



STOP DATA SHARING

Barend Mons SPHN, Zurich, August 2025





STOP DATA SHARING

PERCEPTION:

'Data moves around and outside my control'

But that is not the only reason....

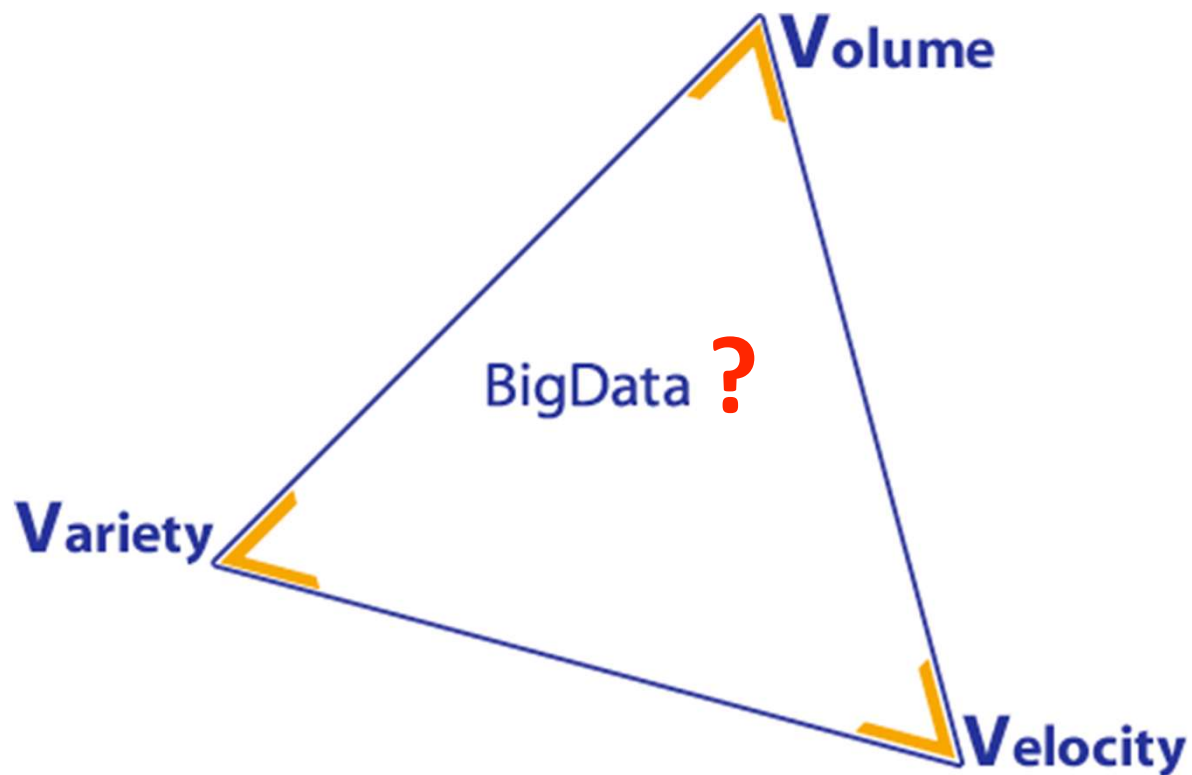


STOP DATA SHARING

Why?

1. Volume and associated issues

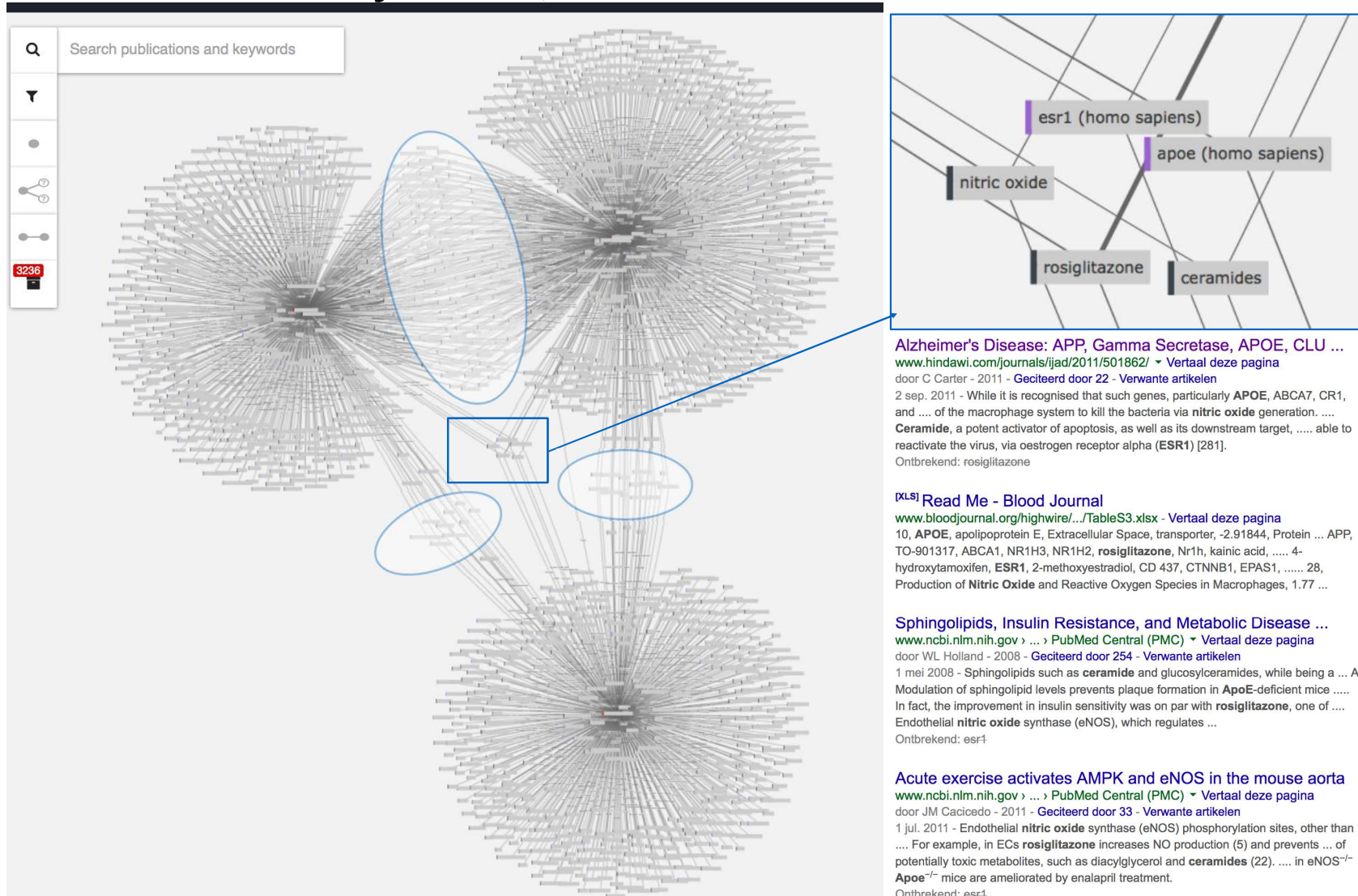
The Explicitome (LS)



10^{14} assertions
of the type S-P-O

Just an example: Evaluate conceptual links between diseases

Here: Metabolic Syndrome, Diabetes Mellitus II and E-O Alzheimer





- One GB data transfer = 10 cents
- One TB data transfer = 100 Euros
- One PB data transfer = 100,000 Euros ?
- DNA sequence data is **doubling every 6-8 months** over the last 3 years and looks to continue for this decade



STOP DATA SHARING

Why?

2. Sensitivity



Privacy ?



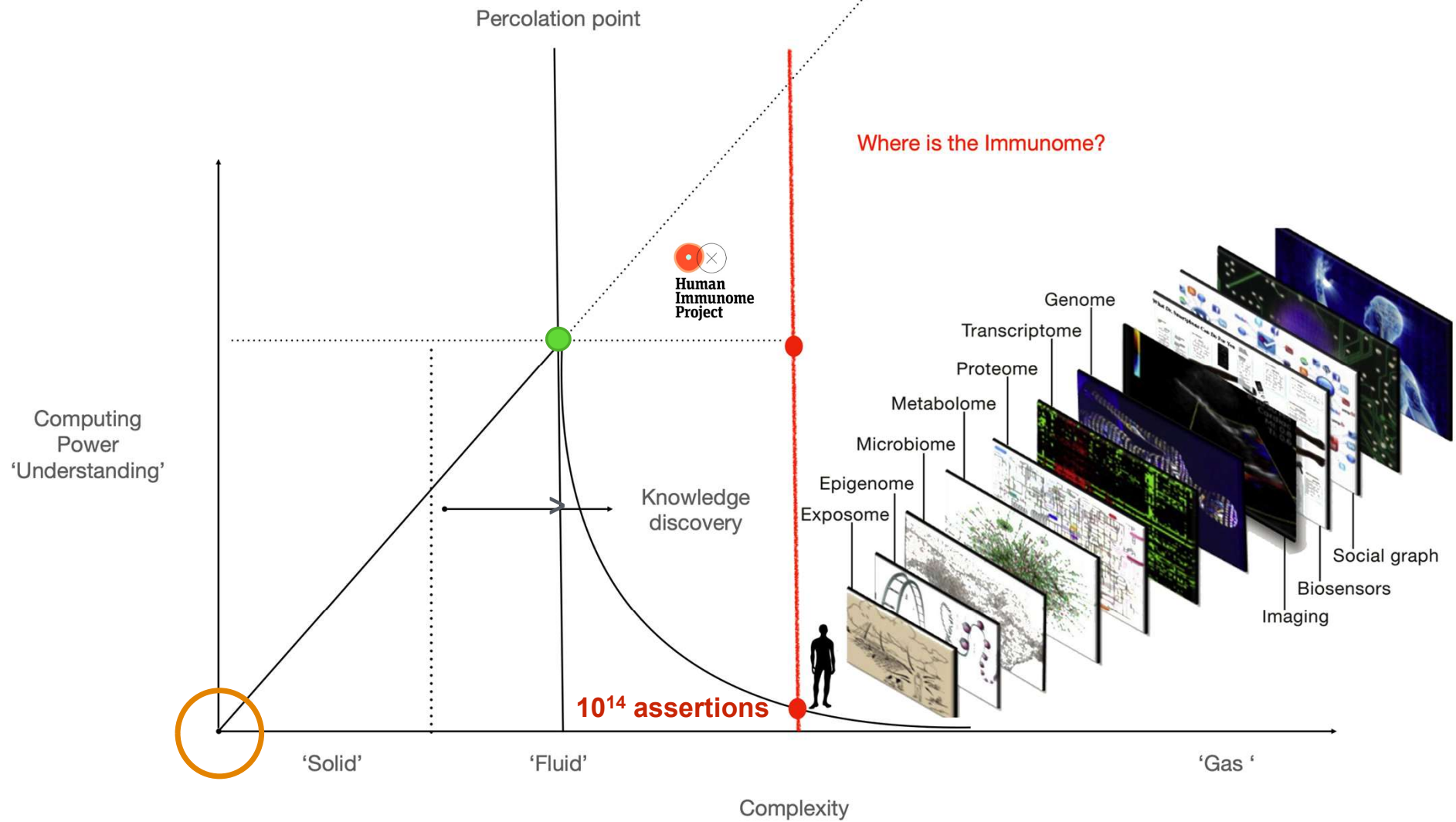


STOP DATA SHARING

Why?

3. Complexity

Chaos is just order, but perceived by a lesser mind



The end of data sharing

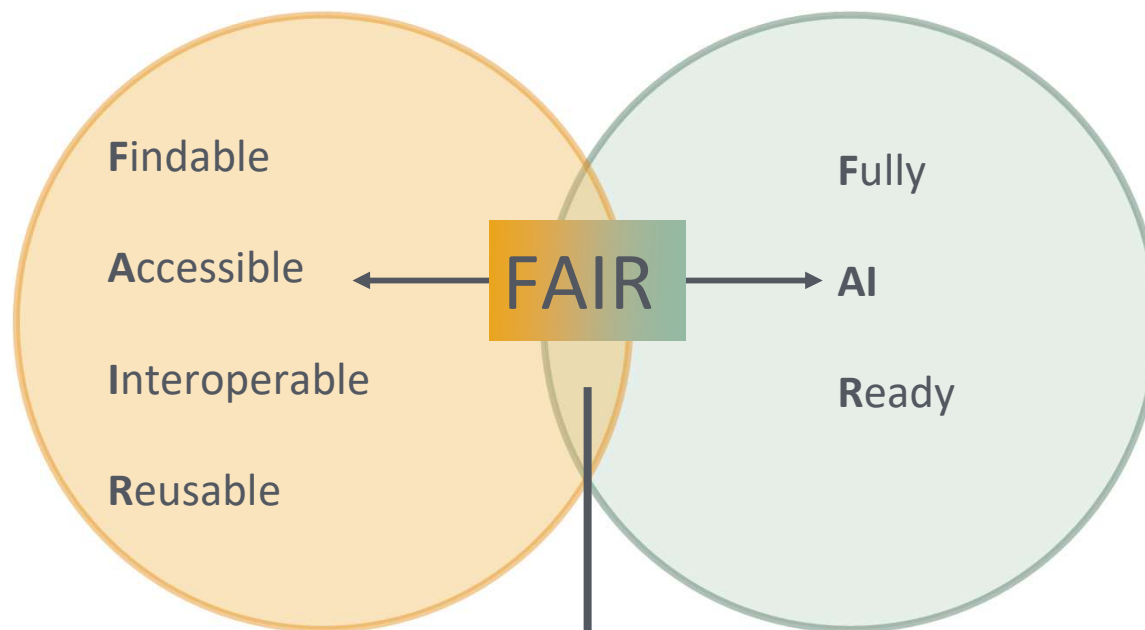


The Dawn of Data Visiting



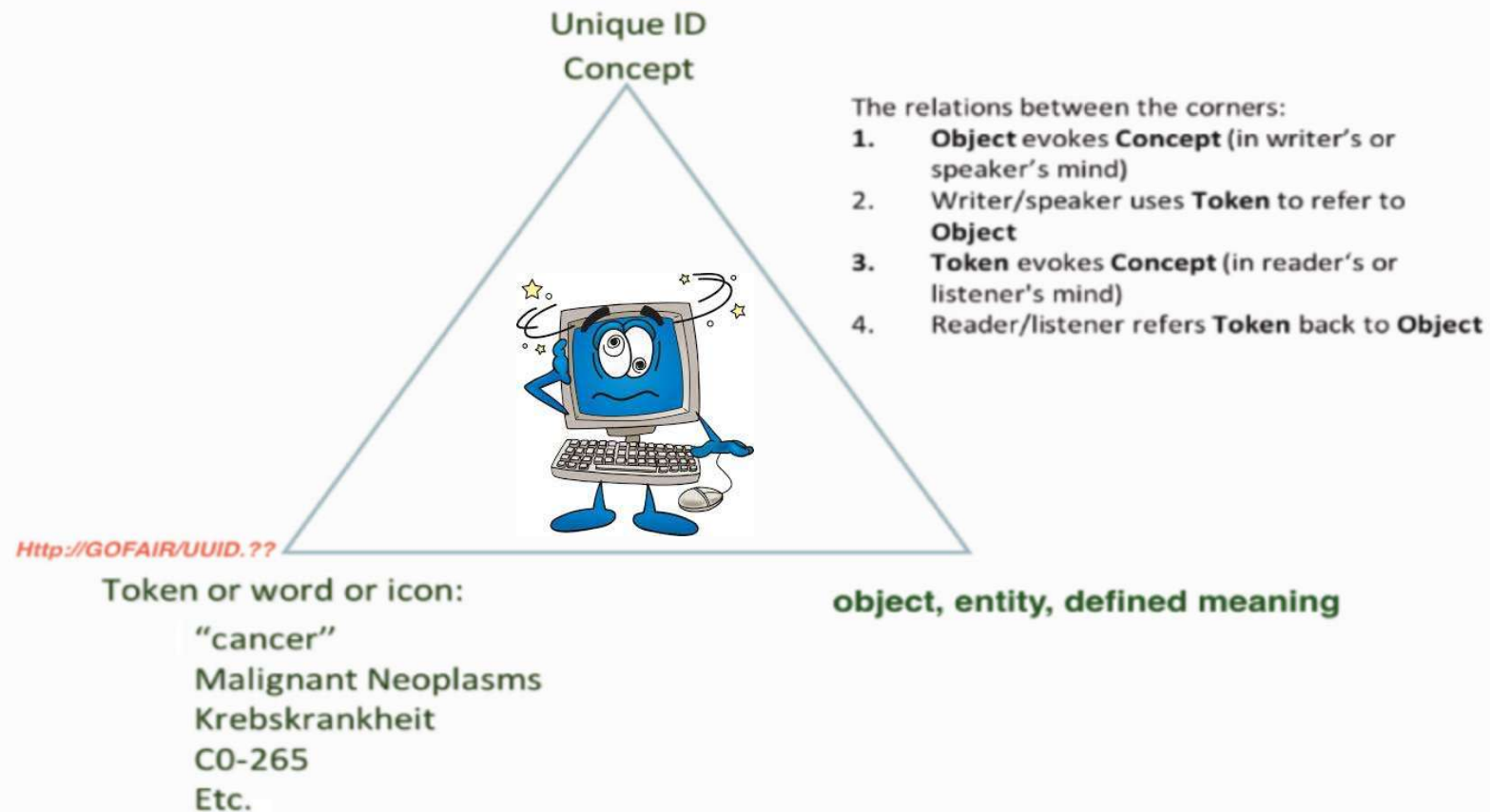
Can SPHN move to data Visiting?

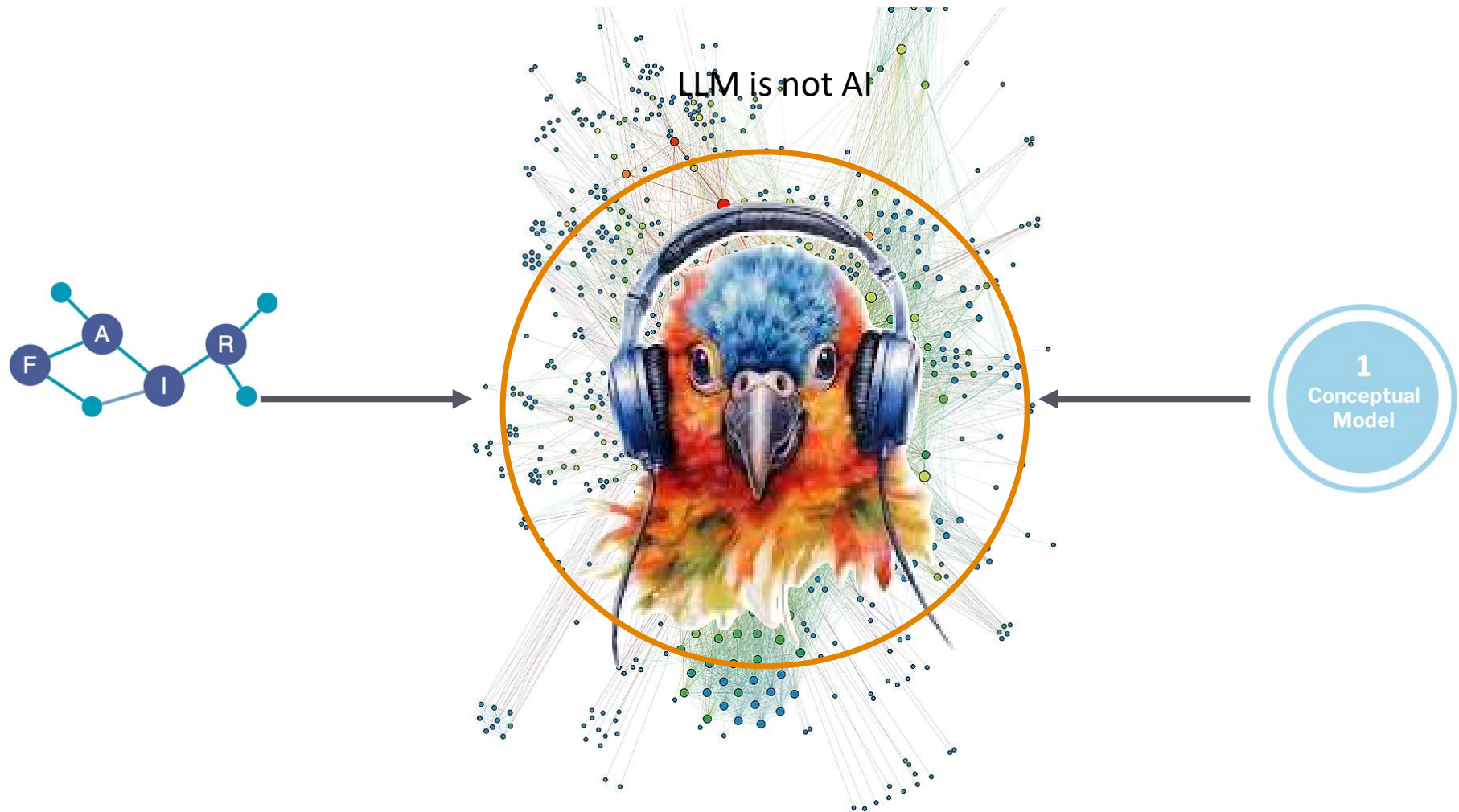
A '*conditio sine qua non*'



'The Machine Knows what I mean'

The Ogden Triangle – Concepts versus words



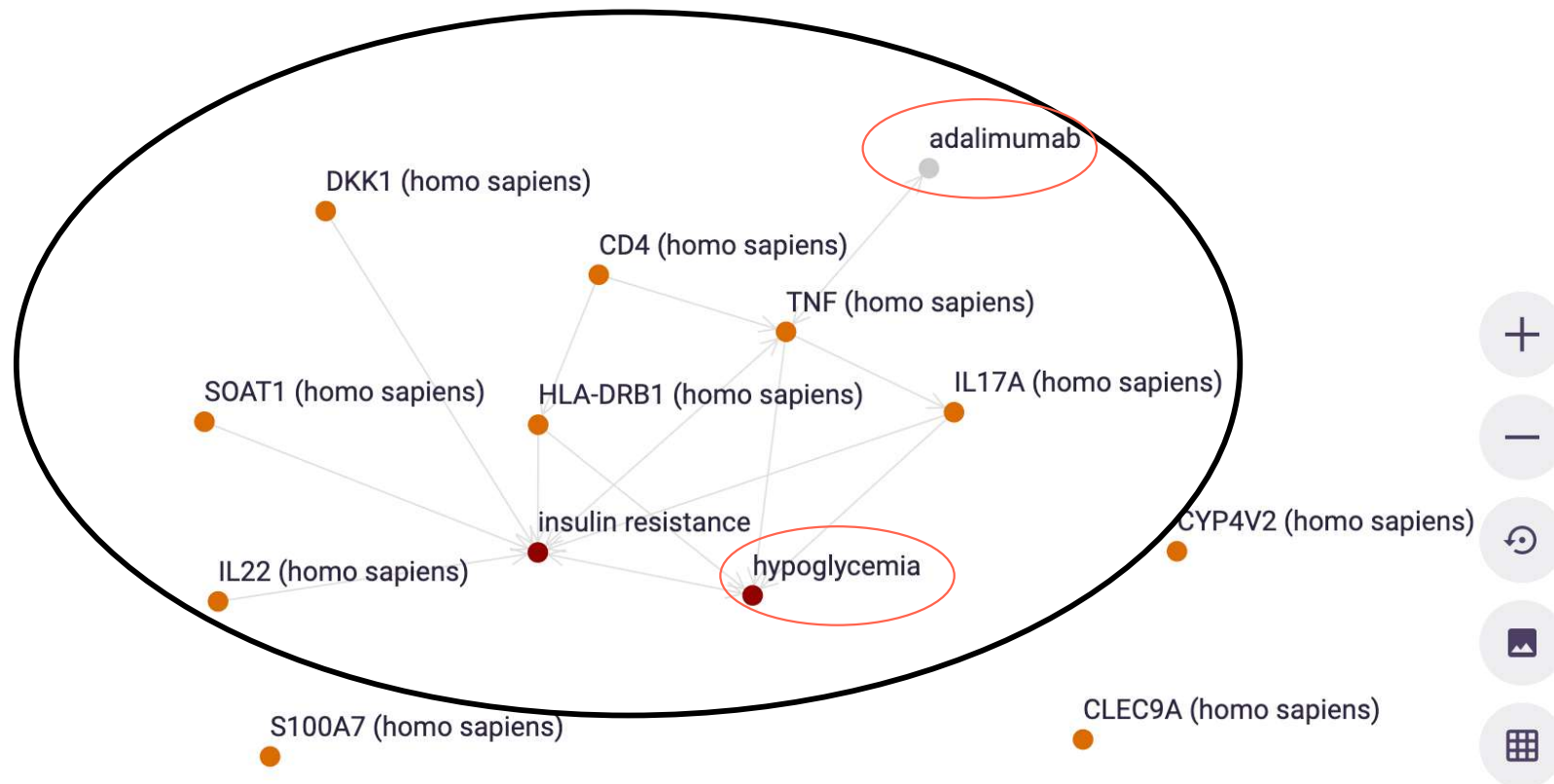


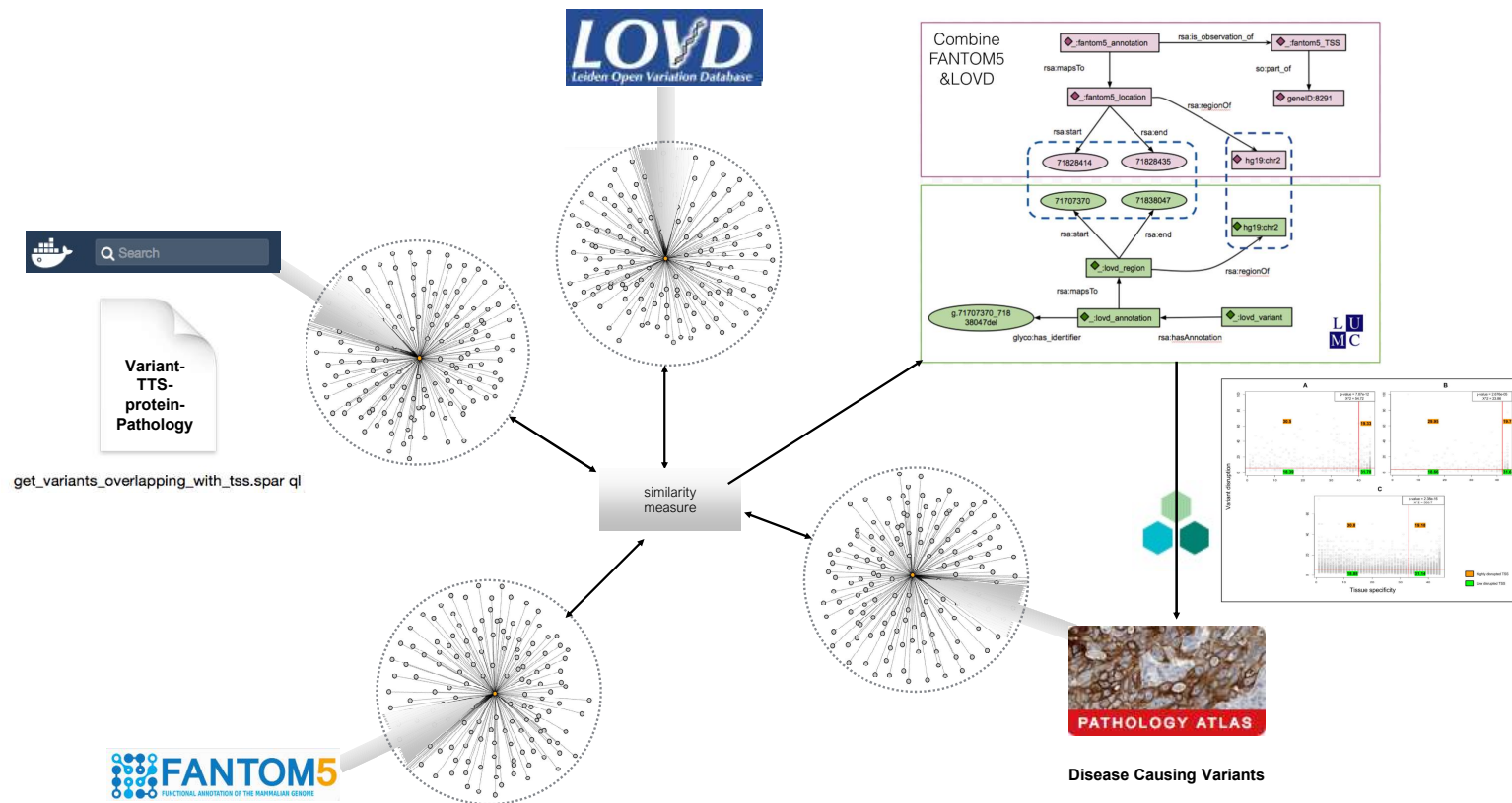
It's all hallucination, but we only call it that when we notice it's wrong. The problem is, large language models are so good at what they do that what they make up looks right most of the time. And that makes trusting them hard.

A few more examples: Drug repurposing (16.000 proteins > targets for disease?)

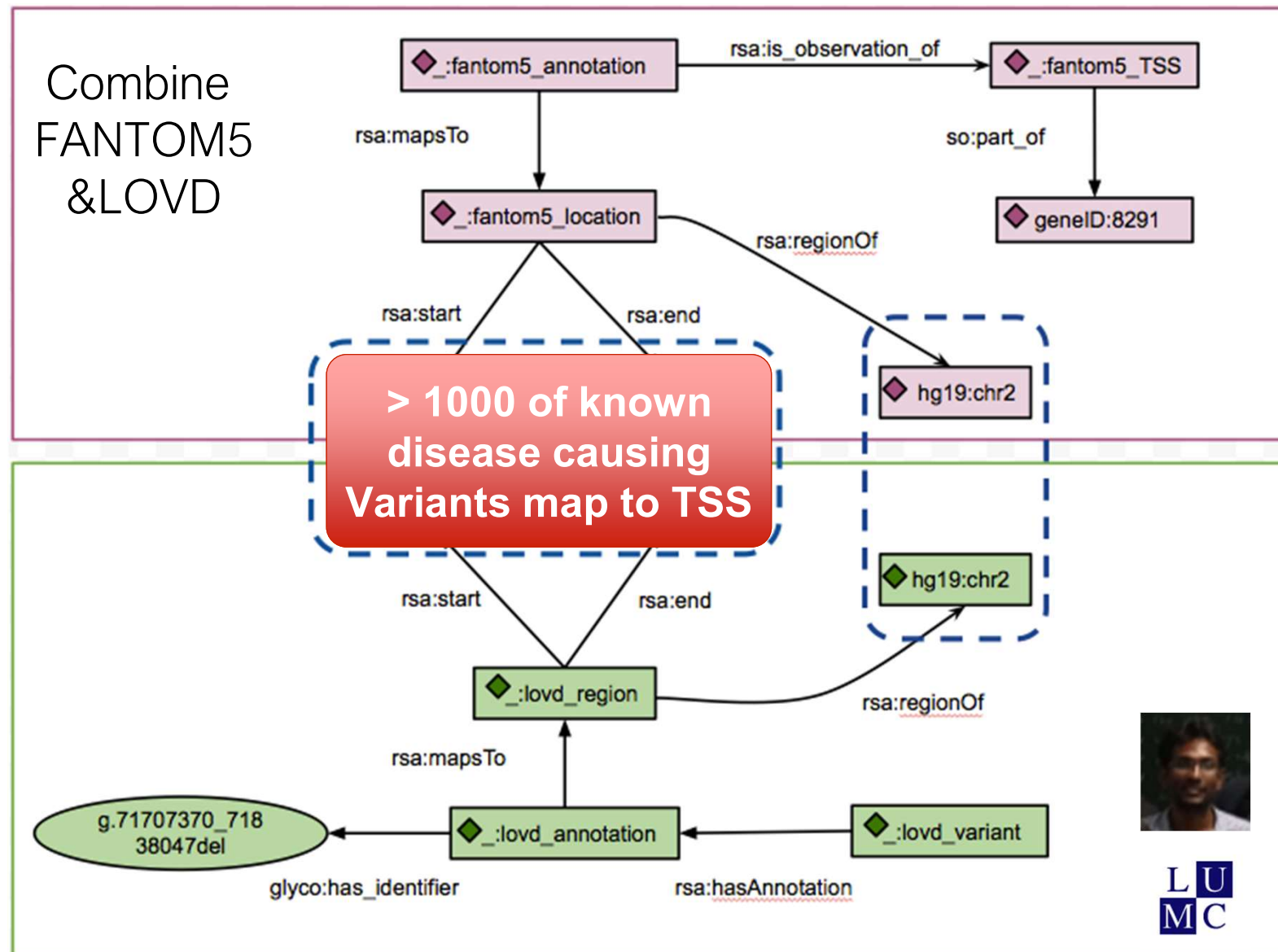
alamab-hypo ✎

6 of the genes co-occurring in sentences in literature with adalimumab and hypoglycaemia are annotated with insulin resistance

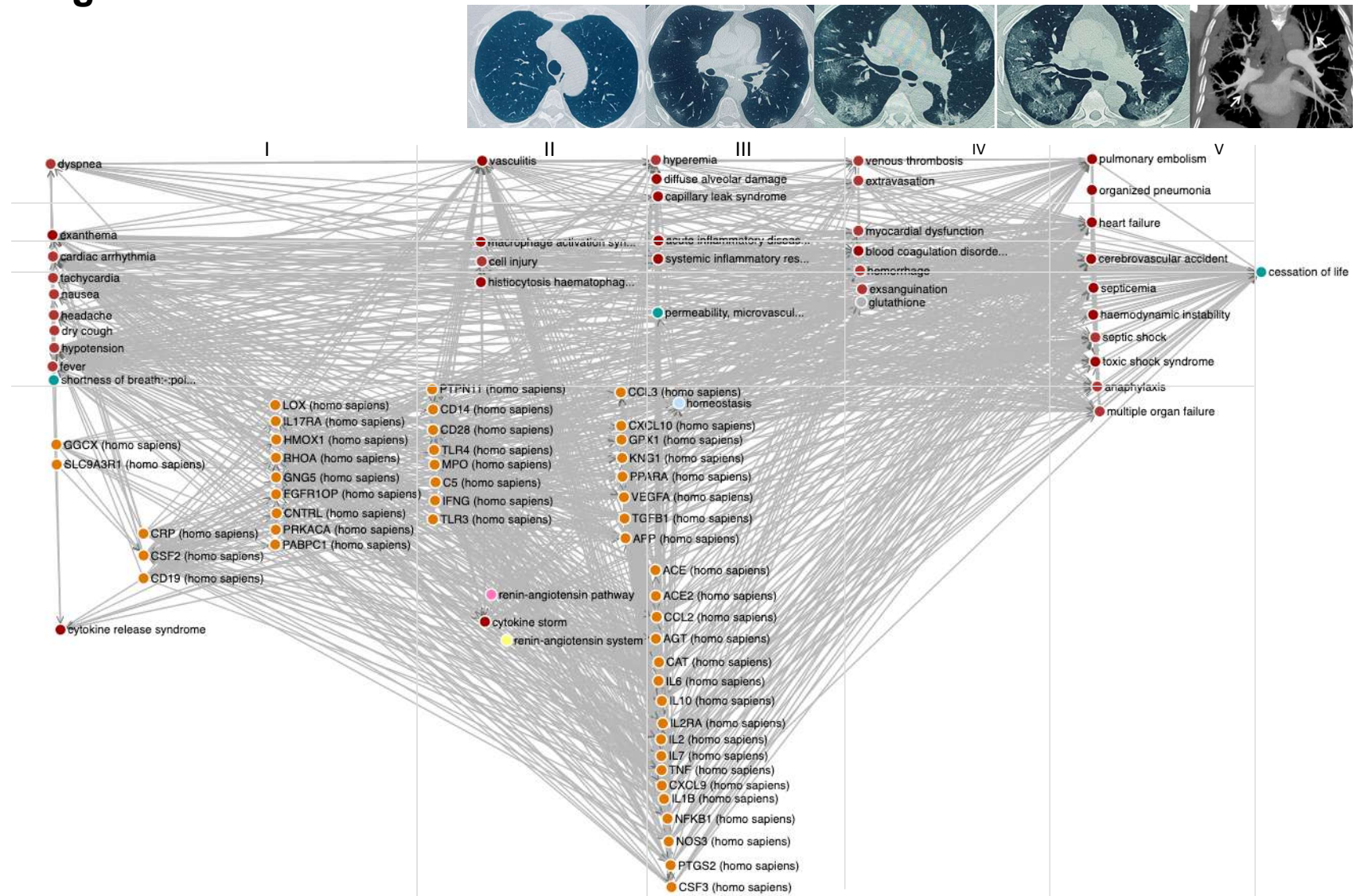




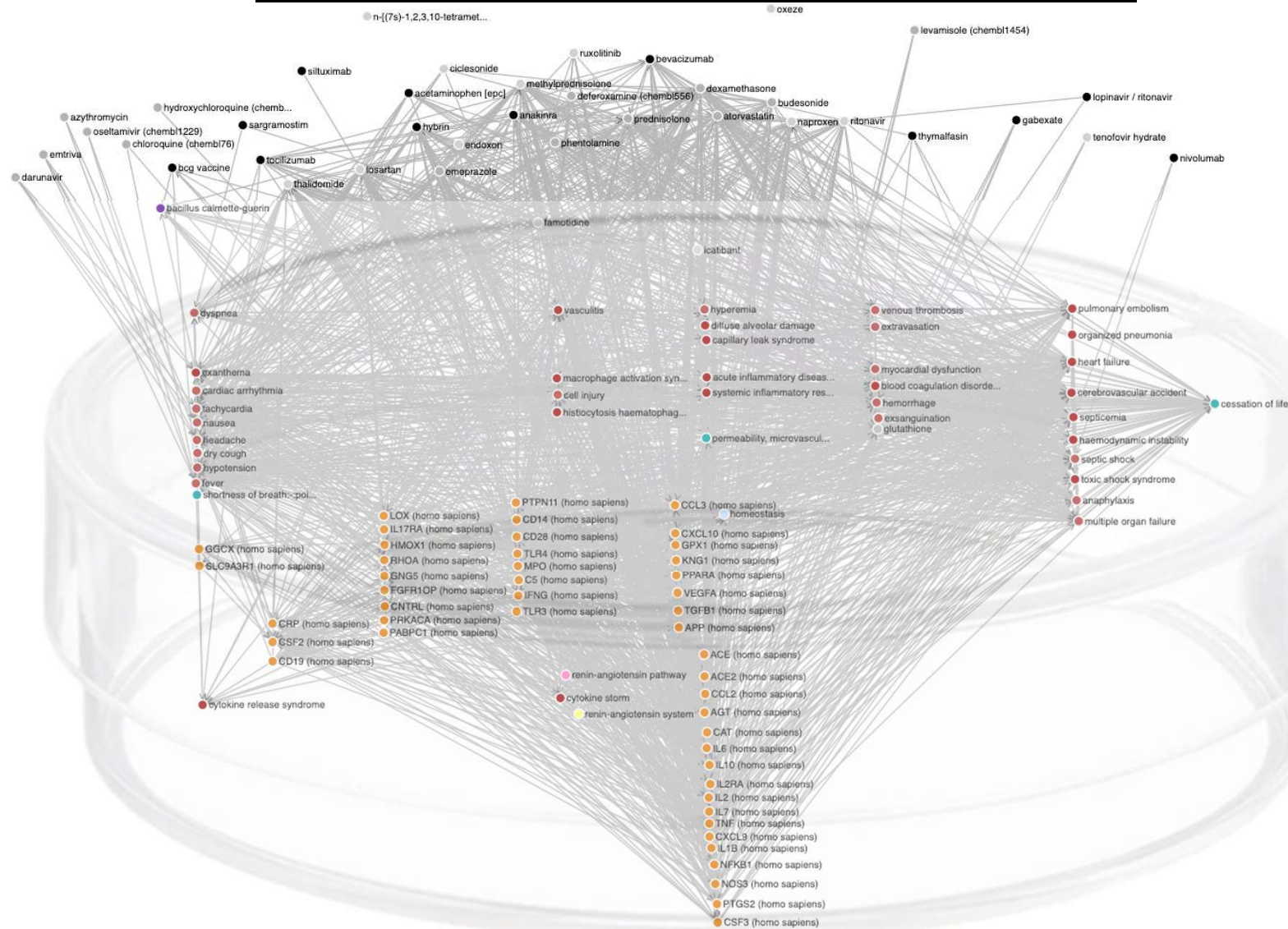
Combine
FANTOM5
& LOVD



Existing Drugs for COVID ?

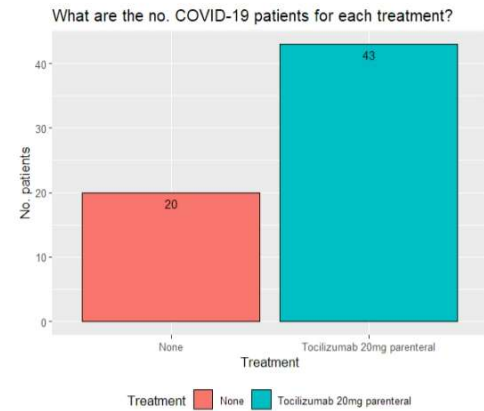


Machine-Assisted Science

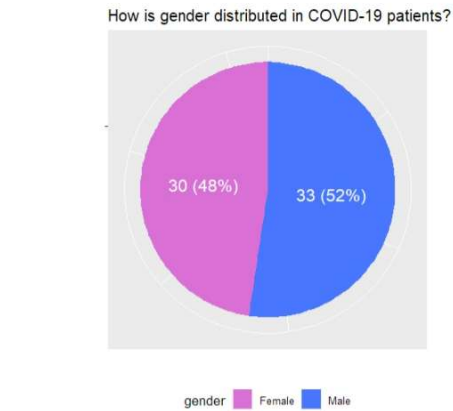


Specific data pulled from FDPs with patient data (RWO) and knowledge (EK)

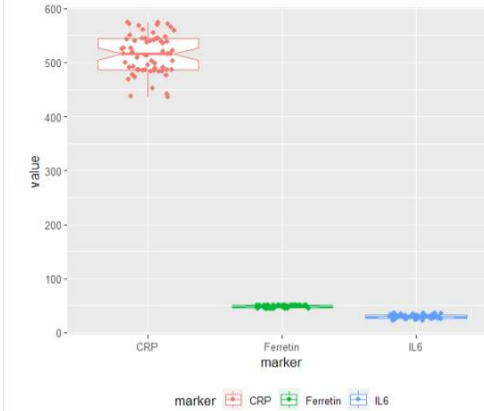
Barchart of Study Counts:



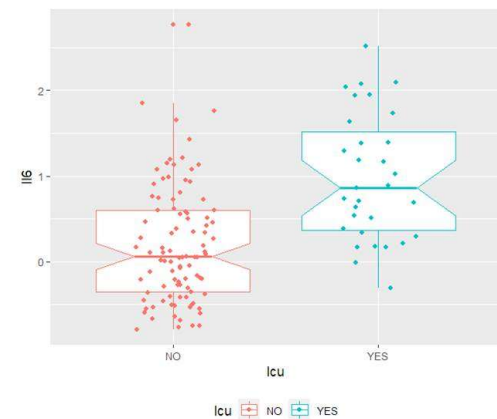
Piechart of Study Breakdown:



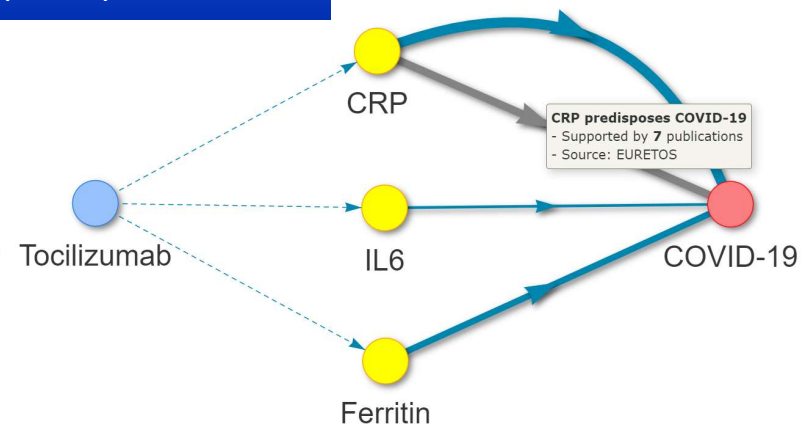
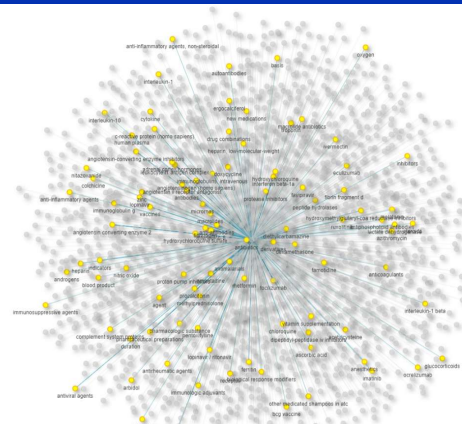
Boxplot of Markers:



Boxplot of Marker vs COVID Severity:



Specific visualization for each type of question



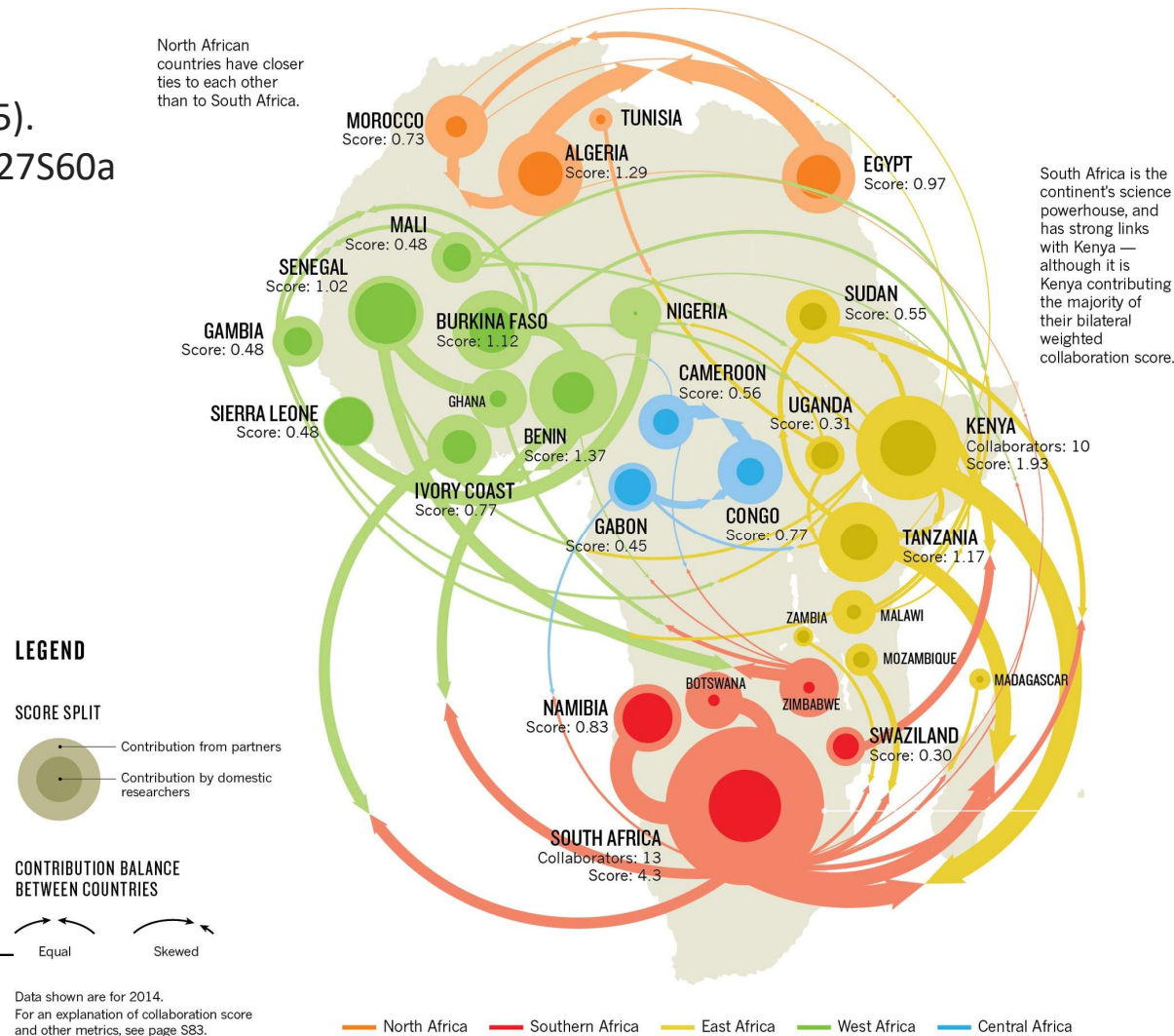
[← Home](#)

Personal Health Train – Distributed Learning Community

People and organizations interested in applying the Personal Health Train concept or distributed learning techniques to health-related topics are working together in the Personal Health Train network. The PHT concept enables health care innovators and researchers to work with health data from various sources.

DEVELOPING PARTNERSHIPS THAT MATTER

Nature **527**, S60–S63 (2015).
<https://doi.org/10.1038/527S60a>

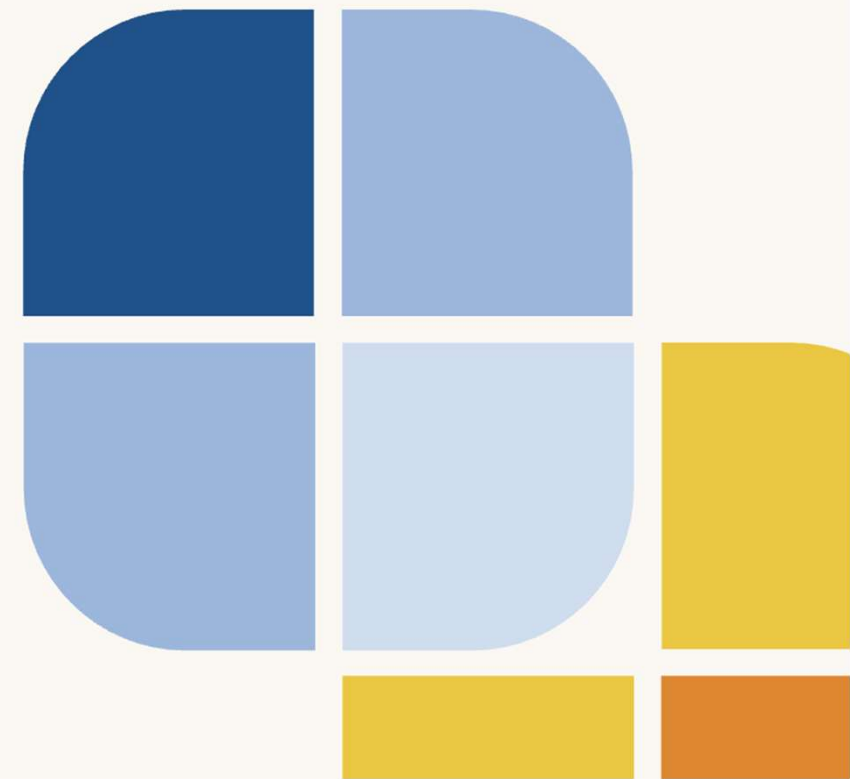




The **Leiden**
Initiative for **FAIR** &
Equitable Science

FAIR, equitable and sustainable data reuse

LIFES is a global association that unites public and private data owners, users, ASPs and communities to jointly evolve the ecosystem for FAIR, equitable and sustainable data reuse for science & innovation



Vision, Mission & Strategy

Vision

Global, **FAIR** and **Equitable Science** and **Innovation** as the new normal.

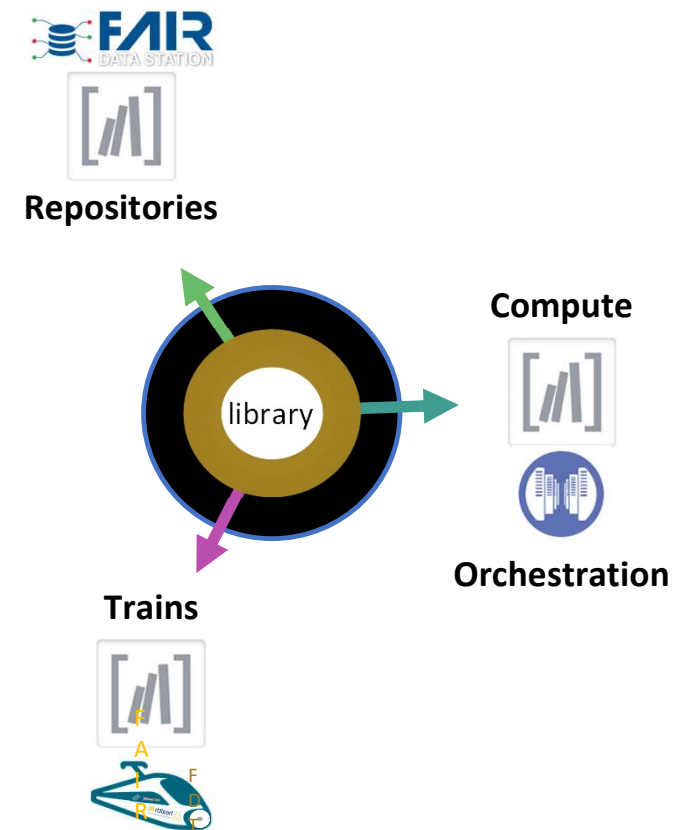
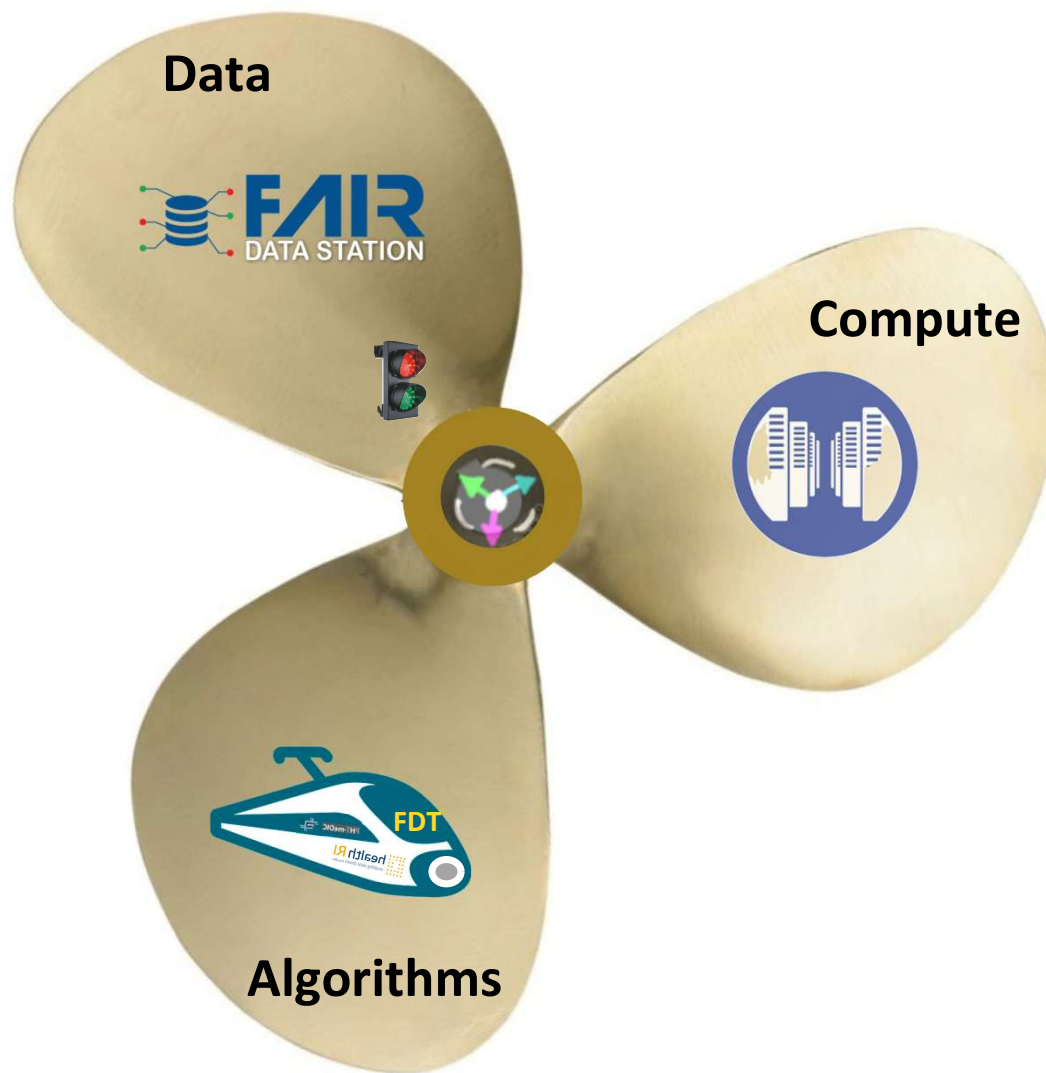
Mission

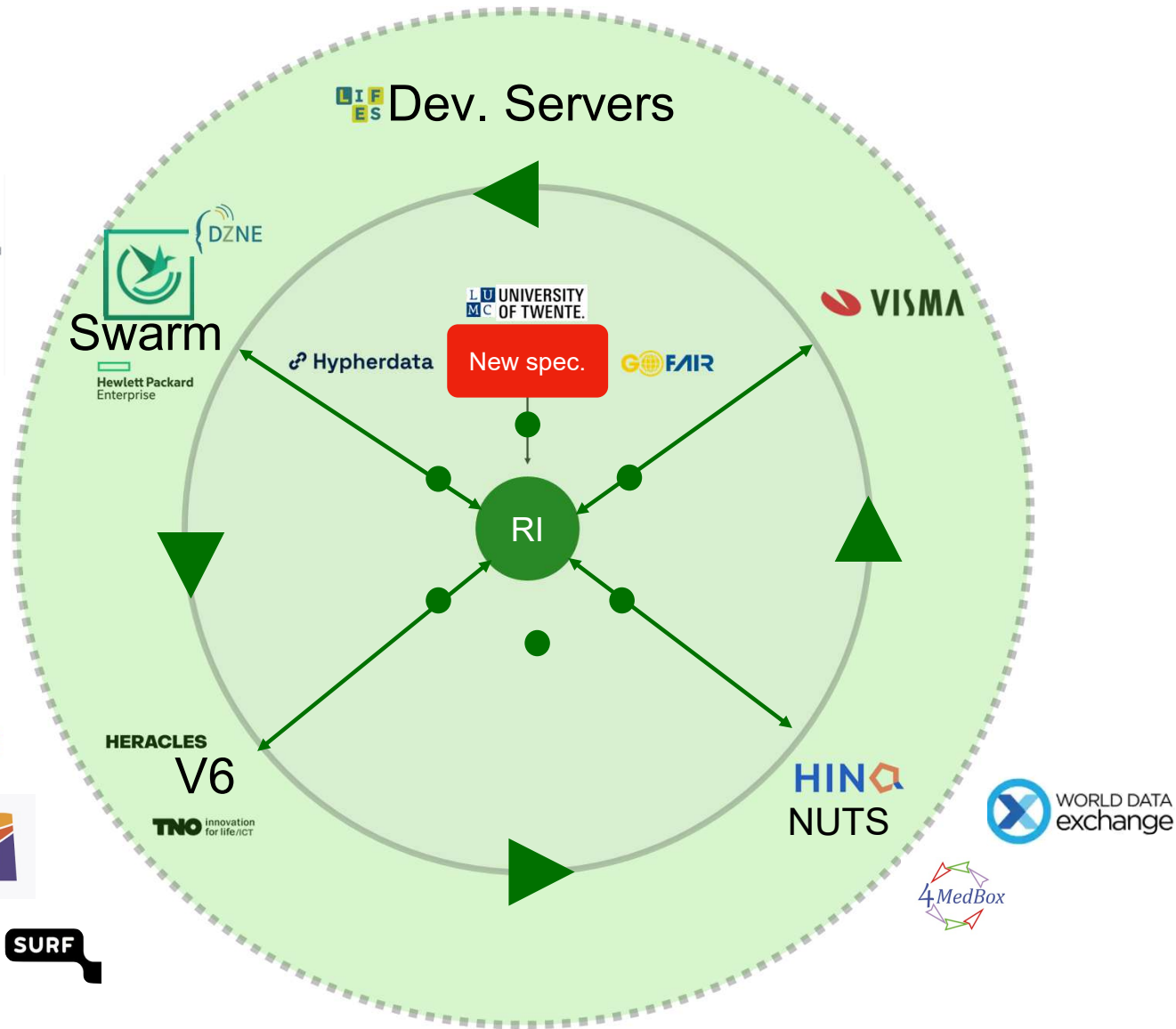
Facilitate **FAIR based**, **privacy preserving**, and **distributed** data **generation** as well as **equitable reuse**.

Strategy

Create and facilitate the Leiden Institute for FAIR and Equitable Science (LIFES) as a **public-private partnership** where FAIR data and services are the norm. LIFES will help develop and maintain the required **expertise**, **capabilities** and **ecosystem** for the equitable and privacy preserving reuse of data and services.

A Semantic FAIR library of instructions connects them all







Can SPHN move to data Visiting?