

# 3. LIFE SCIENCES CLUSTER BASEL LUNCH

**Deborah Strub, Abteilungsleiterin Cluster & Initiativen, Mitglied der  
Geschäftsleitung**

5. Juli 2022



# PROGRAMM

- ▶ Begrüssung
- ▶ Das Swiss Personalized Health Network (SPHN), ein smartes Framework für eine verantwortungsvolle und effiziente Sekundärnutzung von Gesundheitsdaten
- ▶ Lunch



# VIEL VERGNÜGEN

- ▶ **Dr. Sabine Österle**, Teamleiterin Daten-Interoperabilität, Personalized Health Informatics Group, SIB Swiss Institute of Bioinformatics



# A smart framework for a responsible and efficient secondary use of health data

## The Swiss Personalized Health Network (SPHN)

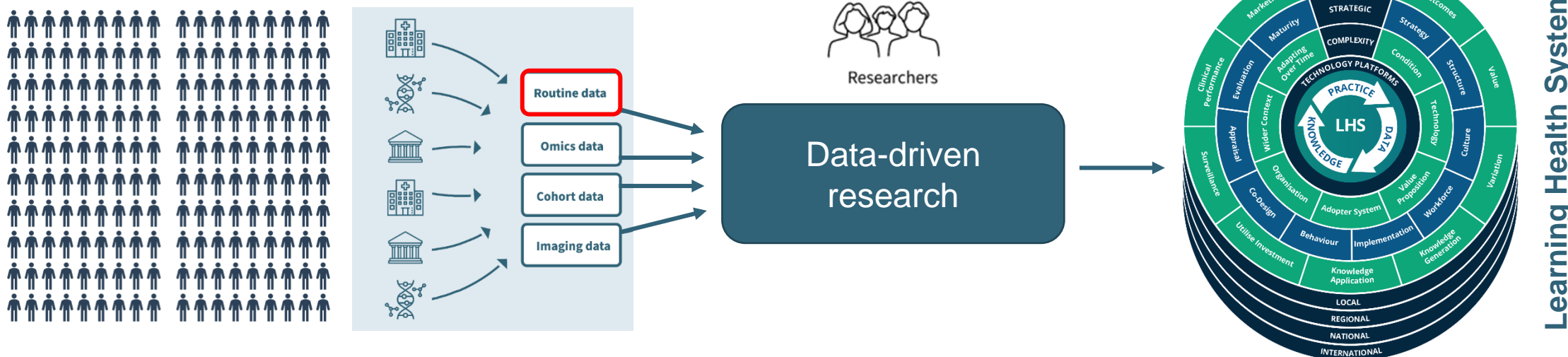
Dr. Sabine Österle, Team Lead Data Interoperability

SIB Personalized Health Informatics Group and SPHN Data Coordination Center

A project of

# The road to Personalized Health

Given the vast amounts of data captured in the healthcare setting, healthcare decisions should no longer be based on population averages, but should take into account individual patient characteristics.



Learning Health System

# Data types and data sources in SPHN

## Data from health care institutions

Routine data (EHR, RWD, med. registries) e.g. diagnoses, medication, vitals, procedures, lab values, imaging data, outcome data, etc.

→ not captured for research, widely unstructured, lacking standards, limited explanatory information

## Molecular and \*omics data

Genomics, epigenomics, transcriptomics, proteomics, metabolomics, biomarkers, etc.

→ no common standards, missing meta-data, clinical versus research grade

## Clinical research data

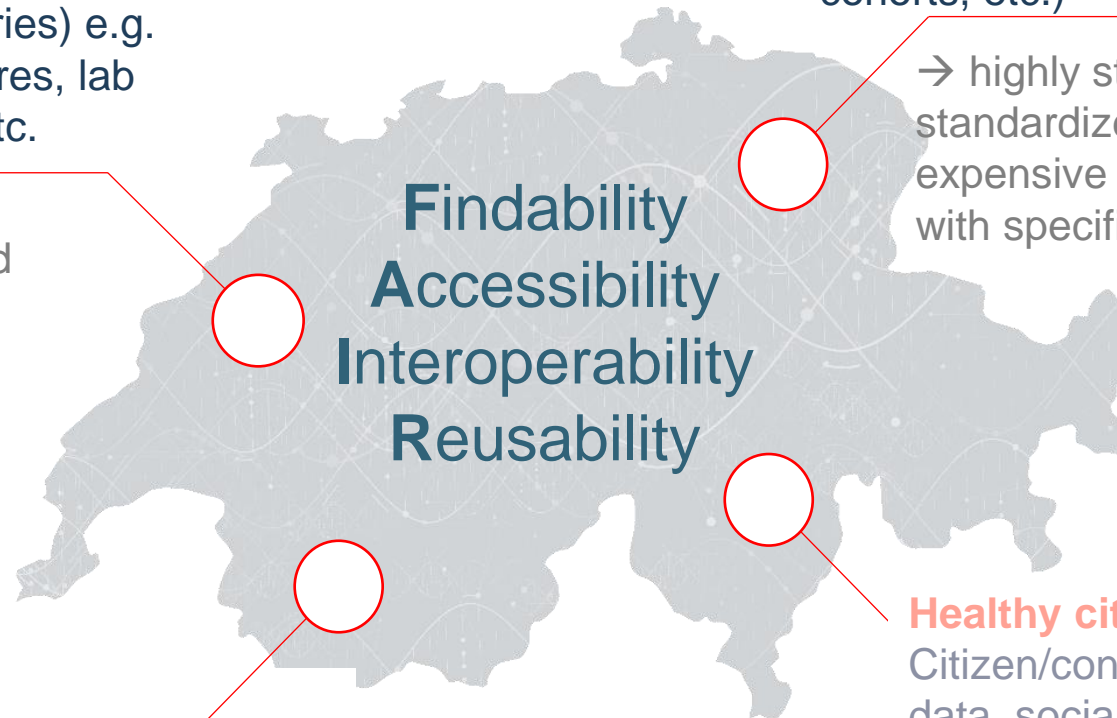
High-quality patient-oriented health data (studies, trials, public health registries, longitudinal cohorts, etc.)

→ highly structured, merely vertically standardized, no cross-use of routine data, expensive and complex collection, often with specific consent

## Healthy citizen data

Citizen/consumer health data, lifestyle data, social media data, wearable devices, etc.

→ Technical lock-in, no standards, lacking technical interoperability



# Tapping routine health care data for research purposes

## Challenges:

- Data management at the UH: Heterogeneity, data integration into Clinical data warehouses/ data lakes
- Interoperability of data: Semantics and standardization (at the source)
- Harmonization of data preparation (quality & standards) and delivery processes
- Data governance, data access (incl. willingness to share)
- Feasibility queries and findability / discoverability of data (in a federated way)
- Meta data catalogues (incl. imaging, multi-media files, etc.)

# Goals

→ Creation of a scalable and sustainable data-enabling environment (health data ecosystem)

→ Enabling researchers to access, integrate, analyze, and share interoperable health data



Manage  
 • Manager initiative  
 • Daily op



# SPHN Funding streams

## **Infrastructure Implementation** (top-down, 'horizontal', Leistungsvereinbarung):

- Collaboration Agreements with 5 University Hospitals
- HospFAIR
- TI4Health pilot

## **Driver/Development Projects** (bottom-up, 'vertical', project grants)

- Driver: 6 completed, 5 ongoing
- Development: 13 completed

## **National Data Streams** (bottom-up, 'vertical', project grants)

## **Demonstrator projects** (call open)

# National Data Streams (NDS)

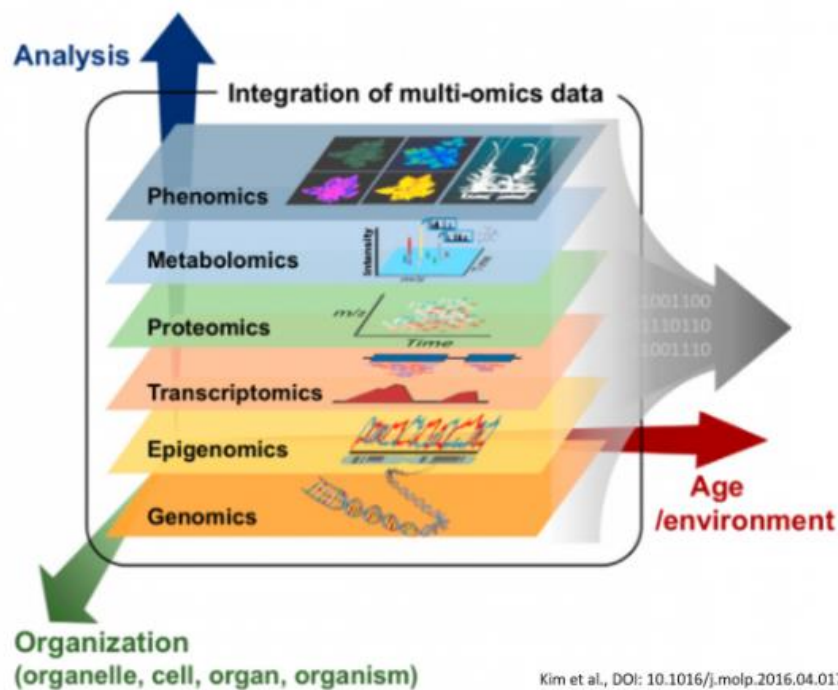
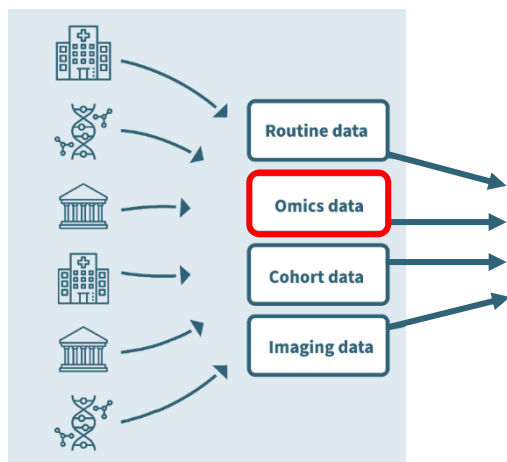
4 NDS starting in summer 2022:

- Personalized, data-driven prediction and assessment of **infection-related outcomes in Swiss ICUs (IICU)**
- Swiss Personalized **Oncology** National Data Stream (SPO-NDS)
- **Pediatric** personalized research network Switzerland (SwissPedHealth) – a Joint Pediatric National Data Stream
- **Low Value of Care** in Hospitalized Patients (LUCID) – a National Data Stream on Quality of Care in Swiss university hospitals

# The biological layer of Personalized Health



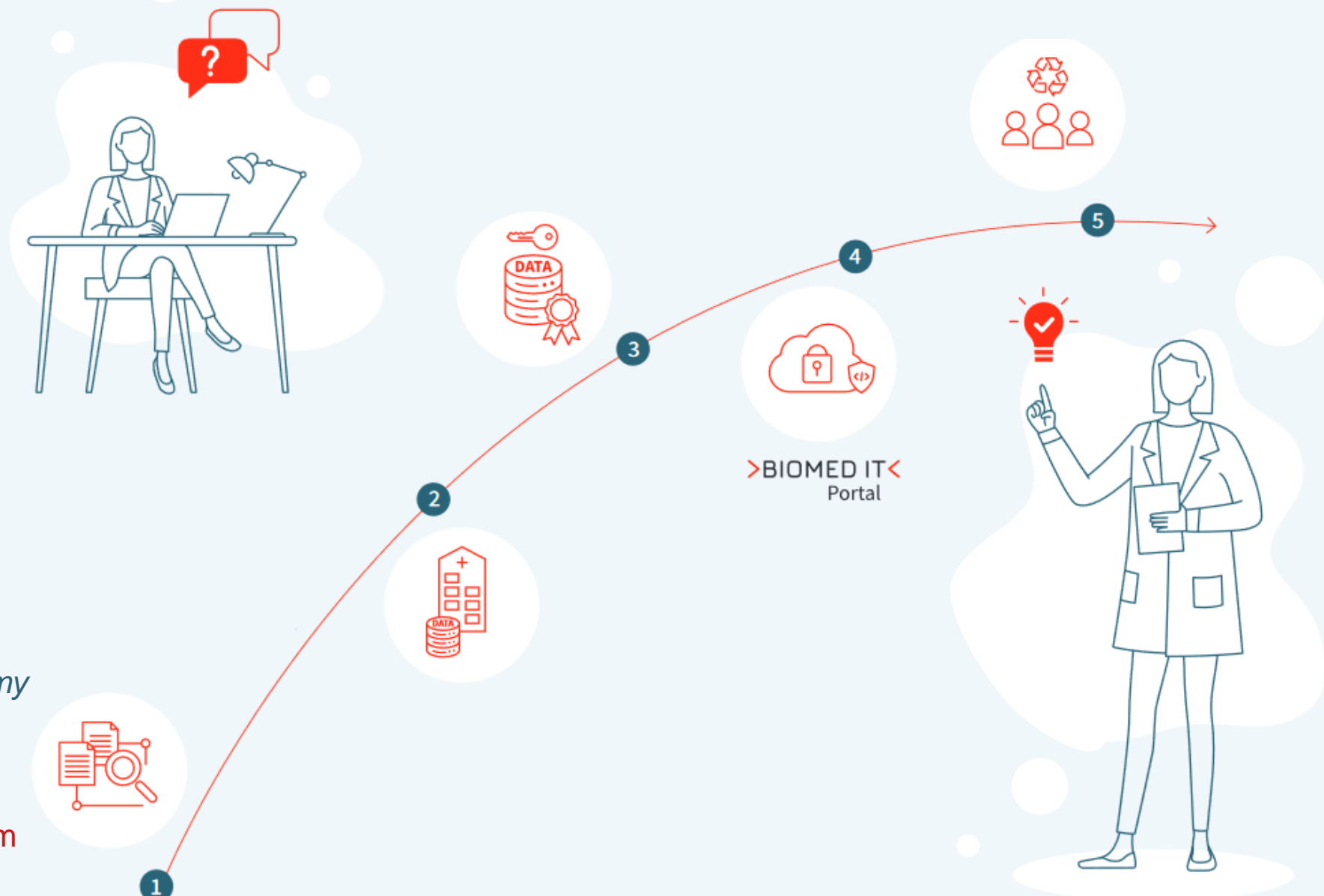
Strategic Focus Area  
**Personalized Health and Related Technologies**



Genomics	Proteotyping	Metabolomics & Lipidomics
<b>CGAC</b> Clinical Genomic Analysis Center	<b>CPAC</b> Clinical Proteotype Analysis Center	<b>CMAC</b> Clinical Metabolomics Analysis Center
Clinical Grade Sequencing (ISO 15189 accredited)	Quantitative Proteotyping	Small molecule analysis
Whole Genome Sequencing (WGS)	Proteotype analysis (DDA, DIA, PRM)	Targeted metabolomics
Whole Exome Sequencing (WES)	Post-translational modification analysis	Untargeted metabolomics
RNA Sequencing (RNASeq)	Spatial proteotype analysis	Lipidomics
<a href="#">Further information and pricing</a>	<a href="#">Further information and pricing</a>	<a href="#">Further information and pricing</a>
Fast-track genomic data analysis	Fast-track proteomic data analysis	Fast-track metabolomic data analysis
<b>Integrated Multi-Omic Data Interpretation</b>		
Integration, visualization, and analysis of omics data mapped onto biological networks and pathways		
BioMedIT integration for data lineage tracking, data management, data sharing, secondary and tertiary analysis		

ETH PHRT Swiss Multi-Omics Center: <http://smoc.ethz.ch>

# The SPHN value chain



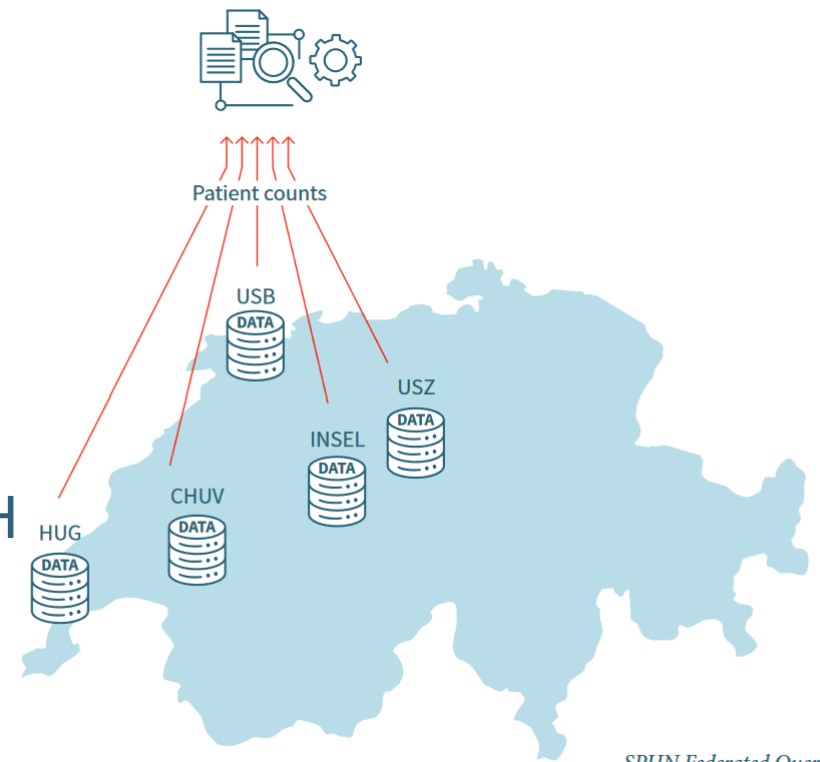
- 1) Project planning**  
*What data is available for my research project?*
- Data Management Plan (DMP) Guidelines
  - **Federated Query System**

# Feasibility – the SPHN Federated Query System

<https://www.youtube.com/watch?v=Jj-wLIFeWNo>

Allow researchers to assess whether and where patients or patient data potentially suitable for a specific research question exists

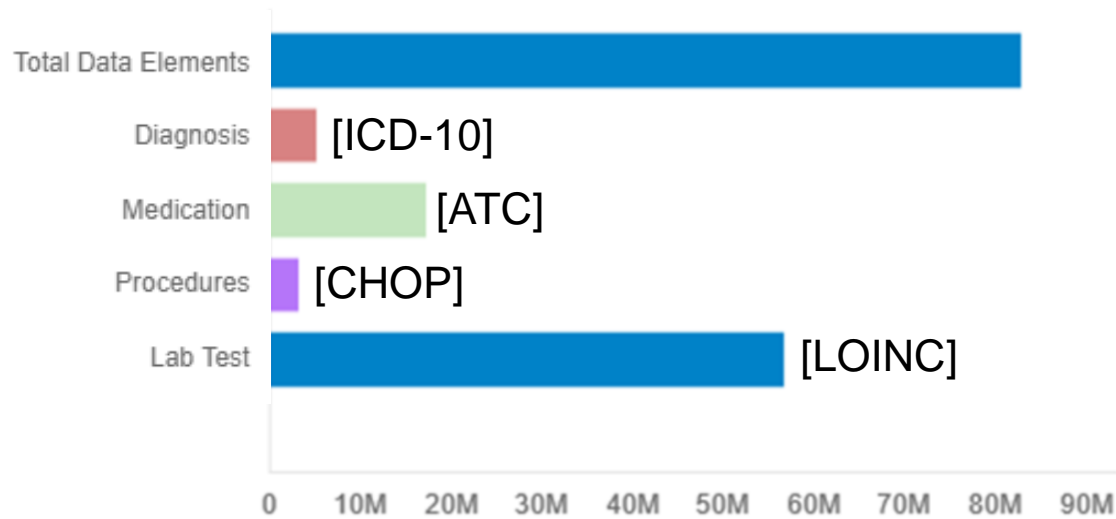
- Currently accessible to researchers from the 5 UH
- Soon available to biomedical researchers from all swissuniversities research institutions



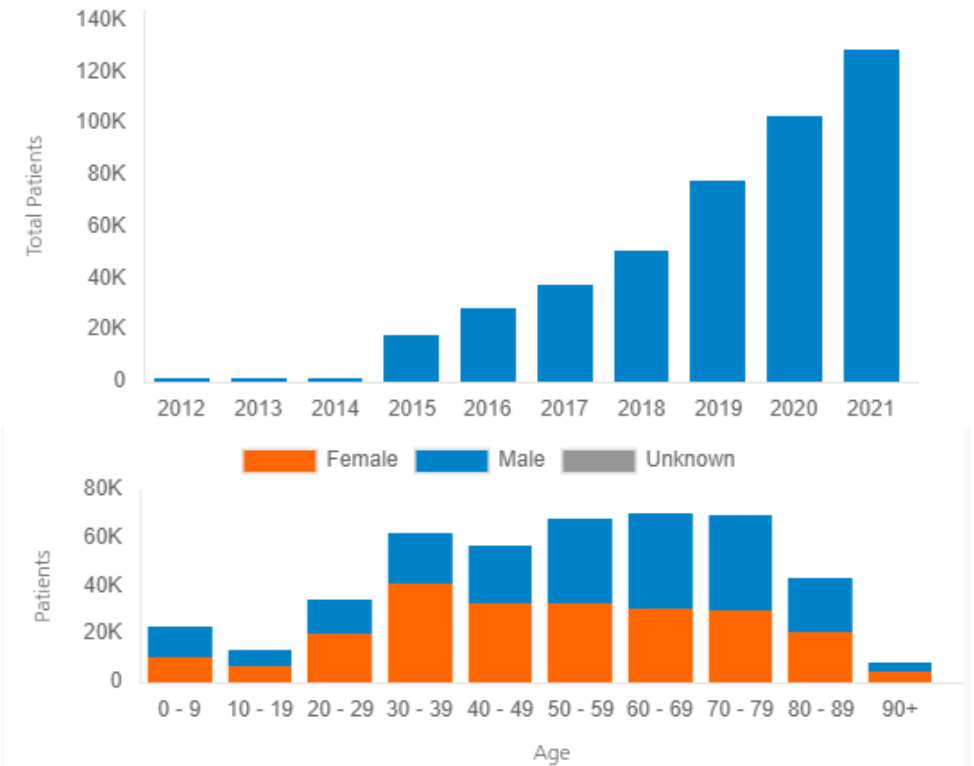
# Feasibility – the SPHN Federated Query System

Data from the Clinical Data Warehouses of all five Swiss university Hospitals

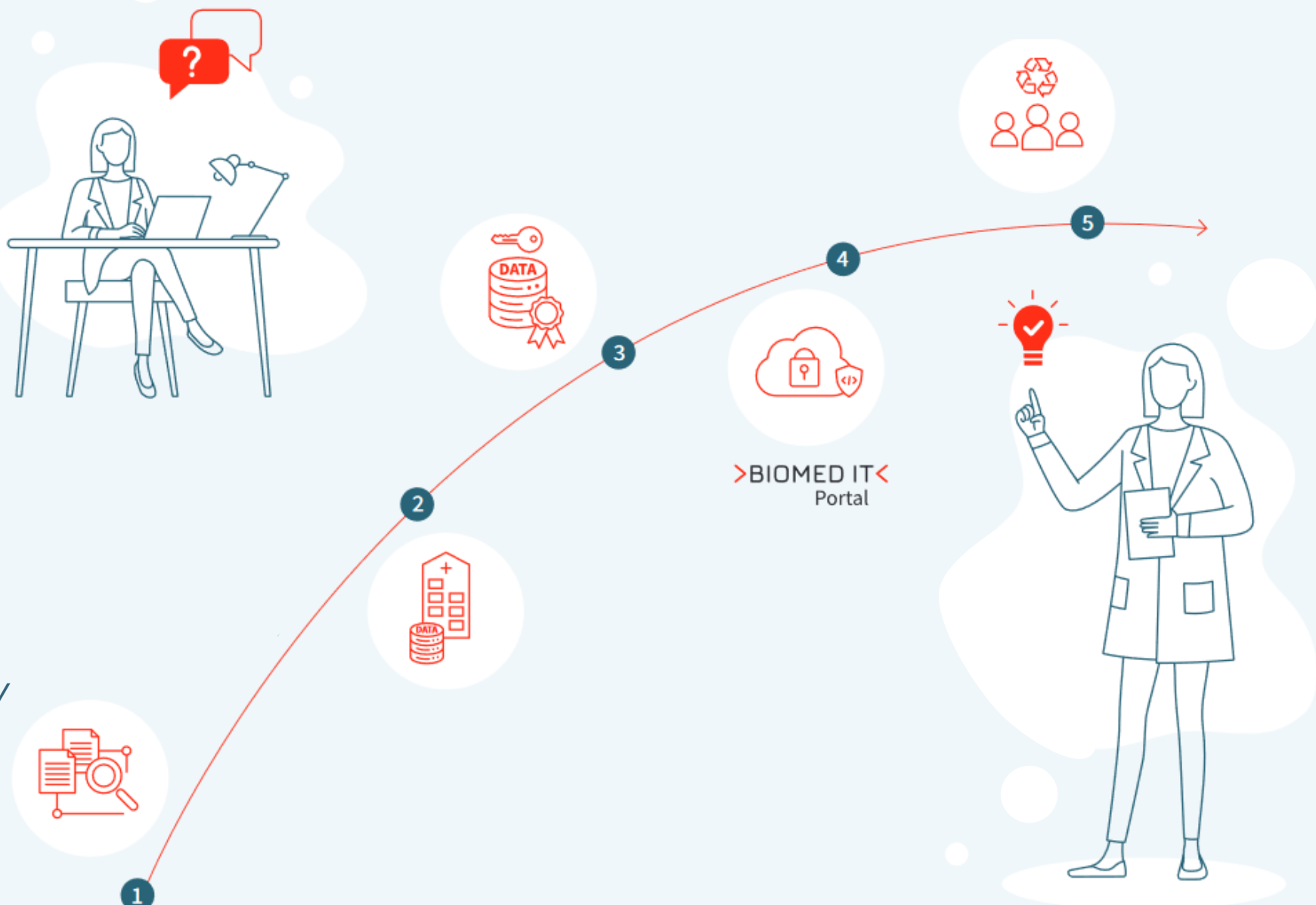
Total ~95'000'000 data elements



Total 495'000 patients (with general consent)



# The SPHN value chain



## 1) Project planning

*What data is available for my research project?*

- Data Management Plan (DMP) Guidelines
- Federated Query System
- **Metadata Catalogues**

# Findability of cohort data

## Swiss Personalized Health Network Cohort Consortium



Switzerland can rely on a number of high quality prospective cohort studies ranging from population-based to nation-wide patient-based studies, some of which span decades. Following the increasing demand for the implementation of collaborative environments, the Swiss Personalized Health Network ...

10	9	145,935
Individual Studies	Individual Studies with Variables	Individual Study Variables



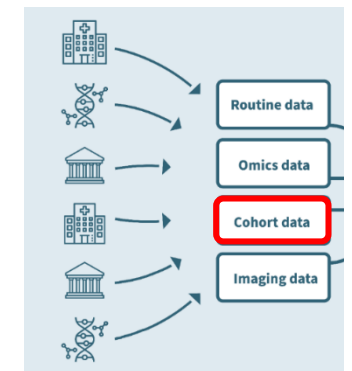
<https://www.maelstrom-research.org/network/sphn-cc>

## Maelstrom Catalogue:

- Powerful toolkit to improve documenting study metadata from around the world
- >310 individual studies, 23 networks covering a wide range of research areas

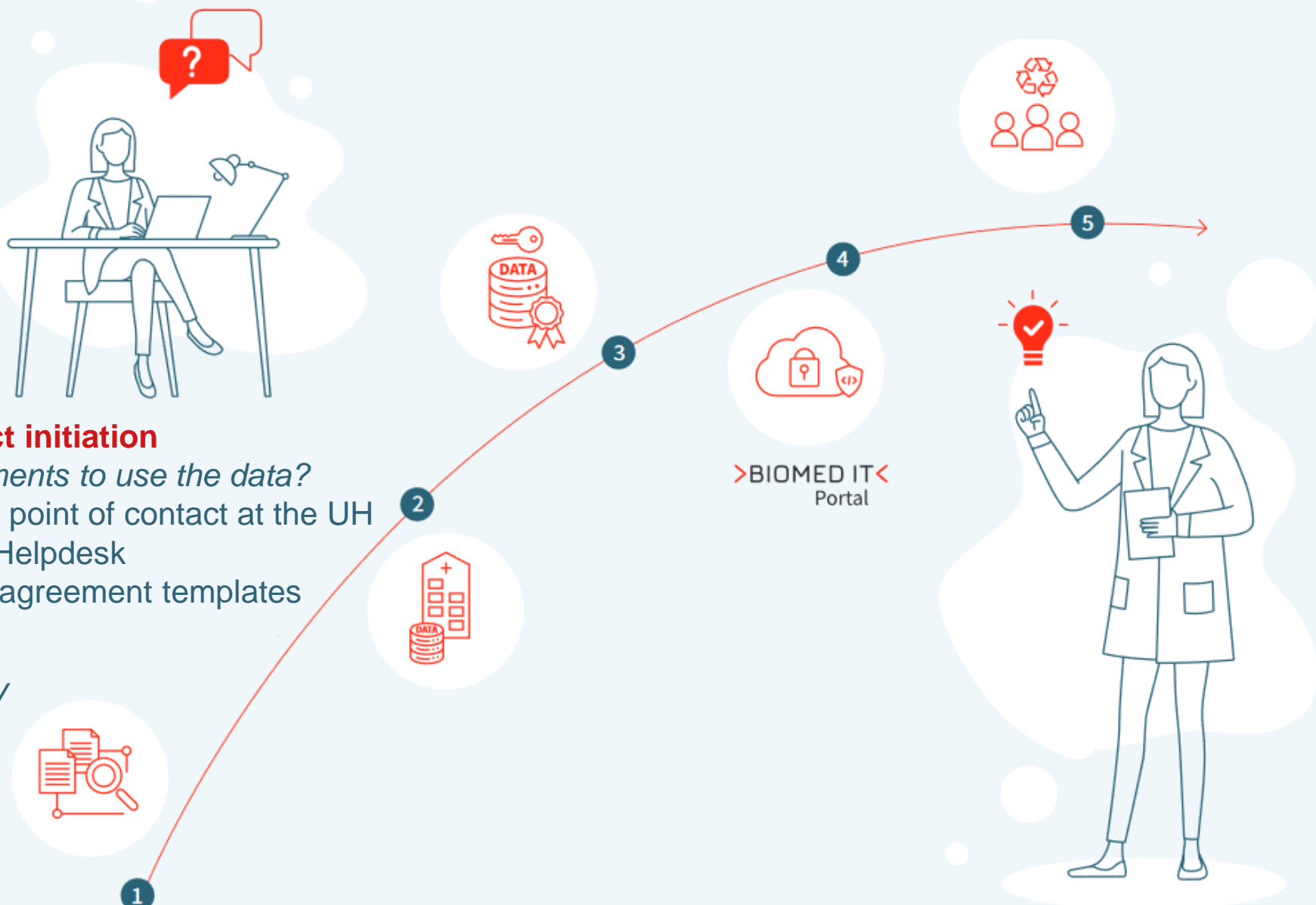
## SPHN Cohort Consortium:

- 10 high-quality population-based and disease-specific cohort studies
- Search full study descriptions and query >146'000 annotated variables





# The SPHN value chain



## 2) Project initiation

Requirements to use the data?

- Single point of contact at the UH
- ELSI-Helpdesk
- Legal agreement templates

## 1) Project planning

What data is available for my research project?

- Data Catalogues
- Federated Query System
- Data Management Plan (DMP) Guidelines

# The SPHN value chain

## 1) Project planning

*What data is available for my research project?*

- Data Catalogues
- Federated Query System
- Data Management Plan (DMP) Guidelines



1

## 2) Project initiation

*Requirements to use the data?*

- Single point of contact at the UH
- ELSI-Helpdesk
- Legal agreement templates



2



## 3) Data preparation

*Is the data available and interoperable?*

- Established data clinical data warehouse and delivery pipelines at university hospitals



3



>BIOMED IT<  
Portal

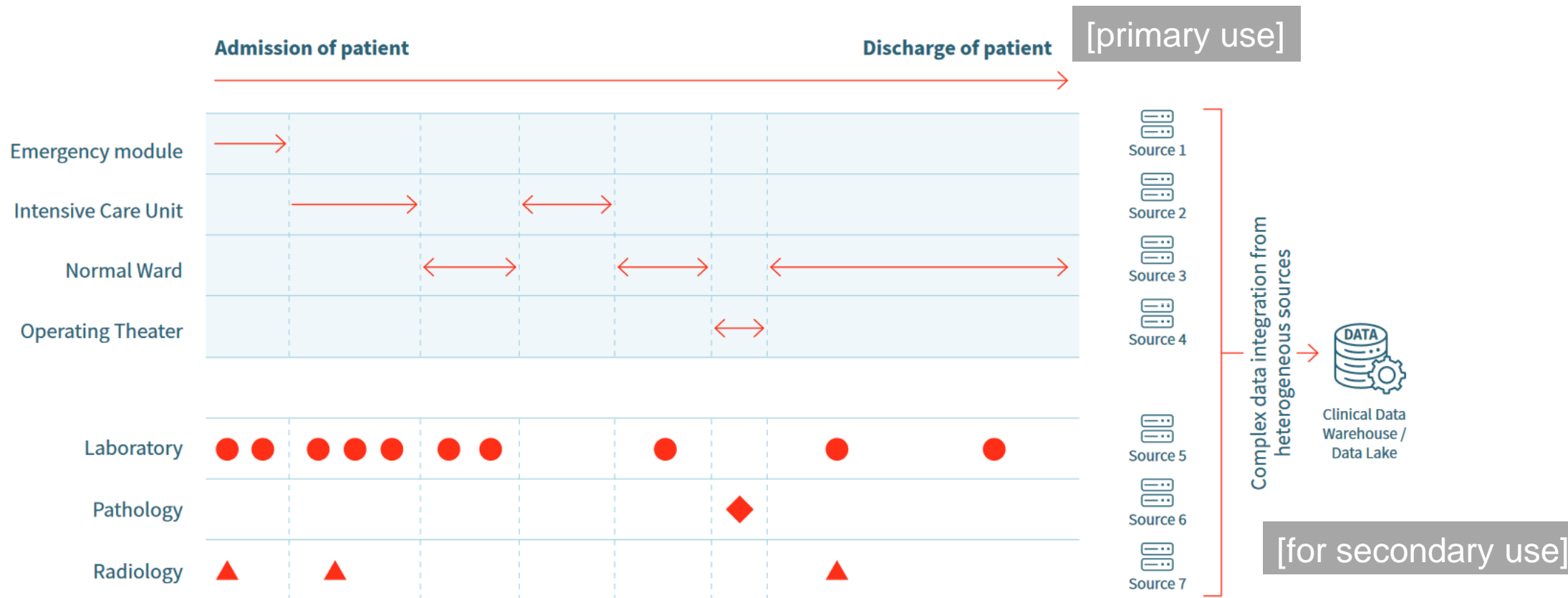
4



5



# Routine health care data capture in hospitals



# The SPHN value chain

## 1) Project planning

*What data is available for my research project?*

- Data Catalogues
- Federated Query System
- Data Management Plan (DMP) Guidelines



1

## 2) Project initiation

*Requirements to use the data?*

- Single point of contact at the UH
- ELSI-Helpdesk
- Legal agreement templates



2



## 3) Data preparation

*Is the data available and interoperable?*

- Established data clinical data warehouse and delivery pipelines at university hospitals
- SPHN Interoperability Framework ensures interoperability and adds value to the data



3



4

>BIOMED IT<  
Portal



5



# Diversity of clinical routine data

Does the patient have fever? Yes

Fieber

Der Patient hat eine  
stark erhöhte Temperatur

Pyrexia

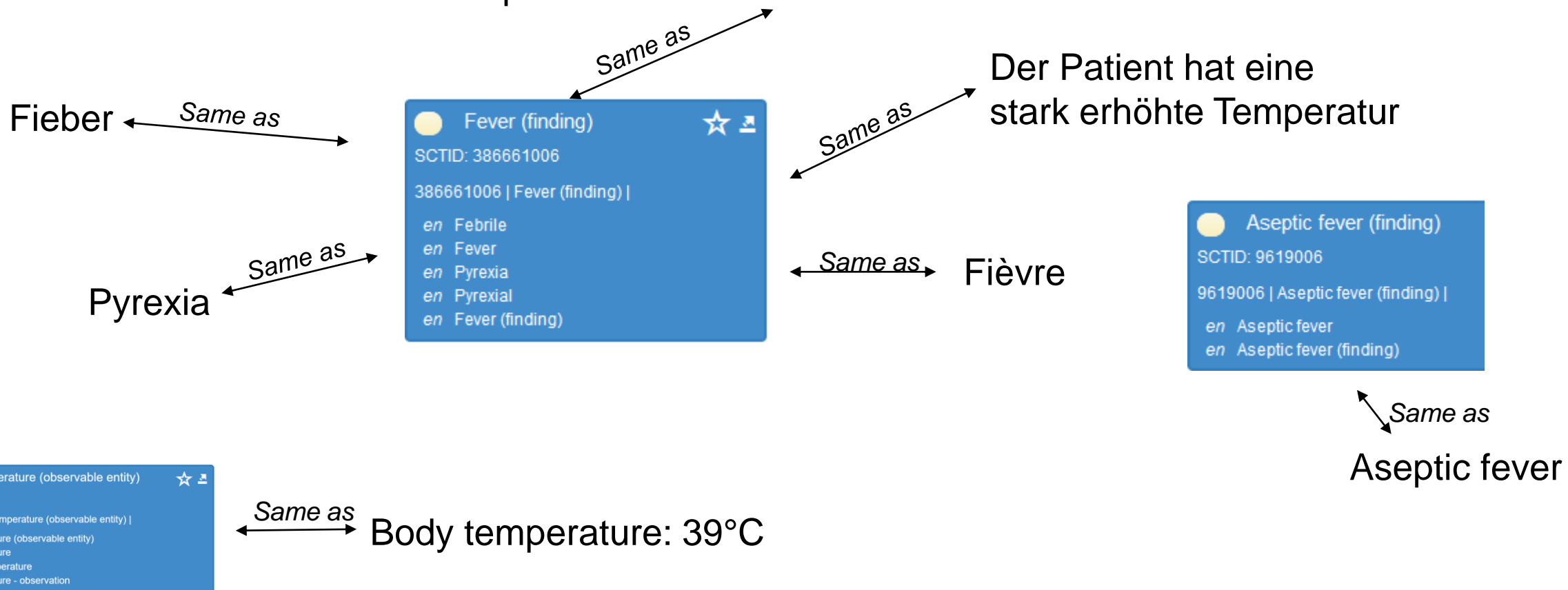
Fièvre

Aseptic fever

Body temperature: 39°C

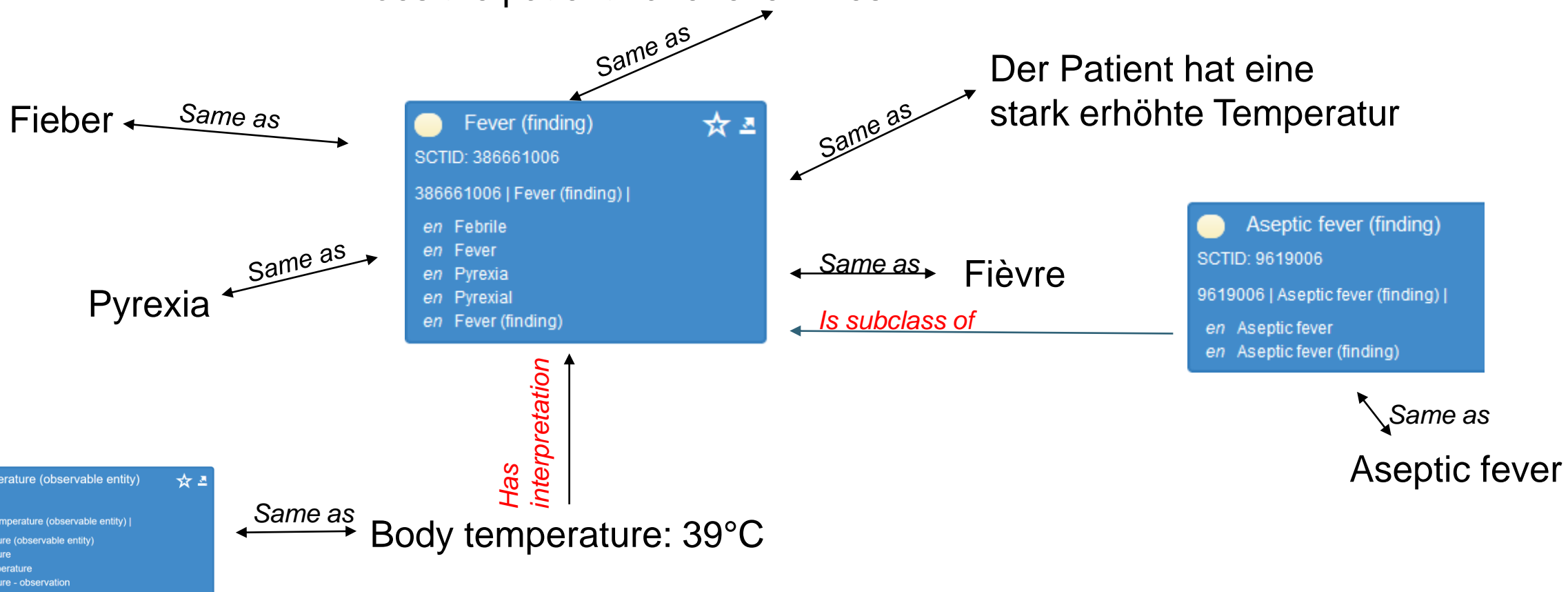
# Semantic description of the data

Does the patient have fever? Yes



# Semantic description of the relations

Does the patient have fever? Yes



# Making sense of the data

Clinical routine data

Other health-related data



## SPHN Interoperability Framework

### Semantics

- Use **controlled vocabulary** for concepts, valuesets and data
- **Standards:** ATC, SNOMED CT, ICD-10-GM, CHOP, UCUM, LOINC, NANDA, ICD-O, MedDRA...
- Multi-level coding and mappings

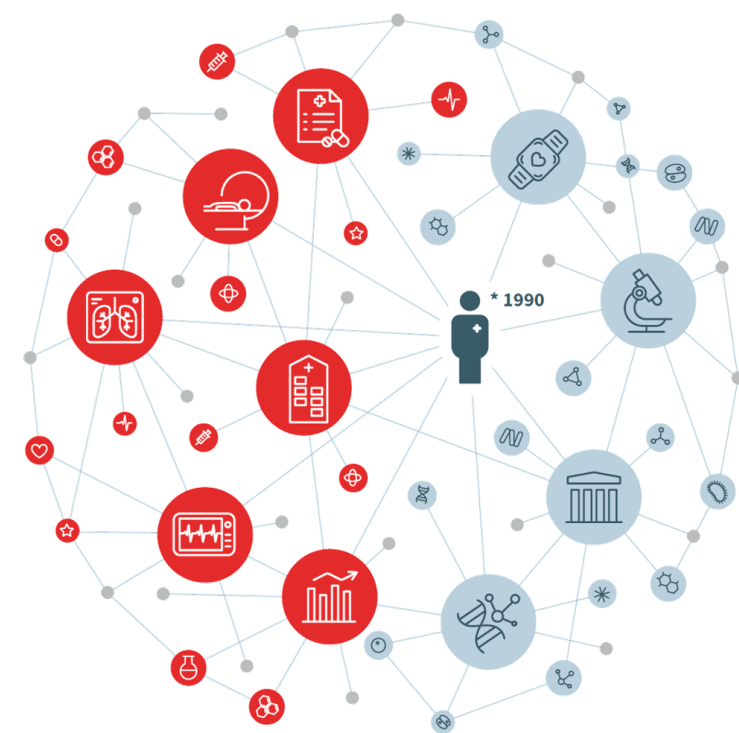
### Language for knowledge representation

- **Semantic Web** (RDF Schema)

### Quality control framework

- **Semantic Web** (SHACL, SPARQL)

→ **FAIR research data**





# The SPHN value chain

## 1) Project planning

*What data is available for my research project?*

- Data Catalogues
- Federated Query System
- Data Management Plan (DMP) Guidelines



1

## 2) Project initiation

*Requirements to use the data?*

- Single point of contact at the UH
- ELSI-Helpdesk
- Legal agreement templates



2



## 3) Data preparation

*Is the data available and interoperable?*

- Established data clinical data warehouse and delivery pipelines at university hospitals
- SPHN Interoperability Framework ensures interoperability and adds value to the data
- Deidentification of data



3



4

>BIOMED IT<  
Portal

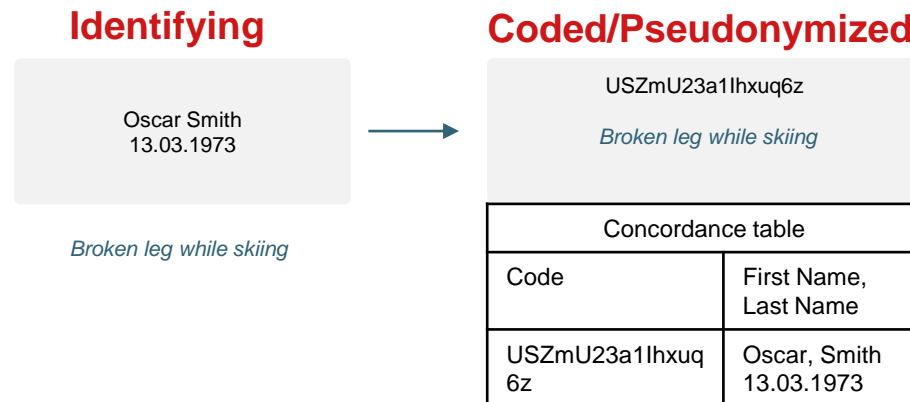


5

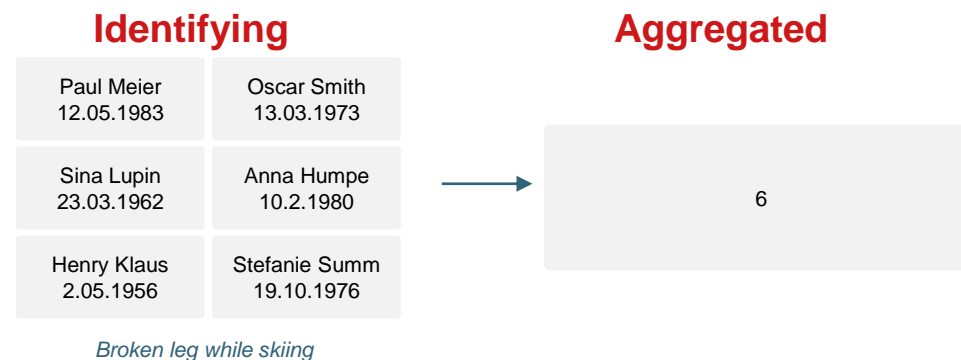


# When is personal data sensitive and when not?

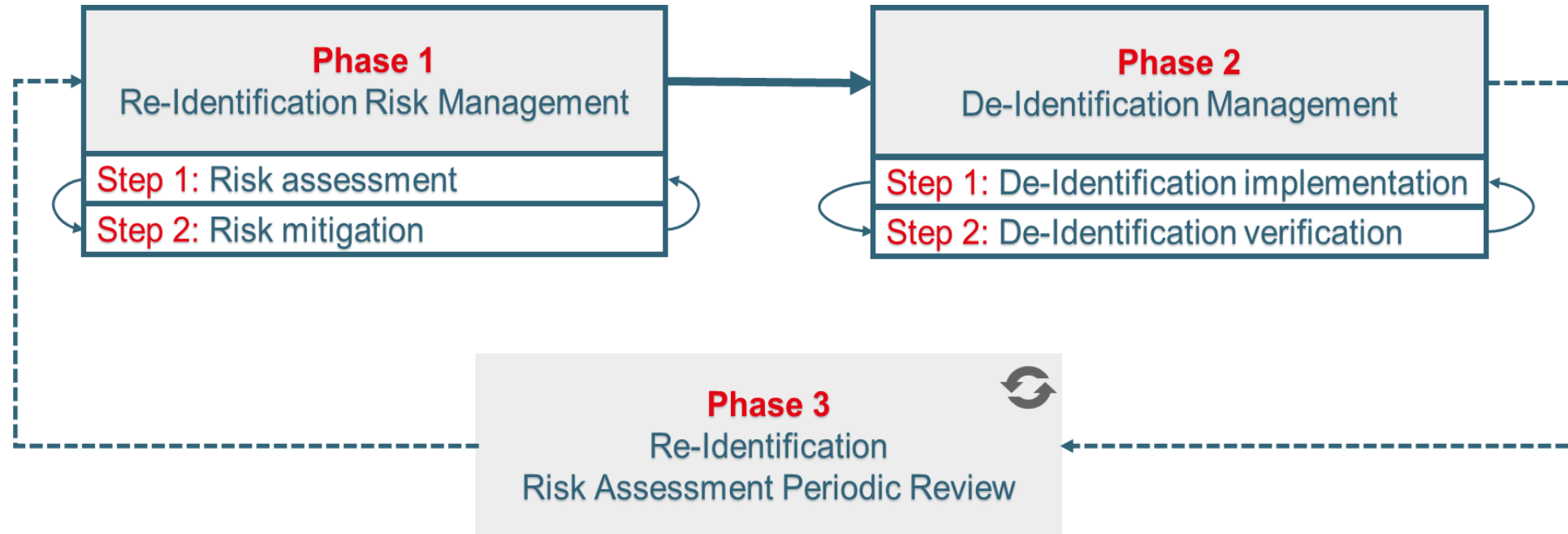
**Pseudonymized/coded** data is sensitive and personal data



**Aggregated** data is not sensitive and not personal data



# De-identification of health data in compliance with Swiss legislation



Phase 1 comprises the re-identification risk management assessing and mitigating patients' re-identification risk

Phase 2 contains the implementation and verification of risk mitigation actions specified in phase 1

Phase 3 describes the periodic review of the risk assessment performed according to project specifications

# The SPHN value chain

## 1) Project planning

*What data is available for my research project?*

- Data Catalogues
- Federated Query System
- Data Management Plan (DMP) Guidelines



1

## 2) Project initiation

*Requirements to use the data?*

- Single point of contact at the UH
- ELSI-Helpdesk
- Legal agreement templates



2



## 3) Data preparation

*Is the data available and interoperable?*

- Established data clinical data warehouse and delivery pipelines at university hospitals
- Deidentification of data
- SPHN Interoperability Framework ensures interoperability and adds value to the data



3



>BIOMED IT<  
Portal

## 4) Data transfer and analysis:

*Where can I analyze sensitive data?*

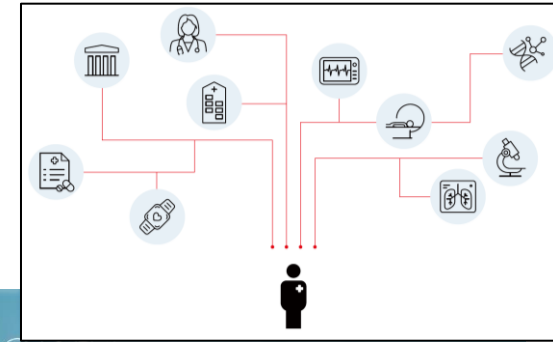
- Entering the BioMedIT Network through the BioMedIT Portal
- Secure transfer via sett
- Link to analytical platforms and biobanks
- B-spaces to analyze data



5

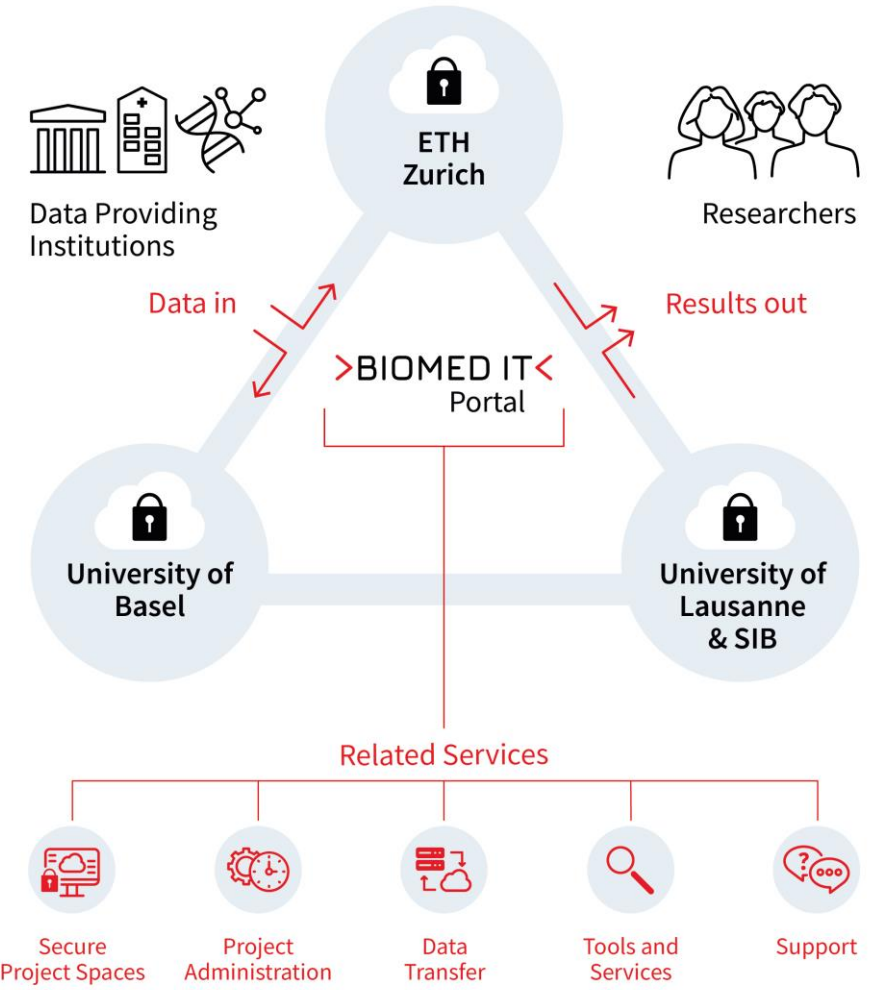
# Systematic research use of large amounts of health data: What does it take?

- Strong capabilities in clinical bioinformatics, computational biology, and computational service infrastructure
- High-performance IT infrastructures for big data computing and storage
- Secure data mobilization
- Security measures for ICT systems to protect confidential information from unauthorized access, unauthorized use, etc.

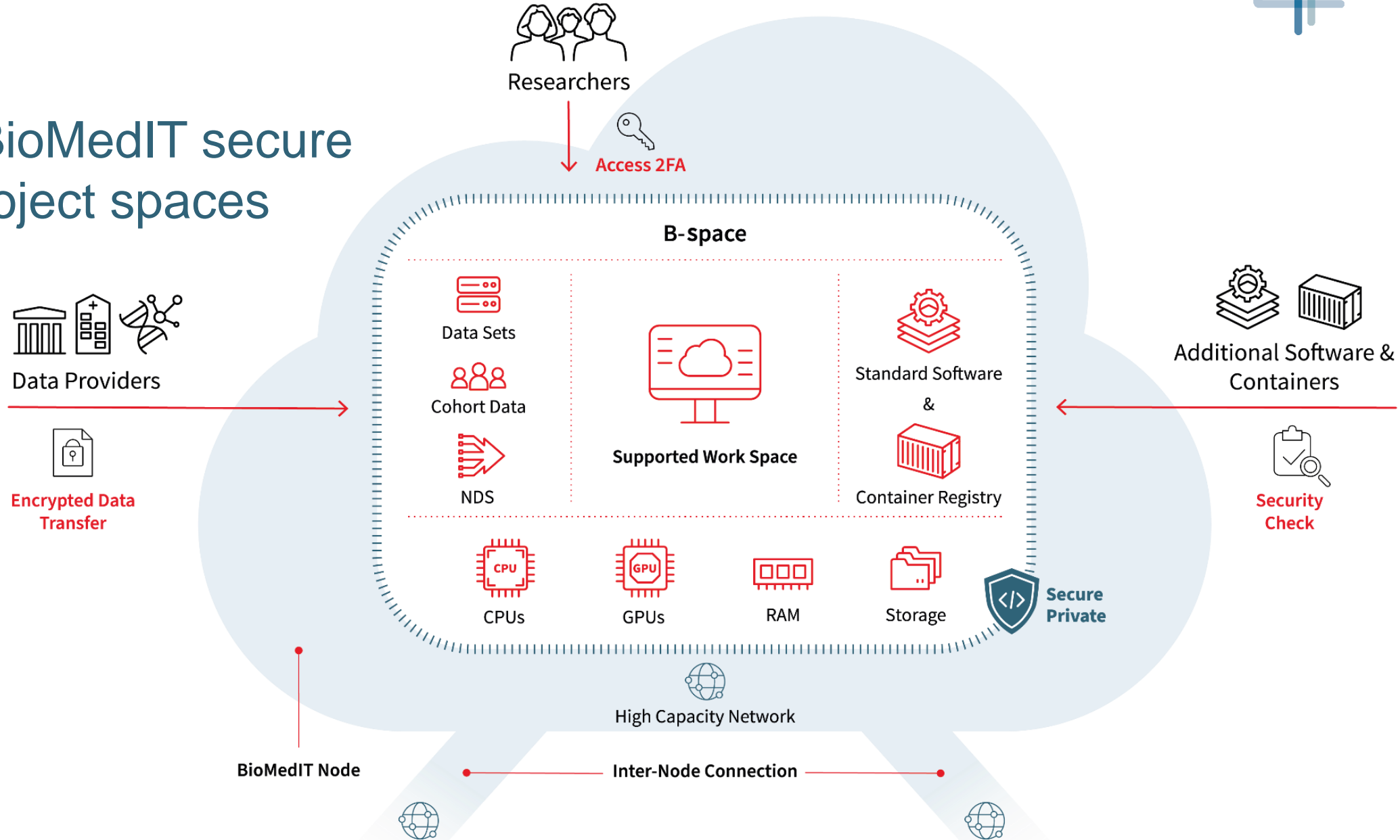


# The BioMedIT network

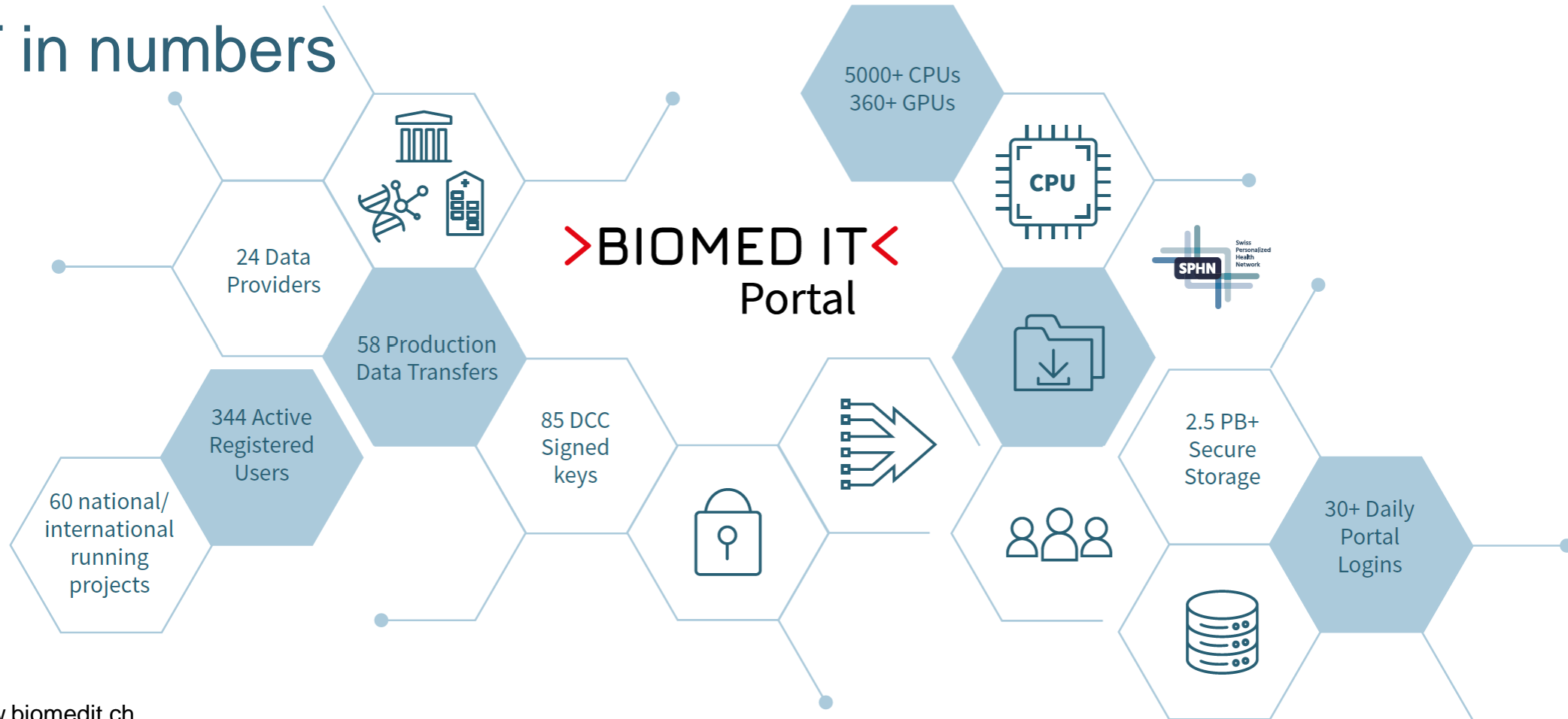
- Switzerland’s secure environment for research using sensitive data
- Allows for the mobilization and processing of health data, as well as leveraging it for research purposes
- Encrypted data comes into the environment, data remains on BioMedIT, only non-sensitive data (results) are taken out



# B-spaces: BioMedIT secure research project spaces



# BioMedIT in numbers



<https://www.biomedit.ch>



# The SPHN value chain

## 1) Project planning

*What data is available for my research project?*

- Data Catalogues
- Federated Query System
- Data Management Plan (DMP) Guidelines



1

## 2) Project initiation

*Requirements to use the data?*

- Single point of contact at the UH
- ELSI-Helpdesk
- Legal agreement templates



2



## 3) Data preparation

*Is the data available and interoperable?*

- Established data clinical data warehouse and delivery pipelines at university hospitals
- Deidentification of data
- SPHN Interoperability Framework ensures interoperability and adds value to the data



3



>BIOMED IT<  
Portal

## 4) Data transfer and analysis:

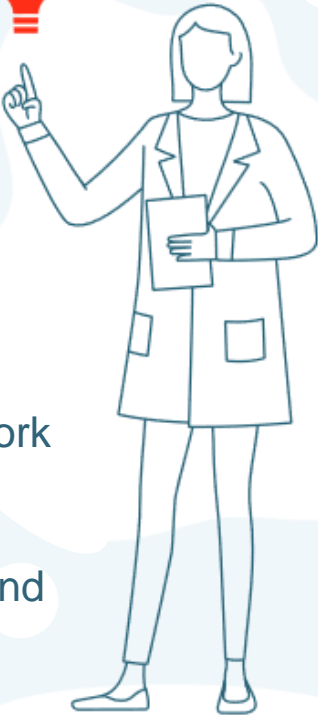
*Where can I analyze sensitive data?*

- Entering the BioMedIT Network through the BioMedIT Portal
- Secure transfer via sett
- Link to analytical platforms and biobanks
- B-spaces to analyze data

## 5) Data reuse

*How can I make data accessible to other researchers?*

- FAIR data repositories
- SPHN training modules help discovering data sources, tools and services



5

# Key contributions of SPHN

- Development of Clinical Data Management Systems in all 5 University Hospitals
- Funding of projects to drive the infrastructure development
- Interoperability framework for FAIR health data
- Secure IT environment for research (BioMedIT)
- Active support for researchers regarding ELSI matters
- Legal framework for multi-site research projects
- Federated Query System (>90 Mio data elements)
- Meta-data Catalogues (planning phase)
- Swiss Federated Genomics Network (planning phase)



# Beyond 2024: The future Data Coordination Center

- SERI mandated SPHN to set out the options for the continuation of the data infrastructures after 2024 and to carry out the necessary clarifications
- A first report has been presented in April 2022 and was well received
- Refinement of the report ongoing, to be handed in end of 2022

## Vision

The future DCC is the central coordination and competence center in Switzerland for health data in the research domain and enables production, exchange and use of FAIR health data across the country in collaboration with data providers, data producers, data holders, data users, and the network partners.

# Acknowledgements

## The PHI Group:

Katrin Crameri, Jan Armida, Julia Maurer, Kristin Gnodtke, Martin Fox, Michael Müller-Breckenridge, Owen Appleton, Patricia Fernandez Pinilla, Petar Horki, Shubham Kapoor, Simone Guzzi, Vasundra Touré, Christian Ribeaud

The **SPHN NSB** and **NAB**, **Task Forces** & **WGs**

The **BioMedIT Board** and **workforces** @ ETHZ, Unibas, Uni/SIB

The **Hospital workforces** @ USZ, USB, CHUV, Insel, HUG

## The SPHN Management Office:

Thos Geiger, Liselotte Selter, Sarah Vermij, Cédric Petter



@SPHN\_ch



[Sabine.Oesterle@sib.swiss](mailto:Sabine.Oesterle@sib.swiss)



[www.sphn.ch](http://www.sphn.ch) | [www.sib.swiss/phi](http://www.sib.swiss/phi)  
[www.BioMedIT.ch](http://www.BioMedIT.ch)

# SAVE THE DATES!

- ▶ **LIFE SCIENCES CLUSTER FORUM**, Montag, 14. November 2022, 17:00 bis 19:00 Uhr  
U.a. mit Severin Schwan, Lukas Engelberger und special Guest.  
Im auditorium der roche
- ▶ **4. LSC Lunch**, 6. September 2022, 11.30-13.00 Uhr



[www.lifesciencesbasel.ch](http://www.lifesciencesbasel.ch)

