
to exchange and communicate the vocabulary when validating it with domain experts
to publish it when it is shared on the web.

## OWL

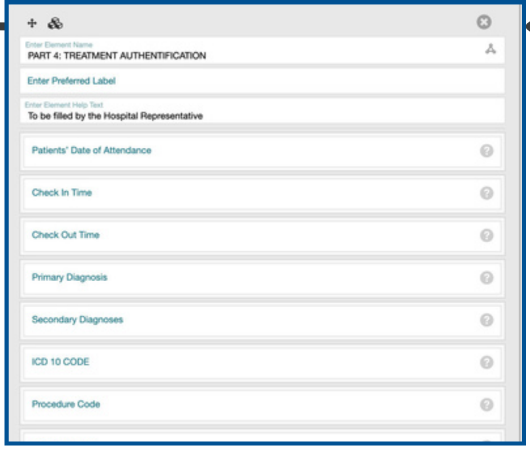
**Open Web Language**
RDF or JSON LD

## Bioportal

**Ontology repository**

**Link:** https://bioportal.bioontology.org/

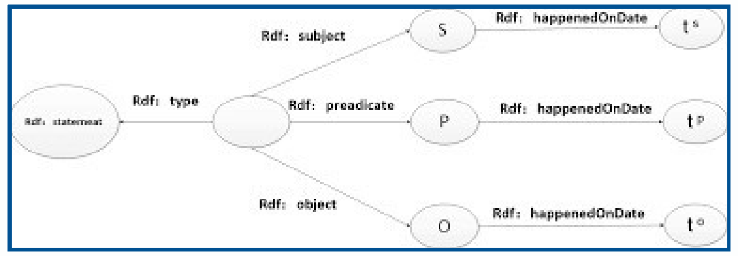
## CEDAR template




**Link:** https://more.metadatacenter.org/tools-training/cedar-template-tools

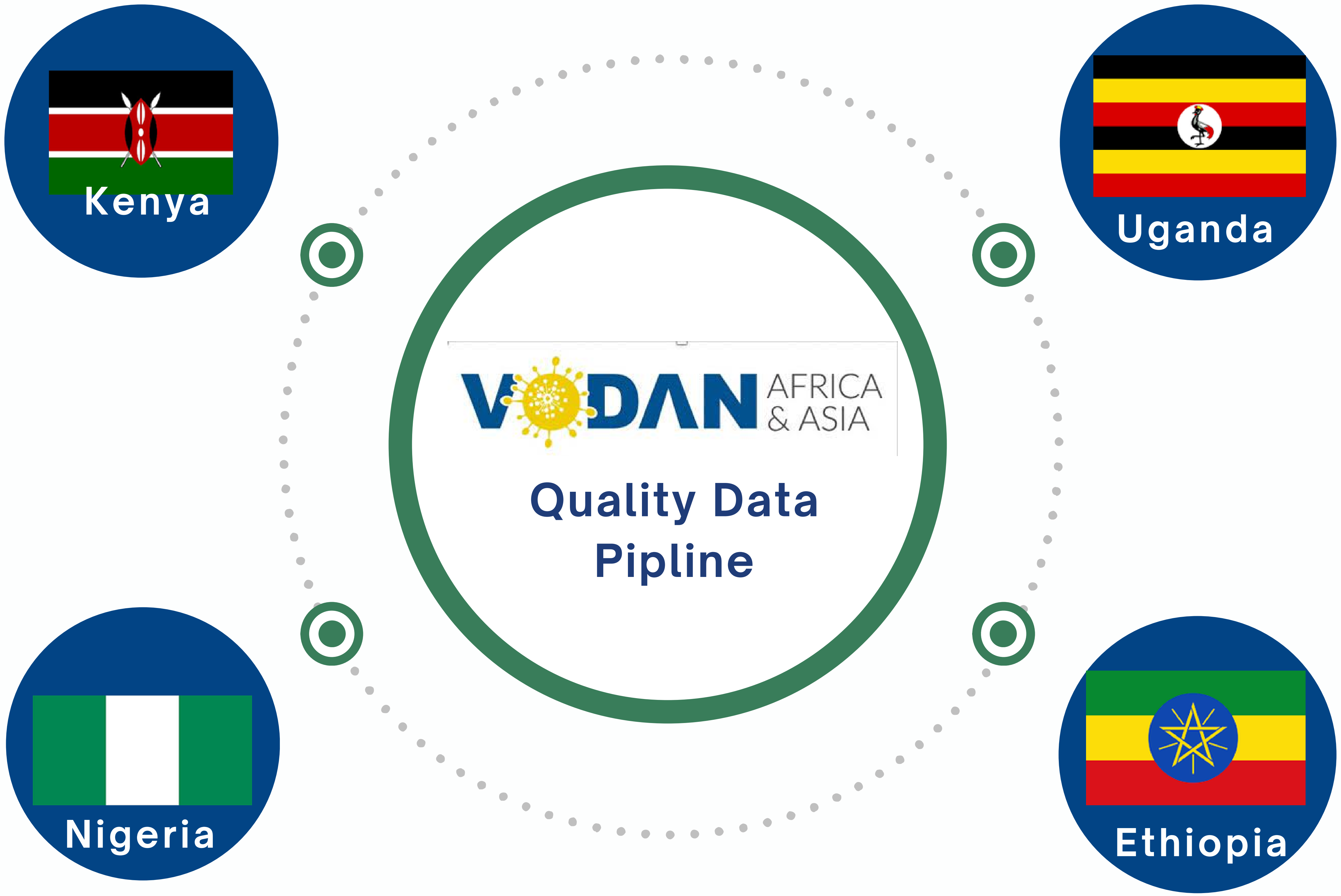
## Triple store

Storage and data analytics




**Source**: Haixia Li and Li Yan. 2021. A Temporal RDF Model for Multi-grained Time Information Modeling. In 2021 4th International Conference on Data Science and Information Technology (DSIT 2021), July 23-25, 2021, Shanghai, China. ACM, New York, NY, USA, 9 Pages. https://doi-org.tilburguniversity.idm.oclc.org/10.1145/3478905.3478908

# Countries enrolled quality data pipeline - 2023



# Quality Data Pipeline - 2023

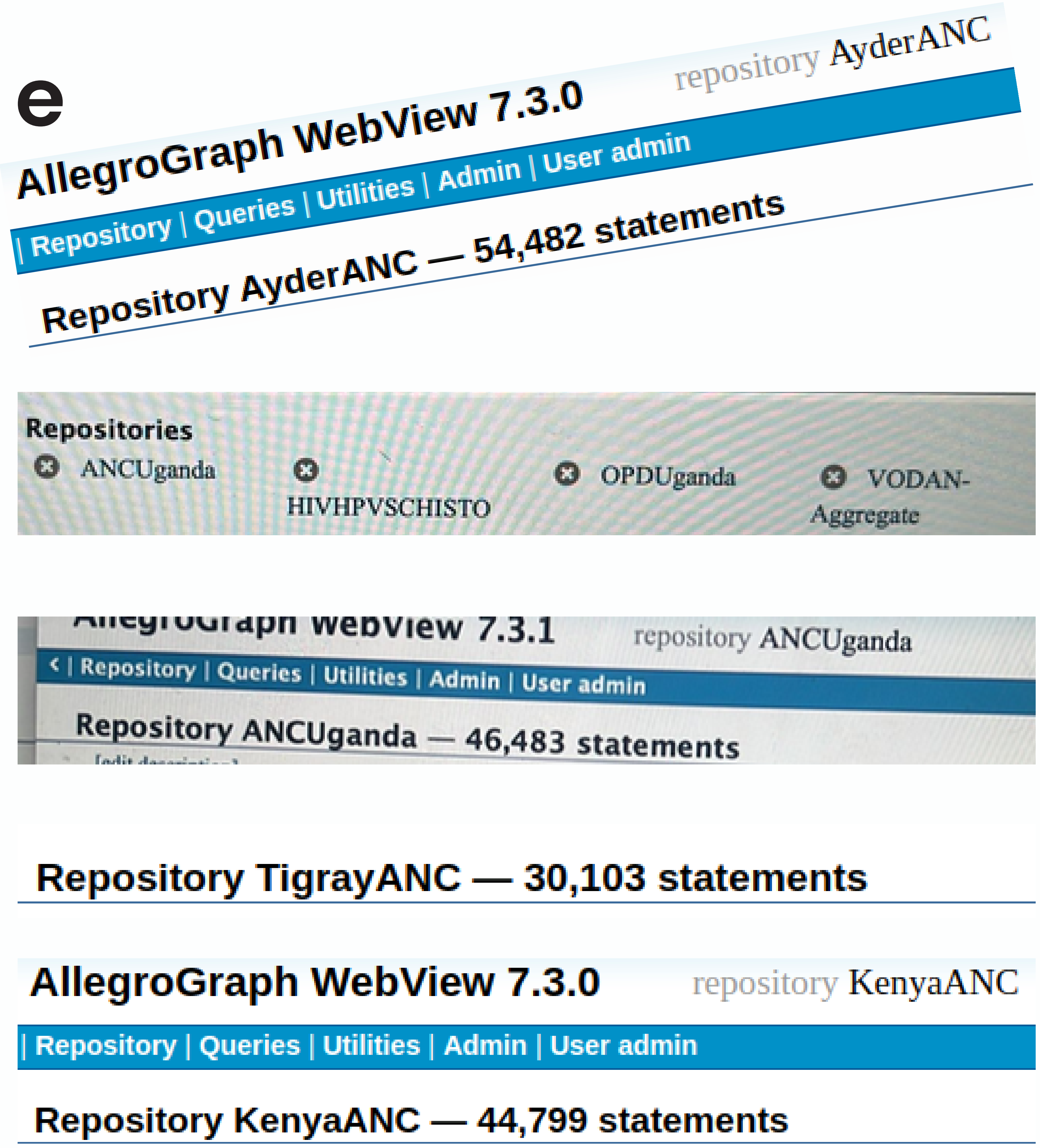
| Patient data produced in 12 health facilities for 9 months in 2023 | | | | |
|---|---|---|---|---|
| **S/N** | **Country** | **Type** | **Patient's Case count** | **Triples** |
| 1 | Ethiopia | ANC | 2,630 | 84,585 |
| 2 | Kenya | ANC | 1,181 | 44,799 |
| 3 | Uganda | ANC | 910 | 46,483 |
| | | OPD | 1,838 | 91,900 |
| 4 | Nigeria | ANC | 1,606 | 80,300 |
| | | OPD | 3,707 | 185,350 |
| | | | 11,872 | 533,417 |

- Antenatal Care (ANC)
- Outpatient Department(OPD)

| Research data FAIRification | | | |
|---|---|---|---|
| **S/N** | **Country** | **Type** | **Case count** |
| 1 | Ethiopia | Neonatal Care | patients from NICU |
| 2 | Uganda | DHIS 2 - HPV Vaccinations | national aggregates |
| | | DHIS2 - HIV | national aggregates |
| | | DHIS2 - HPV Vaccinations | national aggregates |
| 3 | Tunisia | Refugees COVID | media reports |
| | | Refugees - Care | interview data |
| 4 | LUMC | Vaccination trial | repeated experimental design |

# Quality Data Pipeline

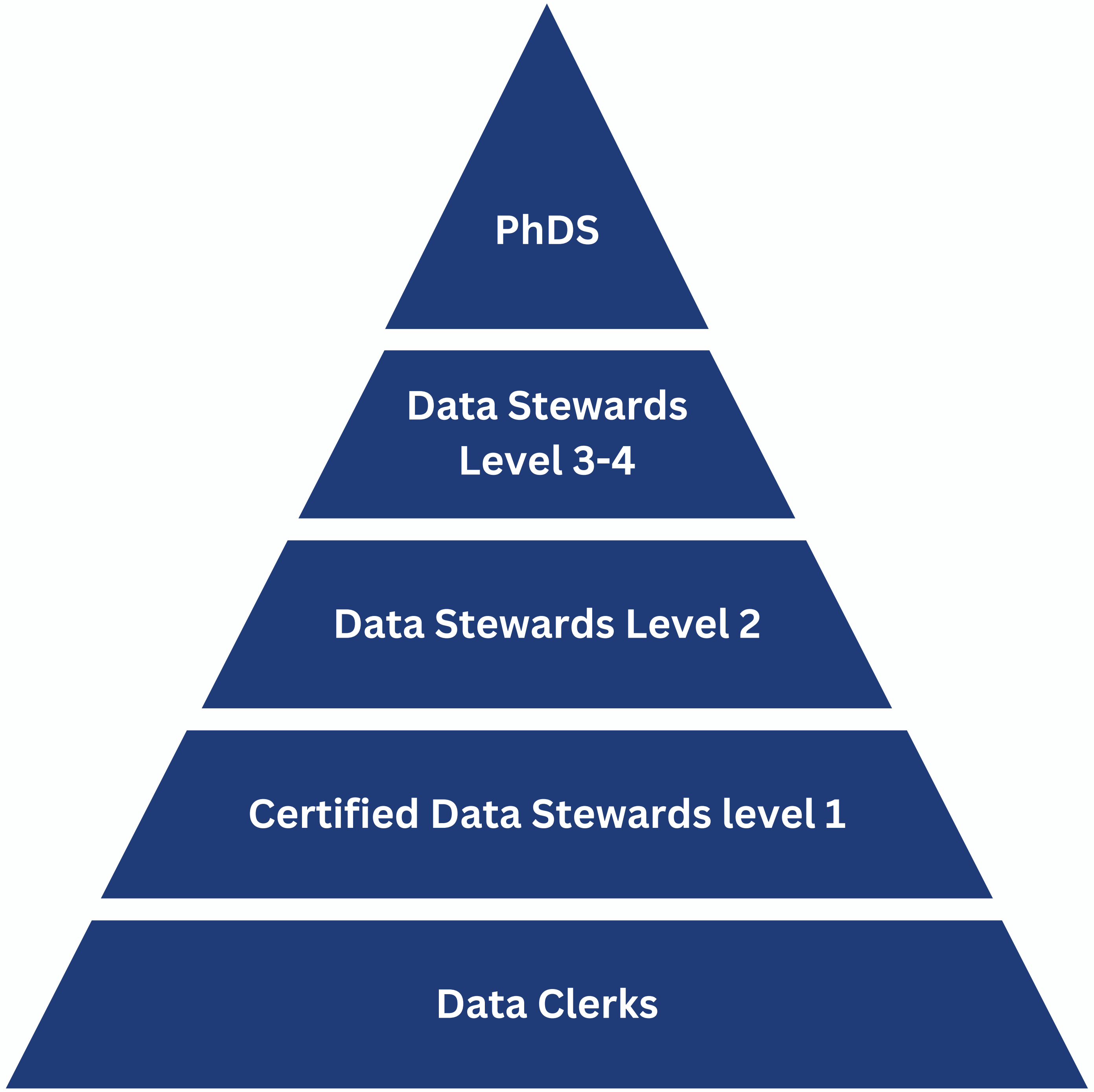
| S/N | Name of the hospital | Country |
|---|---|---|
| 1.1. | Railway Clinic | Nigeria |
| 1.2. | GH Lapai | Nigeria |
| 1.3. | OOU | Nigeria |
| 1.4. | FUL | Nigeria |
| 2.1. | PUMWANI | Kenya |
| 2.2. | Beacon of Hope | Kenya |
| 3.1. | Lira Hospital | Uganda |
| 3.2. | Ark specialist Hospital | Uganda |
| 3.3. | Kampala International University | Uganda |
| 3.4. | Hoima referral Hospital | Uganda |
| 4.1. | Ayder Referral Hospital | Ethiopia - Tigray |
| 4.2. | Mekelle Hospital | Ethiopia - Tigray |



03

# Data Stewardship & Research

Samson Yohannes




## Researchers

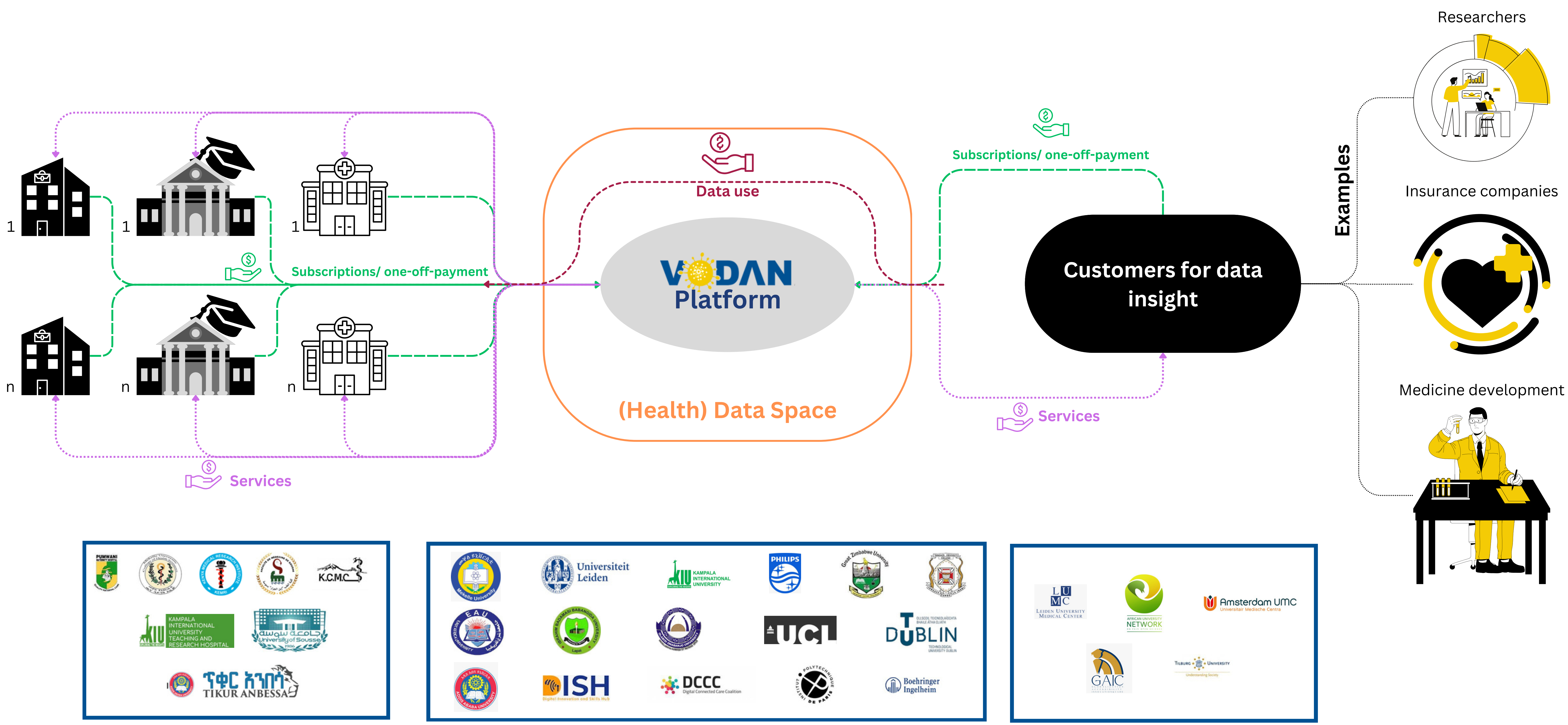
| | |
|---|---|
| Abdullahi Kawu | FAIR-data based Interoperable Digital Generated Data |
| Aliya Aktau | Vocabulary creation for interoperable FAIR-data |
| Charles Kahiro | Models for ethical value creation on federated FAIR Data production |
| Getu Tadele Taye | Modeling a regional surveillance health system of FAIR-data |
| Ibrahim Bwaga | Identifying communities of differentiated risk profiles |
| Joëlle Stocker | FAIRification of soundscape data for resilience modeling |
| Kai Smits | Human Trafficking analytics in Libya |
| Kudakwashe Kindoza | Deployment factors of federated FAIR-Data for interoperable solutions |
| Liya Manu | Integration of harmonised federated FAIR Data information for off-line use |
| Mariam Basajja | Creation of a FAIR-data based digital information system in Uganda |
| Mildred Akandinda | Identifying communities of differentiated profiles in Uganda |
| Morgane Wirtz | Migrants health analytics in Tunisia |
| Natascha Buchs | Business disruption through the FAIR-data based Federated Data Space |
| Putu Hadi Purnama Jati | GDPR-based access and control permission architecture |
| Rens Kievit | Automated permission controls for a GDPR compliant secure dynamic architecture |
| Ruduan Plug | Statistical models for federated FAIR Data models of privacy data produced |
| Samson Yohanes Amare | Federated software services for FAIR-data |
| Tesfit Gebremeskel | Modeling of vocabularies of federated FAIR-data |

04

# Business case

Prof. Mirjam van Reisen

| Customers data stewardship | Services | Customers third party data re-use |
|---|---|---|



# VODAN-Africa presents appreciation for support obtained from:

Accenture
Achmea
CordAID
DCCC
Go-FAIR Foundation
Nuffic - OKP
Invest International
IMS
Philips
Philips Foundation
and other partners

Kampala International University
Mekelle University
Addis Abeba University
Olabisi Onabanjo University
Tangaza University
East Africa University
Eastern College
Tangaza University College
Great Zimbabwe University
University of Sousse
and all other partners

LUMC & Leiden University
Leiden Institute  of Advanced Computer Science (LIACS)
Universityé Polytechnique et societé de Paris
Dublin Technical University
University College London
Boehringer Ingelheim
Amsterdam UMC
Wageningen University
and all other partners

DISH - Digital Innovation and Skills Hub

Globalisation, Accessibility, Innovation and Care Research network - GAIC

Africa University Network on FAIR Open Science

05

# Governance

Prof. Mirjam van Reisen

## KIU Senate (Uganda)

## Presidency of the VODAN-Africa Foundation

Mohamed Mpezamihigo; Francisca Oladipo; Han Baartmans, James Wilderspin

## VODAN Executive Board

| Country coordinators | Technical Team | Medical & Outreach Team | Research Team |
| --- | --- | --- | --- |
| Reginald Nalugala<br>Mariam Basajja<br>Araya Medhanyie<br>Ephrem Biruk<br>Ibrahim Abdullahi<br>Sakinat Folorunsa<br>Bernard Chazovacchii<br>Jeremy Pyuza<br>Jamal Mohamed Warsawe<br>Julia Duncan-Cassell<br>Meriam Ghardalou | Samson Yohannes<br>Ruduan Plug | Frank Kaharuza<br>Lieve Fransen | Munyaradzi Mawere<br>Mirjam van Reisen<br>Simcha Jong<br>Joshua Pos |

# 06

# Way forward

*Prof. Mirjam van Reisen*

2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030

| ***Proof of Concept*** | ***Minimal Viable Product and adoption assessment*** | ***Africa Health Data Space - Africa data Space with selected partners*** | ***Professional Service Organisation linked to the Africa University Network on FAIR Open Science to root capacities*** | ***Integration of domains in the Global Data Space with private sector customers*** |
|---|---|---|---|---|
| LUMC-KIU | 8 countries in Africa, 88 health facilities, research data | Partnership of 10 countries in Africa, 2 countries in Asia, 3 countries in the EU and US | Africa-based HQ | Leader for service development with new content domains |
| Go-FAIR tooling | CEDAR-based tooling, business proposition | VODAN software, test of business model, strengthen permission control, security and protection | Administrative independent business operation | VODAN software, test of business model |
| Philips Foundation | Philips, Philips Foundation and Invest International | Philips, IMS, Microsoft, Achmea, Accenture, .. | Soft loan and investment partnerships and strengthening of market position | Solid products and strengthening of market position |
| Article advanced genetics | Special Issue & Datastewardship manual | Book FAIR Data Science in Africa, research group established in GAIC Research Network | 6 dissertations published by PhDs | 20 dissertations published by PhDs |
| 2 Datastewards | 30 Datastewards and technical data science team | 40 Datastewards, software engineering team, data quality assessment team, data analytics capacity, permission and security capacity | Leadership available in all business domains, and specialised data-stewardship service | Integration of interoperable services in data production and data analytics and services provision |

# List of publications

[1] van Reisen, M., Amare, S.Y., Plug, R., Tadele, G., Gebremeskel, T., Kawu, A.A., Smits, K., Woldu, L.M., Stocker, J., Heddema, F., Folorunso, F., Kievit, R., Medhaniye, A.A. (2023 - forthcoming). Curation of Federated Patient Data: A Proposed Landscape for the Africa Health Data Space. *Elsevier*.

[2] 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

[3] 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

[4] Kawu, A. A., O'Sullivan, D., Hederman, L., & Van Reisen, M. (2023). FAIR4PGHD: A framework for FAIR implementation over PGHD. *In: FAIR Connect*, 1(1), 35-40. IOS Press. DOI: https://doi.org/10.3233/FC-230500

[5] 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

[6] 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

[7] Van Reisen, M., Stokmans, M., Mawere, M., Basajja, M., Ong'ayo, A., Nakazibwe, P., Kirkpatrick, K., Kudakwashe, C. (2020). FAIR practices in Africa. *In: Data Intelligence 2(2020)*, 264–275. doi: 10.1162/dint_a_00049 (eds. Mons, B., Jacobsen, A. & Schultes, E.). MIT Press, http://www.mitpressjournals.org/dint

[8] Van Reisen, M., Stokmans, M., Basajja, M., Ong'ayo, A., Kirkpatrick, K., Mons, B., (2020). Towards the Tipping Point of FAIR Implementation. *In: Data Intelligence 2(2020)*, 264–275. doi: 10.1162/dint_a_00049 (eds. Mons, B., Jacobsen, A. & Schultes, E.). MIT Press, http://www.mitpressjournals.org/dint

[9] Jacobsen, A., Miranda Azevedo, de R., Juty, N., Van Reisen, M., ..., Schultes, E. et. al. (2020). FAIR principles: interpretations and implementation considerations. *In: Data Intelligence 2(2020)*, 264–275. doi: 10.1162/dint_a_00049 (eds. Mons, B., Jacobsen, A. & Schultes, E.). MIT Press, http://www.mitpressjournals.org/dint

## Special Issue Articles

[10] Van Reisen, M., Mons, B.: Introduction to the Special Issue: Data Intelligence on patient health records. *Data Intelligence* 4(4) (2022). doi: 10.1162/dint_e_00165

[11] 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.: Incomplete COVID-19 data: The curation of medical health data by the Virus Outbreak Data Network-Africa. *Data Intelligence* 4(4) (2022). doi: 10.1162/dint_e_00166

[12] Plug, R., Liang, Y., Aktau, A., Basajja, M., Oladipo, F., Van Reisen, M.: Terminology for a FAIR framework for the Virus Outbreak Data Network-Africa. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00167

[13] Stocker, M., Stokmans, M., Van Reisen, M.: Agenda setting on FAIR Guidelines in the European Union and the role of expert committees. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00168

[14] Lin, Y., Purnama Jati, P.H., Aktau, A., Ghardallou, M., Nodehi, S., Van Reisen, M.: Implementation of FAIR Guidelines in selected non-Western geographies. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00169

[15] Basajja, M., Van Reisen, M., Oladipo, F.: FAIR Equivalency with regulatory framework for digital health in Uganda. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00170

[16] Taye, G.T., Amare, Y.S., Gebremeskel, G.T., Medhanyie, A.A., Ayele, W., Habtamu, T., Van Reisen, M.: FAIR Equivalency with regulatory framework for digital health in Ethiopia. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00172

[17] Purnama Jati, P.H., Lin, Y., Nodehi, S., Cahyono, D.B., Van Reisen, M.: FAIR versus open data: A comparison of objectives and principles. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00176

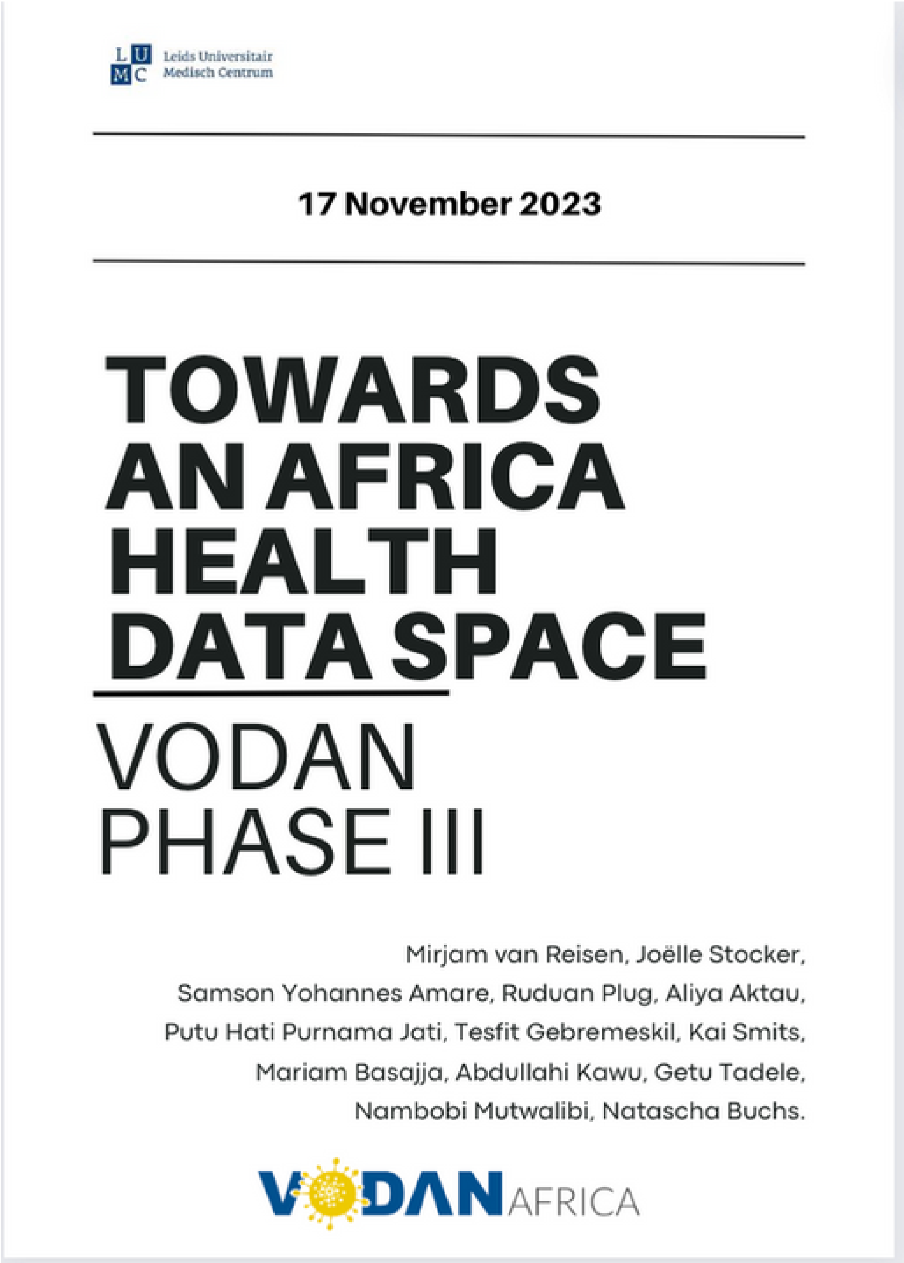
[18] Basajja, M., Suchanek, M., Taye, G.T., Amare, S.Y., Nambobi, M., Folorunso, S., Plug, R., Oladipo, F.O., Van Reisen, M.: Proof of concept and horizons on deployment of FAIR Data Points in the COVID-19 pandemic. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00179

[19] Purnama Jati, P.H., Van Reisen, M., Flikkenschild, E., Oladipo, F., Meerman, B., Plug, R., Nodehi, S.: Data access, control, and privacy protection in the VODAN-Africa architecture. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00180

[20] Ghardallou, M., Wirtz, M., Folorunso, S., Touati, Z., Ogundepo, E., Smits, K., Mtiraoui, A., Van Reisen, M.: Expanding non-patient COVID-19 data: Towards the FAIRification of migrants' data in Tunisia, Libya and Niger. *Data Intelligence 4(4)* (2022). doi: 10.1162/dint_a_00181

[21] Akindele, A.T., Tayo, A.O., Taye, G.T., Amare, S.Y., Van Reisen, M., Berhe, K.F., Gusite, B., Edozie, E.: The impact of COVID-19 and FAIR data innovation on distance education in Africa. *Data Intelligence 4(4)*, (2022). doi: 10.1162/dint_a_00184

# For more information, see report below


Leids Universitair Medisch Centrum

17 November 2023

TOWARDS AN AFRICA HEALTH DATA SPACE

VODAN PHASE III

Mirjam van Reisen, Joëlle Stocker, Samson Yohannes Amare, Ruduan Plug, Aliya Aktau, Putu Hati Purnama Jati, Tesfit Gebremeskil, Kai Smits, Mariam Basajja, Abdullahi Kawu, Getu Tadele, Nambobi Mutwalibi, Natascha Buchs.

VODAN AFRICA

Link to report