

Field Labs

Data Science in Practice

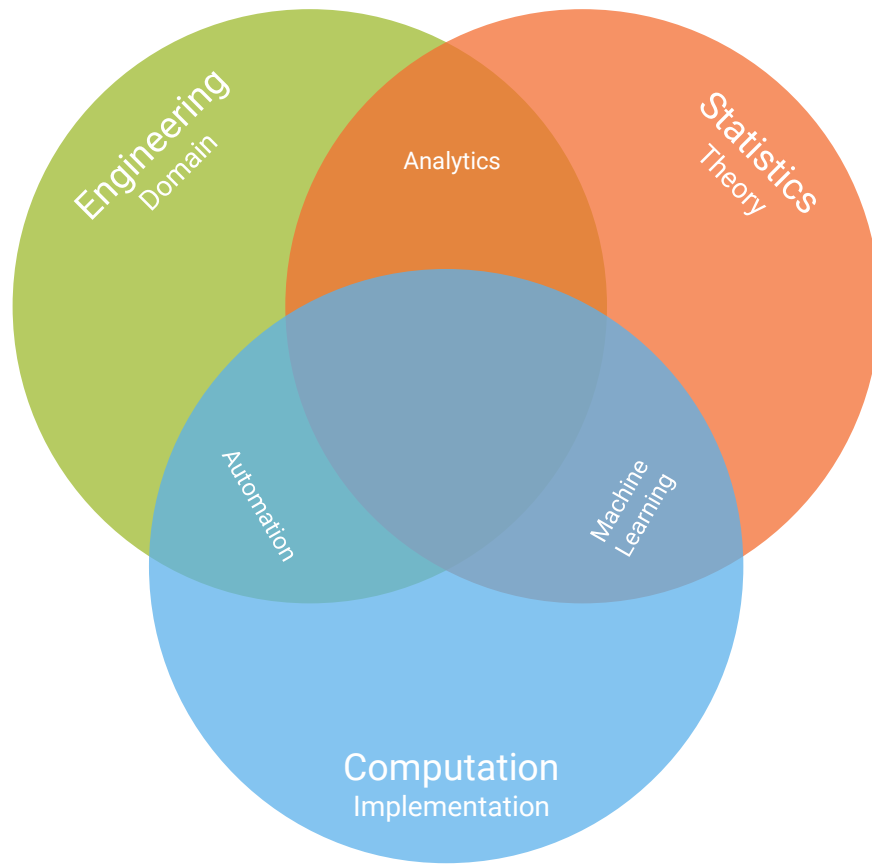


**Universiteit
Leiden**
The Netherlands

Discover the world at Leiden University

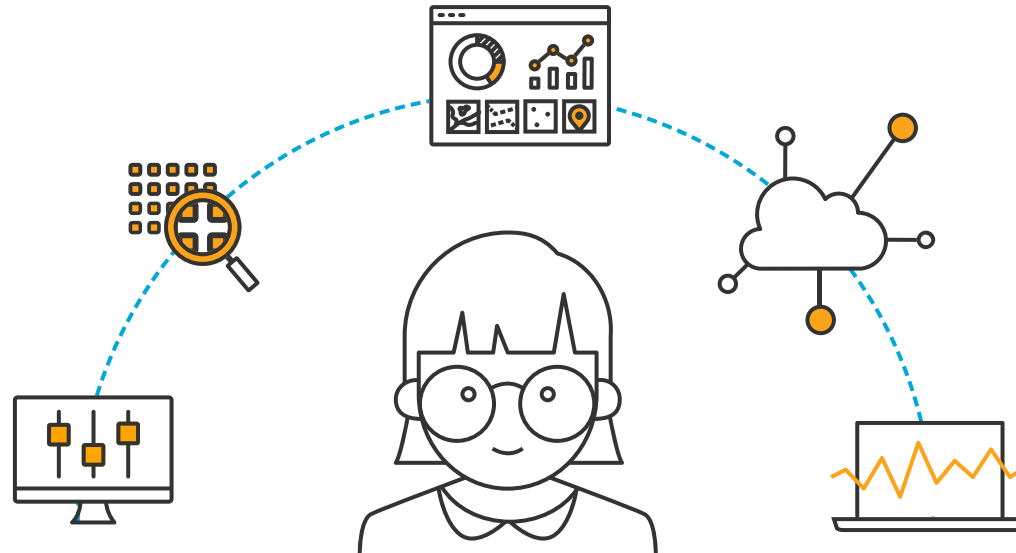
Data Science

In Practice



Data science in practice is a
team-based effort

Stakeholders should be involved in every step

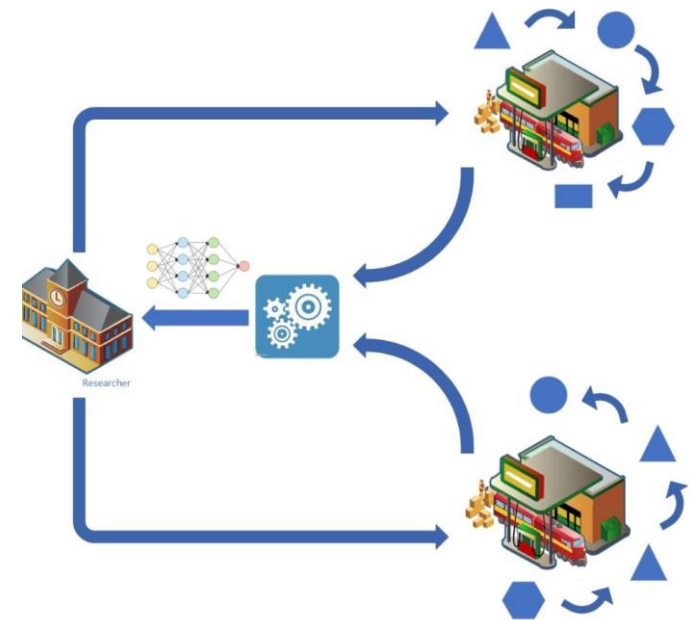


Field Labs

- Field Labs are the main practical component of DSiP
- There are 8 different field labs in total
 - Intradisciplinary
 - Many different organizations
 - Various technical challenges, similar practical challenges
- You will be working with teams of 2-4 people
- Approximately 2 months

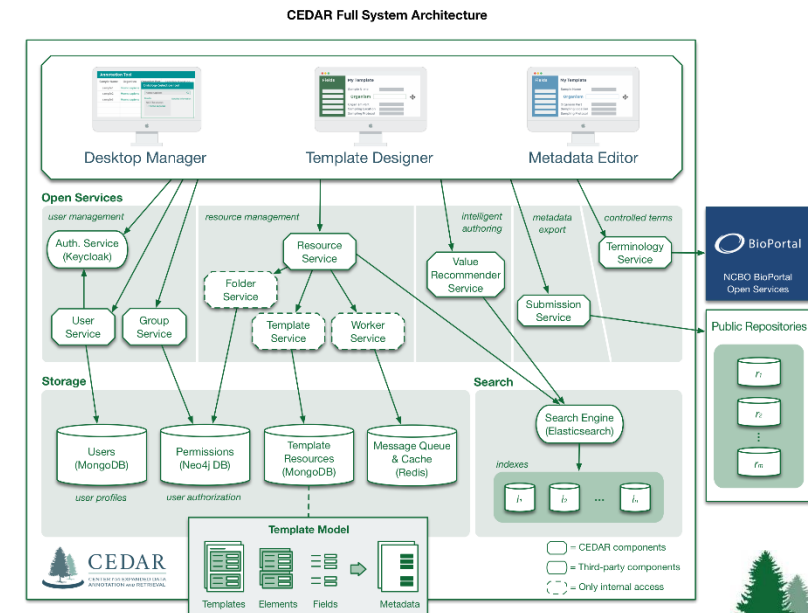
Field Lab 1 – Personal Health Train

- Develop Architecture for Personal Health Train
 - Health application of data federation
 - Specify Engineering requirements with domain stakeholders in health
 - Develop architectural specifications based on recent academic literature



Field Lab 2 – Data Service Deployment

- Deployment of FAIR Metadata Service
 - Develop a Localized Stanford CEDAR service within LUMC
 - Document deployment process and analyze remote deployment strategies
 - Explore possibilities for automation and optimization of deployment



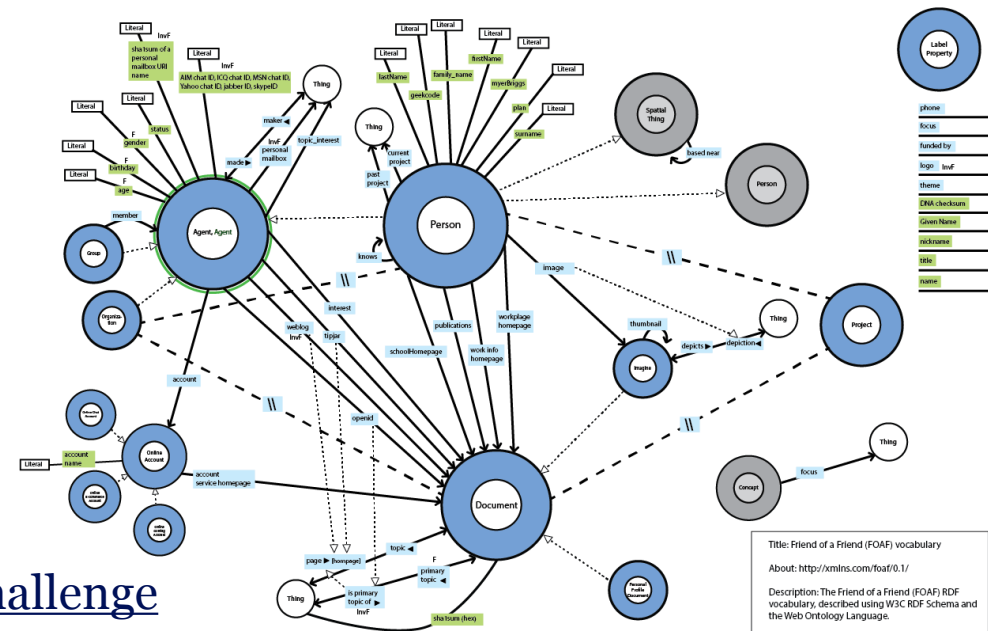
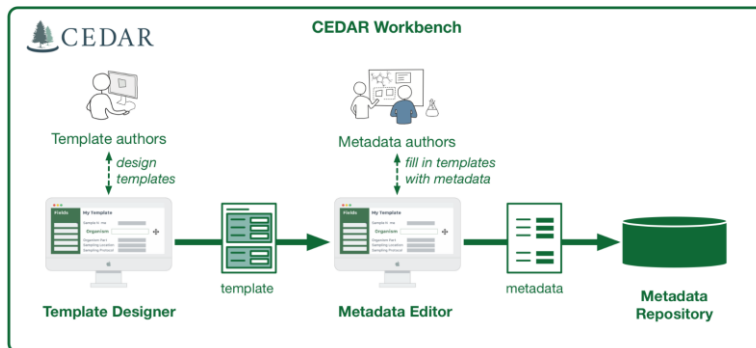
Olzhas Aldabergenov
Erik Flikkenschild

Engineering Challenge



Field Lab 3 – FAIR Vaccine Metadata

- Develop Metadata Templates for New Vaccine Developed by LUMC
 - Based on rigorous controlled vocabulary and ontology specification
 - Requires extensive collaboration with domain experts
 - Implementation of FAIR principles

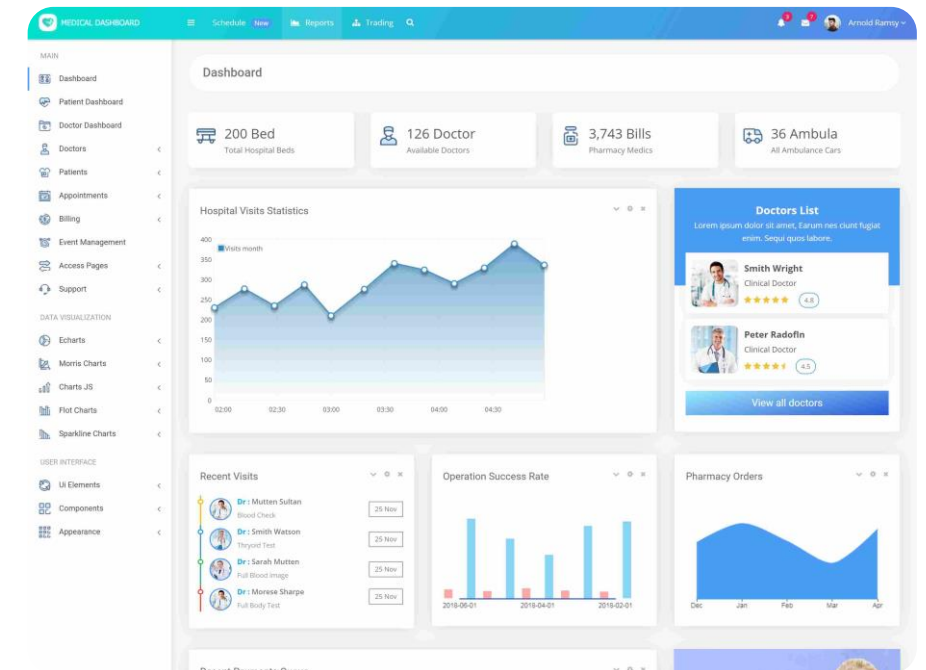


Aliya Aktau

Modelling Challenge

Field Lab 4 – Querying and Visualization

- Development of Analytics and Visualization for Health Facilities
 - Develop from a Full Front-End Prototype Solution for Client-Responsive Visualizations
 - Driven by requirements specifications and value for domain stakeholders
 - Research and design possibilities for interactivity or modularity



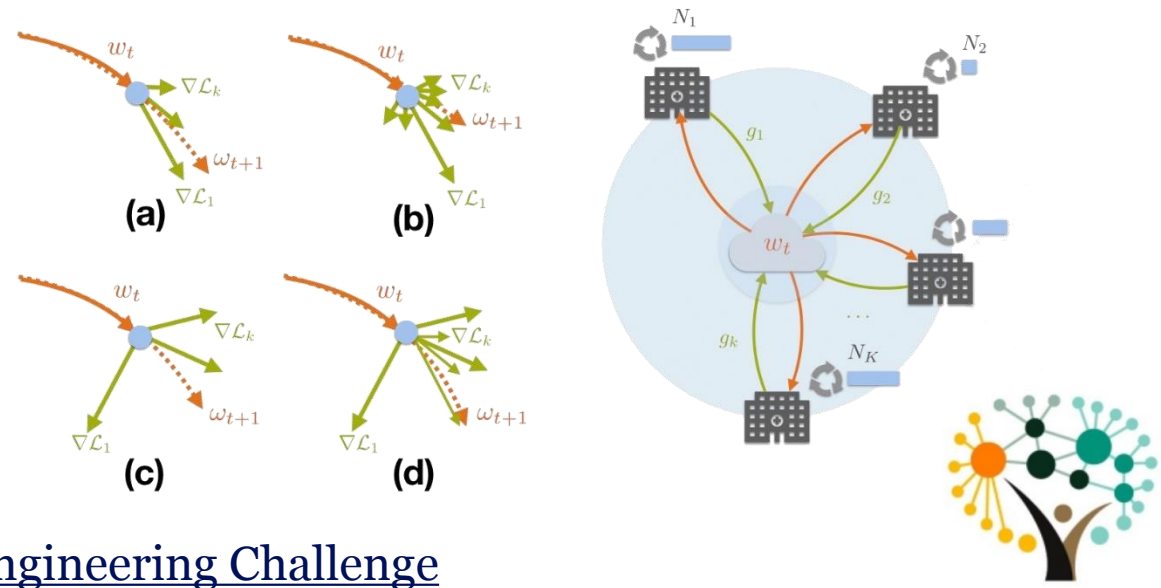
Mariam Basajjia

Engineering Challenge



Field Lab 5 – Federated Data Analytics

- Prototype Service to Query and Analyze Federated Data Sources
 - Implementation of the state-of-the-art federated data model
 - Analysis of the challenges and recent scientific developments
 - Scientifically-rigorous focus



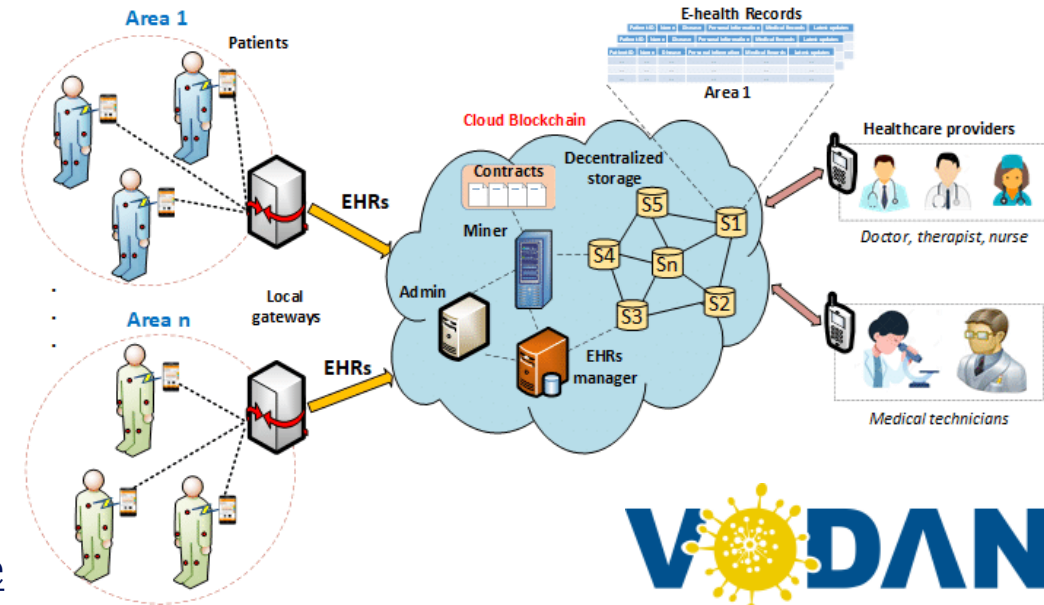
Ruduan Plug

Engineering Challenge



Field Lab 6 – Automated Security Contracts

- Develop an Automated Authorization Model for Data Processing
 - Research possibilities for automated auditing of processed data and models
 - Identify challenges in automated security and privacy analysis
 - Develop Smart contracts based on transparent access and control schemas



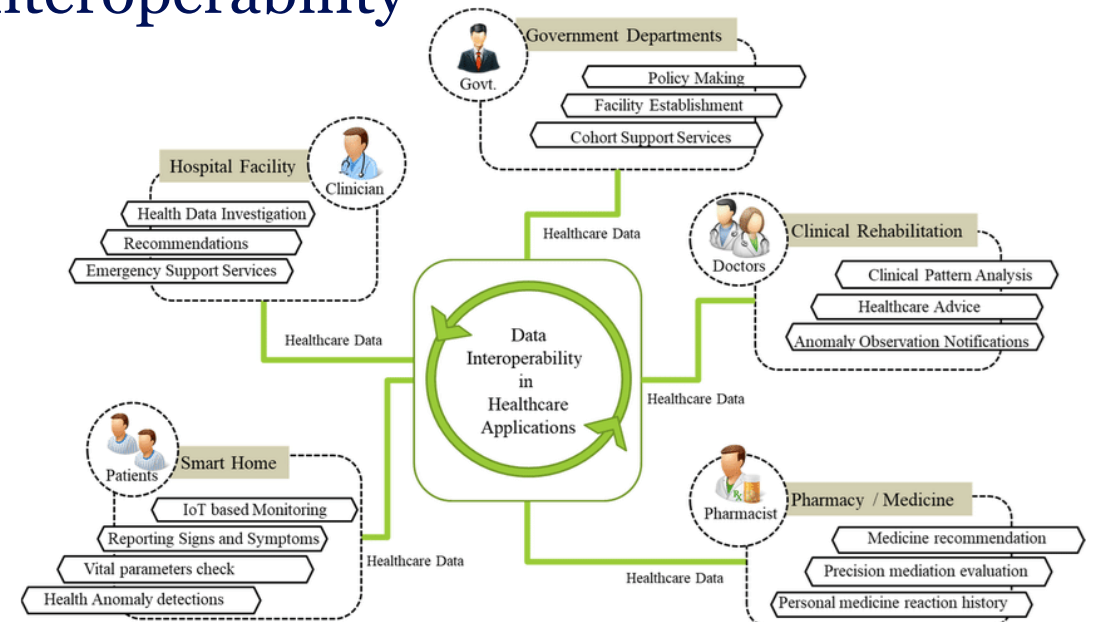
Putu Hadi Purnama Jati

Modelling Challenge



Field Lab 7 – Health Data Interoperability

- Model Data Interoperability on COVID-related Health Programmes
 - Develop business intelligence model for interoperability
 - Produce metadata templates based on common data elements
 - Discover and visualize opportunities for new research



Field Lab 8 – Document Digitalisation

- Develop a Prototype that can Digitize Physical Documents
 - Apply modern optical character recognition strategies on PDF scans
 - Research state-of-the-art methods to process diagrams or pictures
 - Consider the challenges in terms of FAIR and GDPR in digitization

Physical Invoice:

Customer address
Hallium Energy services
Violet garden, 23-10
Telephone : 04519304

Invoice no : 43876324
Dated : 17th Nov 2018
PO No : 76486234

Description	Hours worked	Amount USD
Project: NEH-JNS-HJB-HSA Services rendered by our engineers from 1st Sep to 31st Oct		
Mark mahadeo - Sep efforts	168	4547
Mark mahadeo - Oct efforts	112	3031
Subtotal		7578
Volume discount at 10%		
Total	280	6821

For wicom limited
Authorized signatory

Finance dept:
Wicom limited
Pinnacle mansion
5th road
Telephone : 84874858

Digitized Invoice (with bounding boxes):

Customer Address
Invoice No
Dated
PO Number
Total Amount

Json response:

```
{
  "response": [
    {
      "label": "Invoice No",
      "text": "43876324",
      "bounding_box": { "xmin": 77, "ymin": 86, "xmax": 95, "ymax": 96 }
    },
    {
      "label": "Date of Invoice",
      "text": "17th Nov 2018",
      "bounding_box": { "xmin": 168, "ymin": 94, "xmax": 178, "ymax": 104 }
    },
    {
      "label": "PO Number",
      "text": "76486234",
      "bounding_box": { "xmin": 330, "ymin": 230, "xmax": 342, "ymax": 232 }
    },
    {
      "label": "Customer Address",
      "text": "Hallium Energy services\nViolet garden, 23-10",
      "bounding_box": { "xmin": 220, "ymin": 156, "xmax": 235, "ymax": 165 }
    },
    {
      "label": "Total Amount",
      "text": "6821",
      "bounding_box": { "xmin": 123, "ymin": 132, "xmax": 135, "ymax": 145 }
    }
  ]
}
```

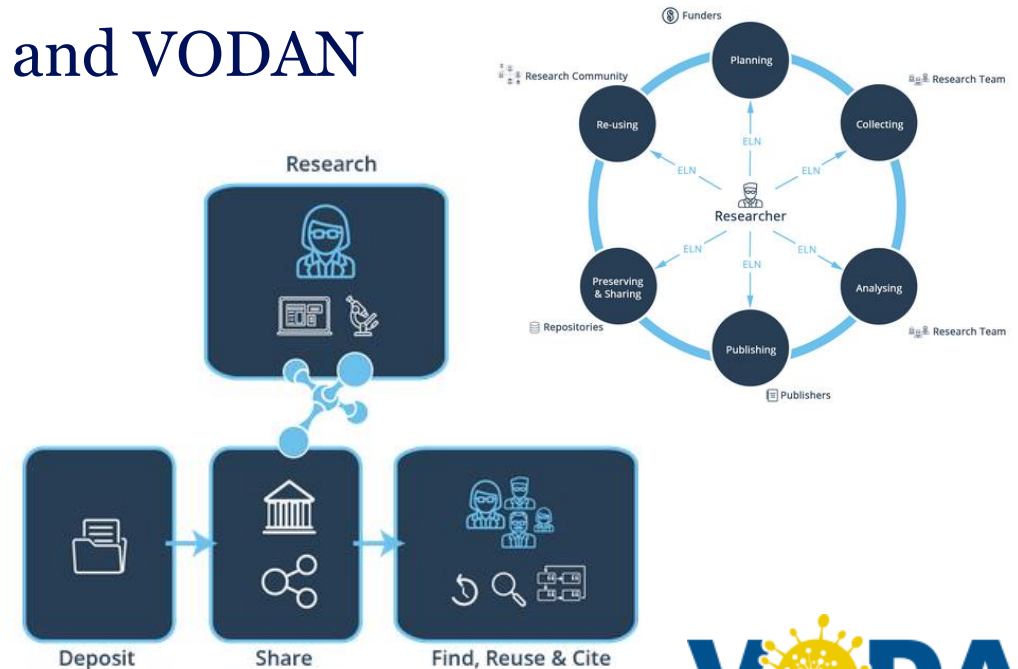
Jelena Prokic
Ruduan Plug

Engineering Challenge



Field Lab 9 – Scientific Data Interoperability

- Explore Possibilities for Interoperability between Research Projects
 - Large Scale Health Research from ZonMw and VODAN
 - Mediate between prominent scientists associated to ministries and NWO
 - Design and document promising leads to reusability of large-scale data



Data Science in Practice



Universiteit
Leiden
The Netherlands