

# Patientenzentriertes Dashboard, Transformation longitudinaler Daten und Pseudonymisierung: Lösungen aus dem DIZ der Charité

MIRACUM-DIFUTURE School 2024

Fabian Prasser und Team

24.09.2024

# Agenda

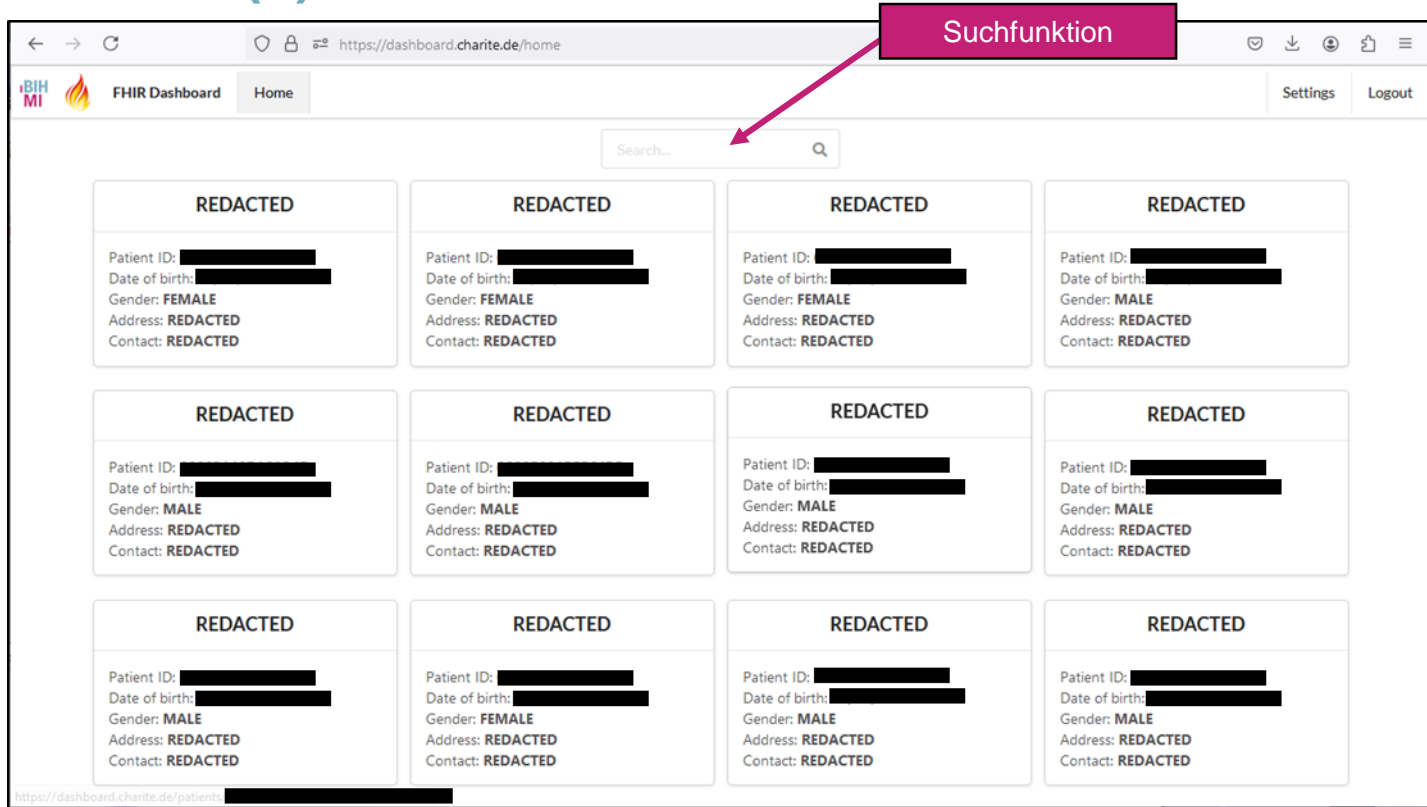
1. Patientenzentriertes FHIR-Dashboard
2. Transformation longitudinaler Daten mit HERALD
3. Pseudonymisierung mit ACE

# 1. Patientenzentriertes FHIR-Dashboard

# Hintergrund

- Verfügbarkeit von FHIR in den DIZen
- Darstellung als patientenzentrierte „Read-Only-Akte“
- Macht das DIZ und den KDS „anfassbar“
- Hilft beim debuggen der FHIR-Daten bzw. Transformationsprozesse
- Perspektivisch Potenzial für klinischen Nutzen, bspw. Befundung
- Integration weitergehender Funktionen, wie Terminologien, und Integration unterschiedlicher Quellsysteme, bspw. für PROMs

# Screenshots (1)



## Screenshots (2)

The screenshot displays the FHIR Dashboard interface for a patient. The top navigation bar includes the BIH logo, 'FHIR Dashboard', 'Home', and tabs for 'Basic clinical parameters', 'Detailed labs and consent', and 'Care intensity'. The patient information section shows 'Patient ID: [REDACTED]', 'Name: REDACTED', 'Date of birth: [REDACTED]', and 'Gender: FEMALE'. Below this, there are four main panels: 'Encounters', 'Diagnoses and conditions', 'Procedures', and 'Medication'. Each panel has a 'Filter...' button. The 'Encounters' panel shows a table with columns for Admission, Discharge, Type, Reason, and Status. The 'Diagnoses and conditions' panel shows a table with columns for Date, Code, and Label. The 'Procedures' panel shows a table with columns for Date, Code, and Label. The 'Medication' panel shows a table with columns for Date, Label, and Status. A tooltip is visible over the 'Diagnoses and conditions' panel, listing various conditions like 'Krankheiten des Blutes und der blutbildenden Organe' and 'Anämie bei chronischen, anderenorts klassifizierten Krankheiten'. A search function is also visible within the 'Medication' panel.

**Encounters**

Admission	Discharge	Type	Reason	Status
[REDACTED] 2024	N/A	INPATIENT ENCOUNTER	N/A	INPROGRESS
[REDACTED] 2024	N/A	AMBULATORY	N/A	INPROGRESS
[REDACTED] 2024	[REDACTED] 2024	INPATIENT ENCOUNTER	N/A	FINISHED
[REDACTED] 2024	[REDACTED] 2024	INPATIENT ENCOUNTER	N/A	FINISHED
[REDACTED] 2024	N/A	AMBULATORY	N/A	INPROGRESS
[REDACTED] 2024	[REDACTED] 2024	INPATIENT ENCOUNTER	N/A	FINISHED
[REDACTED] 2024	N/A	AMBULATORY	N/A	INPROGRESS
[REDACTED] 2024	N/A	AMBULATORY	N/A	INPROGRESS

**Diagnoses and conditions**

Date	Code	Label
[REDACTED] 2024	R52.1	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems
[REDACTED] 2024	G93.6	Aplastische und sonstige Anämien
[REDACTED] 2024	K59.02	Anämie bei chronischen, anderenorts klassifizierten Krankheiten
[REDACTED] 2024	C79.3	Anämie bei Neubildungen C00-D48
[REDACTED] 2024	E86	Anämie bei Neubildungen C00-D48
[REDACTED] 2024	R11	Bronchus und Lunge, mehrere Teilbereiche überlappend
[REDACTED] 2024	D63.0	Anämie bei Neubildungen C00-D48
[REDACTED] 2024	C34.8	Bronchus und Lunge, mehrere Teilbereiche überlappend

**Procedures**

Date	Code	Label
[REDACTED] 2024	3-990	Computergestützte Bilddatenanalyse mit 3D-Auswertung
[REDACTED] 2024	3-990	Computergestützte Bilddatenanalyse mit 3D-Auswertung
[REDACTED] 2024	3-705.0	Ein-Phasen-Szintigraphie
[REDACTED] 2024	3-993	Quantitative Bestimmung von Parametern
[REDACTED] 2024	3-733.x	Sonstige
[REDACTED] 2024	3-733.x	Sonstige

**Medication**

Date	Label	Status
[REDACTED] 2023	Atezolizumab	COMPLETED
[REDACTED] 2023	Bevacizumab	COMPLETED
[REDACTED] 2023	Atezolizumab	COMPLETED
[REDACTED] 2023	Bevacizumab	COMPLETED
[REDACTED] 2023	Atezolizumab	COMPLETED
[REDACTED] 2023	Bevacizumab	COMPLETED

Möglichkeit der Einschränkung auf einen Fall

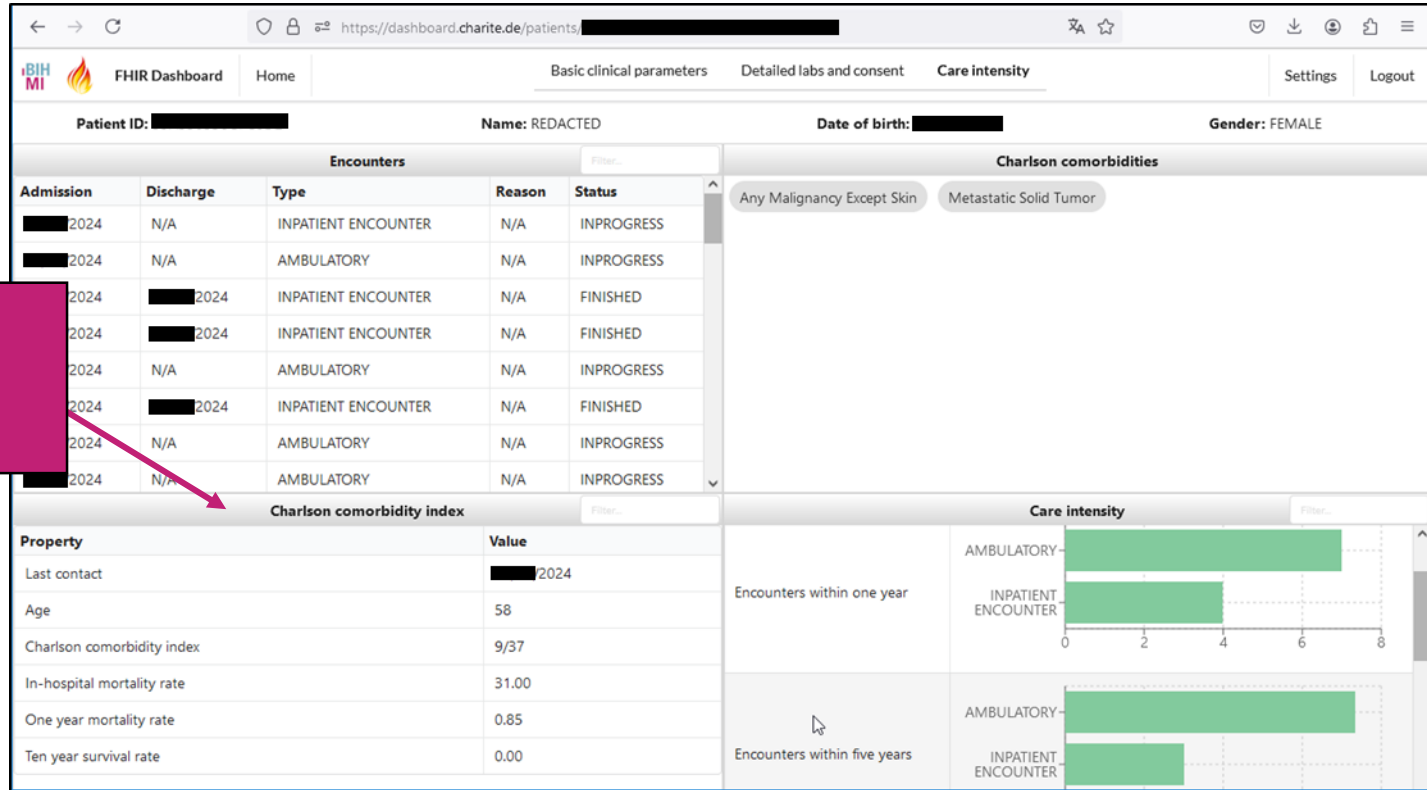
Panels können frei angeordnet werden mit Persistenz im User-Profil

Suchfunktion innerhalb der Panels

## 7



# Screenshots (4)



Nur eine Spielerei

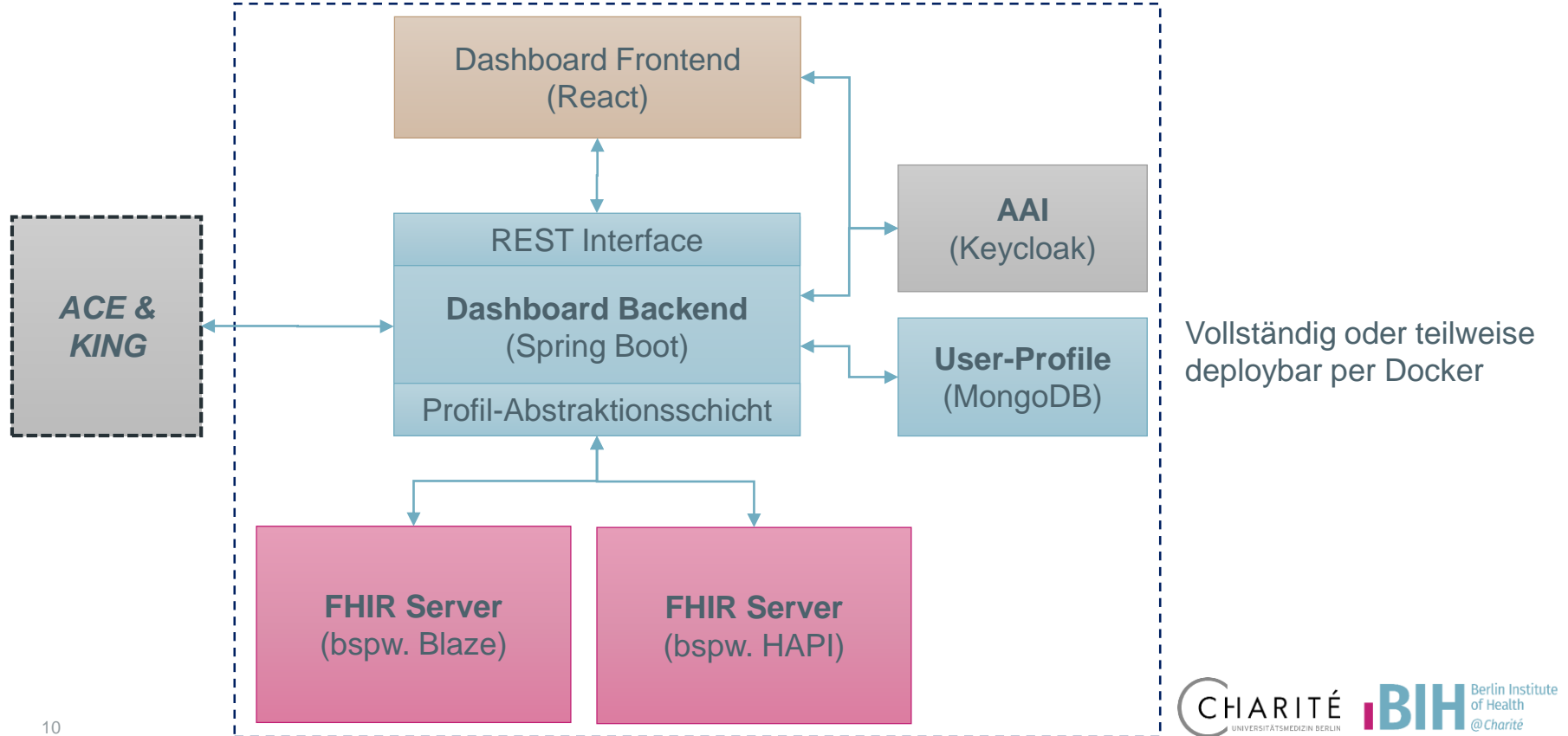


# Screenshots (5)

The screenshot shows the 'Settings' page of the FHIR Dashboard. The browser address bar displays 'https://dashboard.charite.de/settings'. The page has a top navigation bar with 'iBIH MI' logo, 'FHIR Dashboard', and 'Home' tabs. On the right, there are 'Settings' and 'Logout' buttons. The main content area is divided into three columns: 'Workspace 1', 'Workspace 2', and 'Workspace 3'. The 'Workspace 1' column contains a section titled 'Information to display' with a list of checkboxes: 'Diagnoses and conditions' (checked), 'Encounters' (checked), 'Procedures' (checked), 'Lab results and further observations' (unchecked), 'Medication history' (checked), 'Timeline of lab results and observations' (unchecked), 'Clinical notes' (unchecked), 'Diagnostic reports' (unchecked), 'Trend of lab results and observations' (unchecked), 'Charlson comorbidity index' (unchecked), 'Charlson comorbidities' (unchecked), 'Care intensity' (unchecked), and 'Consents' (unchecked). Below this is a 'Layout options (rows x columns)' section with a dropdown menu set to '2 x 2'. The 'Workspace 2' and 'Workspace 3' columns are currently empty. To the right of the workspace columns, there is a 'Language' section with a dropdown menu set to 'en-GB', and a 'Workspace labels' section with three text input fields: 'Workspace 1' (containing 'Basic clinical parameters'), 'Workspace 2' (containing 'Detailed labs and consent'), and 'Workspace 3' (containing 'Care intensity').

Konfiguration der  
individuellen  
Ansicht

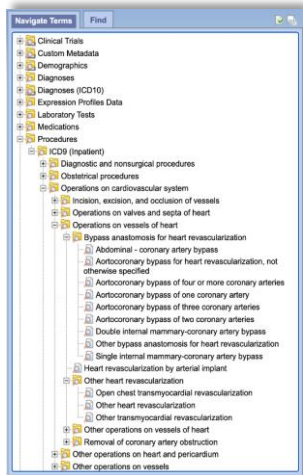
# Architektur und Technologie



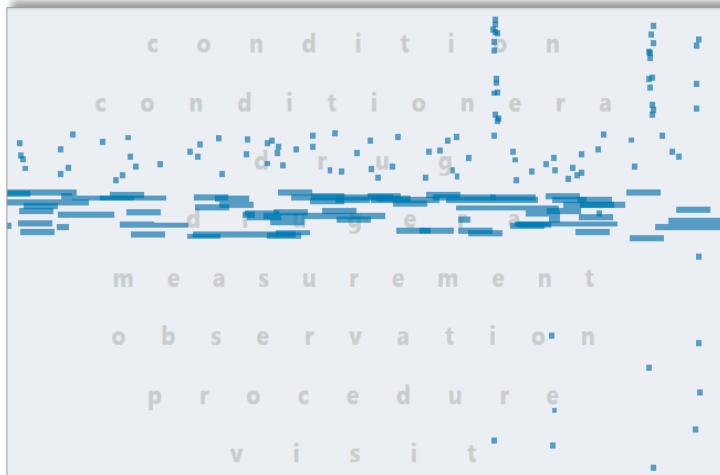
## 2. Transformation longitudinaler Daten mit HERALD

# Herausforderungen

Klinische Routinedaten sind ...



...hierarchisch



...longitudinal

Viele Analysen benötigen...

Patient ID	Age (#Values: 338)	Sex (#Values: 338)	FirstCancerDiagnosis (#Values: 338)	FirstChemoRadio (#Values: 338)	ProcedureCount (#Values: 338)	Medication (#Values: 328)	Hemoglobin (#Values: 151)	Glucose (#Values: 144)
120820	51	M	Non-small cell carcinoma of lung TNM stage 4 (disorder)	Combined chemotherapy and radiation therapy (procedure)	9	PACLitaxel 100 MG Injection	5.90000 %	69.00000 mg/dL
121087	60	M	Non-small cell carcinoma of lung TNM stage 4 (disorder)	Combined chemotherapy and radiation therapy (procedure)	8	Cisplatin 50 MG Injection		
121125	46	F	Non-small cell carcinoma of lung TNM stage 4 (disorder)	Combined chemotherapy and radiation therapy (procedure)	4	PACLitaxel 100 MG Injection		
121287	46	F	Non-small cell carcinoma of lung TNM stage 4 (disorder)	Combined chemotherapy and radiation therapy (procedure)	7	PACLitaxel 100 MG Injection	6.00000 %	82.00000 mg/dL
121505	44	F	Non-small cell carcinoma of lung TNM stage 4 (disorder)	Combined chemotherapy and radiation therapy (procedure)	8	Cisplatin 50 MG Injection	6.90000 %	117.00000 mg/dL
121934	52	F	Non-small cell carcinoma of lung TNM stage 4 (disorder)	Combined chemotherapy and radiation therapy (procedure)	2	Cisplatin 50 MG Injection	6.30000 %	83.00000 mg/dL

...tabulare Querschnittsdaten

# Ansatz: Domänenspezifische Sprache HERALD

A.

```
FirstDiagnosis: FIRST (LABEL = "*Diabetes*" OR LABEL = "*Coronary Heart Disease*") AFTER (1980-01-01)
Comorbidities: EXISTS (LABEL = "*Diabetes*") AND EXISTS (LABEL = "*Coronary Heart Disease*")
BodyMassIndex: AVERAGE (LABEL = "*Body Mass Index*" AND UNIT = "*kg/m2*") CLOSE TO (LABEL = "FirstDiagnosis") BY 6 MONTHS
LDLCholesterol: FIRST (LABEL = "*LDL Cholesterol*" AND UNIT = "*mg/dL*") AFTER (LABEL = "FirstDiagnosis") BY 4 WEEKS
HDLCholesterol: ANY (LABEL = "*HDL Cholesterol*" AND UNIT = "*mg/dL*") CLOSE TO (LABEL = "LDLCholesterol") BY 1 DAY
CholesterolRatio: RATIO BETWEEN (LABEL = "LDLCholesterol") AND (LABEL = "HDLCholesterol")
```

B.

Label for generated  
observations

Selector, aggregator,  
comparator, or query

Temporal  
relationship

FirstDiagnosis:

FIRST (LABEL = "\*Diabetes\*" OR LABEL = "\*Coronary Heart Disease\*") AFTER (1980-01-01)

Selector

Filter expression

Temporal  
condition

Constant  
timestamp

BodyMassIndex:

AVERAGE (LABEL = "\*Body Mass Index\*" AND UNIT = "\*kg/m2\*") CLOSE TO (LABEL = "FirstDiagnosis") BY 6 MONTHS

Aggregator

Filter expression

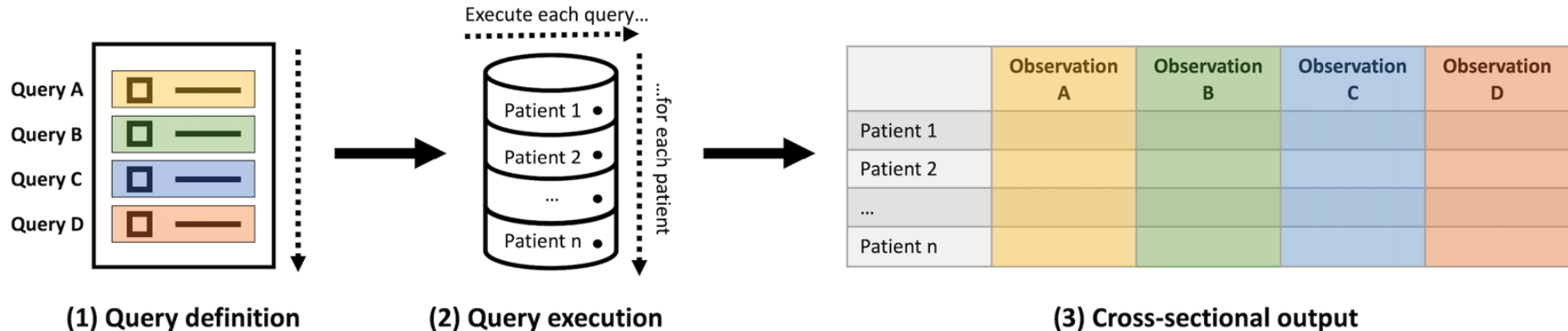
Temporal  
condition

Filter expression

Offset

# Ausführung von HERALD Queries

- Eine Datendefinition ist eine Liste von Queries, die jeweils eine Beobachtung produzieren
- Jede Query verarbeitet alle Beobachtungen als Eingabe und erzeugt eine Beobachtung
- Queries werden nacheinander für jeden Patienten in der aufgeführten Reihenfolge ausgeführt
- Queries können sich auf Beobachtungen beziehen, die von früheren Queries erzeugt wurden



# Beispielimplementierung: Datendefinition

Select cohort(s)

- ☒ Coronary Heart Disease
- ☒ Diabetes

Specify observation(s) to analyze

FirstDiag

FIRST (LABEL = "\*\*Diabetes\*" OR LABEL = "\*\*Coronary Heart Disease\*") AFTER (1980-01-01)

Edit ▾

Comorbidities

EXISTS (LABEL = "\*\*Diabetes\*") AND EXISTS (LABEL = "\*\*Coronary Heart Disease\*")

Edit ▾

BodyMassIndex

AVERAGE (LABEL = "\*\*Body Mass Index\*" AND UNIT = "\*\*kg/m2\*") CLOSE TO ("FirstDiag") BY 6 MONTHS

Edit ▾

LDLCholesterol

FIRST (LABEL = "\*\*Low Density Lipoprotein Cholesterol\*" AND UNIT = "\*\*mg/dL\*") AFTER ("FirstDiag") BY 4 WEEKS

Edit ▾

HDLCholesterol

ANY (LABEL = "\*\*High Density Lipoprotein Cholesterol\*" AND UNIT = "\*\*mg/dL\*") CLOSE TO ("LDLCholesterol") BY 1 DAY

Edit ▾

CholesterolRatio

RATIO BETWEEN ("LDLCholesterol") AND ("HDLCholesterol")

Edit ▾

☒ Patient age

☒ Patient sex

Quick add ▾

Check specification

# Query-Editor

Specify an observation to be used in the analysis

Label: FirstDiag

Selection

FIRST

AFTER

Date

(1980-01-01)

Without offset

Add rule

Add group

AND

OR

Delete

LABEL

equal

\*Diabetes\*

Delete

LABEL

equal

\*Coronary Heart Disease\*

Cancel

OK

# Ontologie-Ansicht

Select field or value

Search...

☒ Include homonyms

☒ Include specializations

- Conditions

- Diagnosis
  - Coronary Heart Disease
    - Value: Coronary Heart Disease
  - Diabetes
    - Value: Diabetes

+ Demographic Information

- Observations

- Body Mass Index
  - Unit: kg/m2
- Low Density Lipoprotein Cholesterol
  - Unit: mg/dL
- High Density Lipoprotein Cholesterol
  - Unit: mg/dL



# Beispielimplementierung: Analysen und Visualisierungen

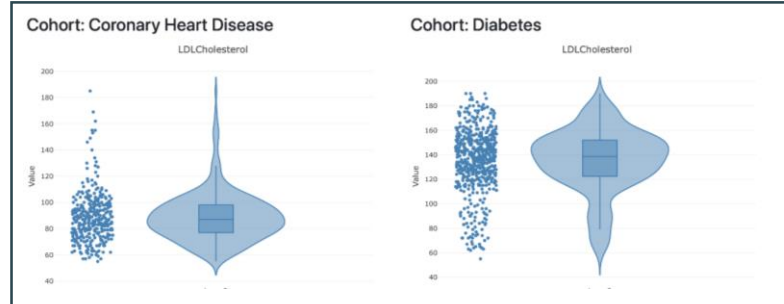
## A. Data quality report

Data quality report for Cohort Coronary Heart Disease					
Label	Completeness (%)	Missing rate (%)	Uniqueness (%)	Data Type	Unit
FirstDiag	100.00	0.00	99.56	Categorical	Unknown
Comorbidities	100.00	0.00	100.00	Categorical	Boolean
BodyMassIndex	66.18	33.82	100.00	Decimal	kg/m2
LDLCholesterol	48.25	51.75	99.70	Integer	mg/dL
HDLCholesterol	45.92	54.08	99.68	Integer	mg/dL
CholesterolRatio	45.92	54.08	100.00	Decimal	(mg/dL) / (mg/dL)

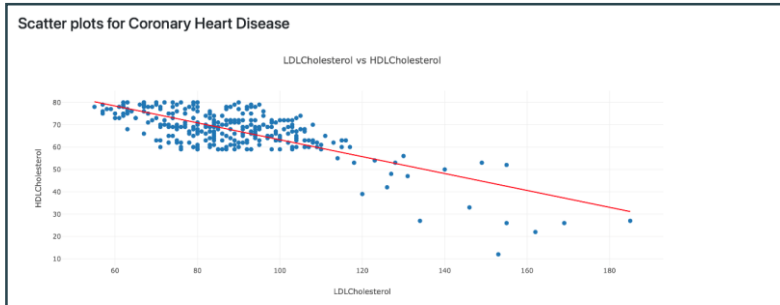
  

Data quality report for Cohort Diabetes					
Label	Completeness (%)	Missing rate (%)	Uniqueness (%)	Data Type	Unit
FirstDiag	100.00	0.00	99.85	Categorical	Unknown
Comorbidities	100.00	0.00	100.00	Categorical	Boolean
BodyMassIndex	76.30	23.70	99.80	Decimal	kg/m2
LDLCholesterol	80.48	19.52	99.81	Integer	mg/dL
HDLCholesterol	78.69	21.31	99.43	Integer	mg/dL
CholesterolRatio	78.69	21.31	100.00	Decimal	(mg/dL) / (mg/dL)

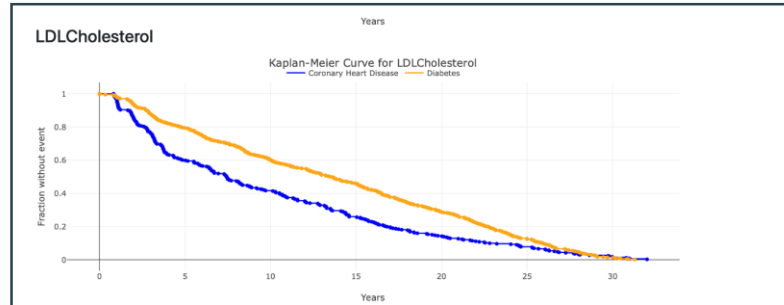
## B. Cohort comparison



## C. Scatter plot



## D. Time-to-event analysis



# Beispielimplementierung: i2b2-Integration

BIHMI CHARITÉ Project: Synthea SyntheticMass Demo User: Lena Baum Find Patients | Analysis Tools | Help | Change Password | Logout

**Navigate Terms** Find

- Allegies
- Care Plans
- Conditions
  - Description
    - [Acute allergic reaction - Primary small cell malignant neoplasm of lung T1M stage 4 (disorder)]
    - Acute allergic reaction
    - Acute bacterial sinusitis (disorder)
    - Acute bronchitis (disorder)
    - Acute viral pharyngitis (disorder)
    - Alzheimer's disease (disorder)
    - Antepartum eclampsia
    - Appendicitis
    - Asthma
    - Atopic dermatitis
    - Attempted suicide - suffocation
    - Bleeding from anus
    - Blighted ovum
    - Blindness due to type 2 diabetes mellitus (disorder)
    - Brain damage - traumatic
    - Bullet wound
    - Cardiac Arrest
    - Child attention deficit disorder
    - Childhood asthma
    - Chronic intractable migraine without aura
    - Chronic obstructive bronchitis (disorder)
    - Chronic pain
    - Chronic paralysis due to lesion of spinal cord
    - Chronic sinusitis (disorder)
    - Closed fracture of hip
    - Complication occurring during pregnancy
    - Concussion injury of brain
    - Concussion with loss of consciousness
    - Concussion with no loss of consciousness
    - Congenital uterine anomaly
    - Contact dermatitis
    - Coronary Heart Disease
    - Cystitis
    - Diabetes
    - Diabetic renal disease (disorder)
    - Diabetic retinopathy associated with type II diabetes mellitus (disorder)
    - Drug overdose
    - End stage renal disease (disorder)
    - Escherichia coli urinary tract infection
    - Facial laceration
    - Familial Alzheimer's disease of early onset (disorder)
    - Fetus with chromosomal abnormality
    - First degree burn
    - Fracture of ankle
    - Fracture of clavicle
    - Fracture of forearm
    - Fracture of rib
    - Fracture of the vertebral column with spinal cord

**Cross-sectional analysis**

Specify data Perform action

Select cohort(s)

☒ Alzheimer's disease  
☒ Early onset Alzheimer's disease

Export to PDF

Specify observation(s) to analyze

Alzheimer'sDiagnosis FIRST (LABEL = "Alzheimer's disease (disorder)" OR LABEL = "Familial Alzheimer's disease of early onset (disorder)") Edit +

Pneumonia FIRST (LABEL = "Pneumonia") Edit +

☒ Patient age ☒ Patient sex Quick add + Check specification

Select analysis to perform

☐ Quality report ☐ Cohort description ☐ Statistical comparison ☒ Time to event analysis ☐ Scatter plots

Visualize

**Pneumonia**


Kaplan-Meier Curve for Pneumonia

Alzheimer's disease Early onset Alzheimer's disease

Years	Alzheimer's disease (Fraction without event)	Early onset Alzheimer's disease (Fraction without event)
0	1.0	1.0
4	1.0	1.0
6	0.85	0.85
8	0.65	0.65
10	0.45	0.45
12	0.25	0.25
14	0.10	0.10
16	0.05	0.05

# Weiterführende Informationen und Online-Demo

## HERALD Project - Web Demo



**HERALD (Human-centric Extraction for Research and Analysis of Longitudinal Data)** is a query language designed to facilitate the extraction and aggregation of longitudinal health data into cross-sectional tables suitable for statistical analysis and machine learning. Its syntax closely mirrors natural language. The implementation consists of a parser and an execution engine, combined with a graphical query editor, and analytics functionalities. Below you'll find three different examples provided as a web demo, that are based on subsets of the synthetically generated patient records from the [Synthea SyntheticMass dataset](#).

### Example 1: Coronary Heart Disease vs. Diabetes

A comparative analysis between two cohorts of patients with a primary diagnosis of either coronary heart disease or type 2 diabetes. The example supports a comparison of demographic variables (age, sex) and common risk factors (BMI, LDL and HDL cholesterol levels) as well as comorbidities.

Open

### Example 2: Alzheimer's Disease

Comparison of patients with early-onset Alzheimer's with those diagnosed at a later stage. The example supports a comparison of basic demographic information (age, sex), Mini-Mental State Examination (MMSE) scores to assess cognitive function, pneumonia incidence, care plans and medications.

Open

### Example 3: Lung Cancer

Comparison of patients diagnosed with non-small cell carcinoma of the lung of TNM stages 1 to 4, who have received combined chemotherapy and radiation therapy. Additional information on basic demographics (age, sex), BMI, related medication and selected lab values (Hemoglobin, Glucose and eGFR) is provided.

Open

### Explore our Code on GitHub

HERALD Lang

Repository with the core components of HERALD, including a parser, an execution engine, and a graphical editor.

HERALD Demo

Repository containing the code of this online demo.

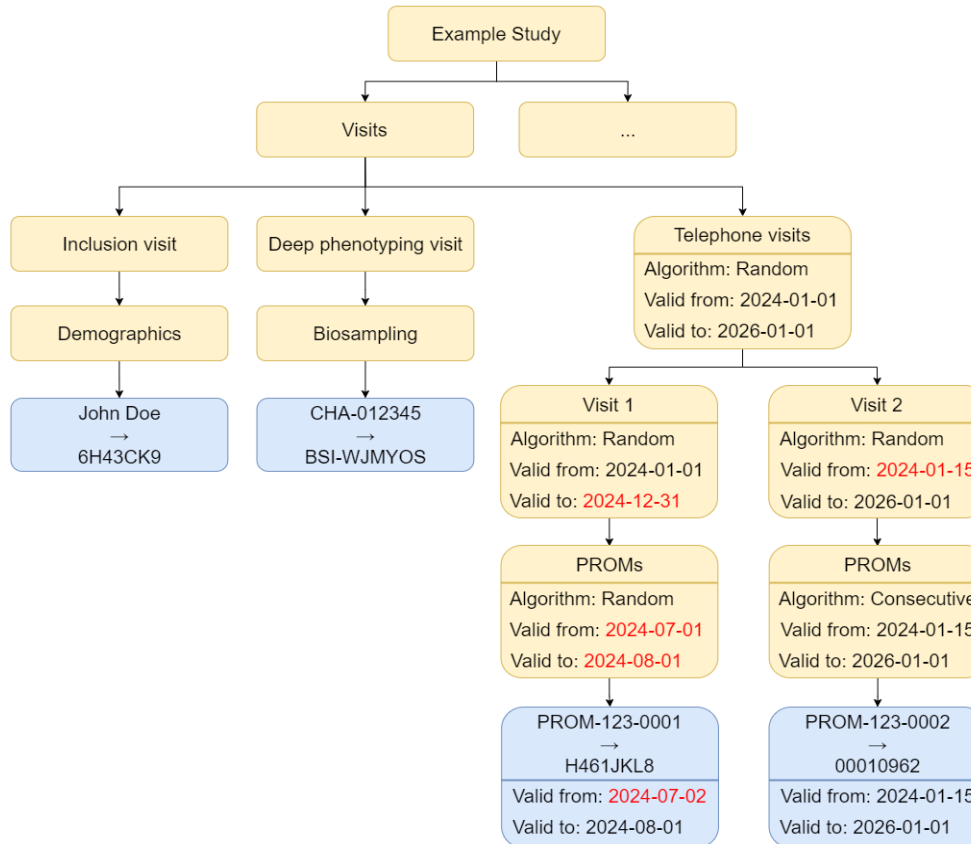
© 2024 Berlin Institute of Health

[Home](#) | [Imprint](#) | [Privacy](#)

<https://www.herald-lang.org>

### 3. Pseudonymisierung mit ACE

# Herausforderung und Lösung



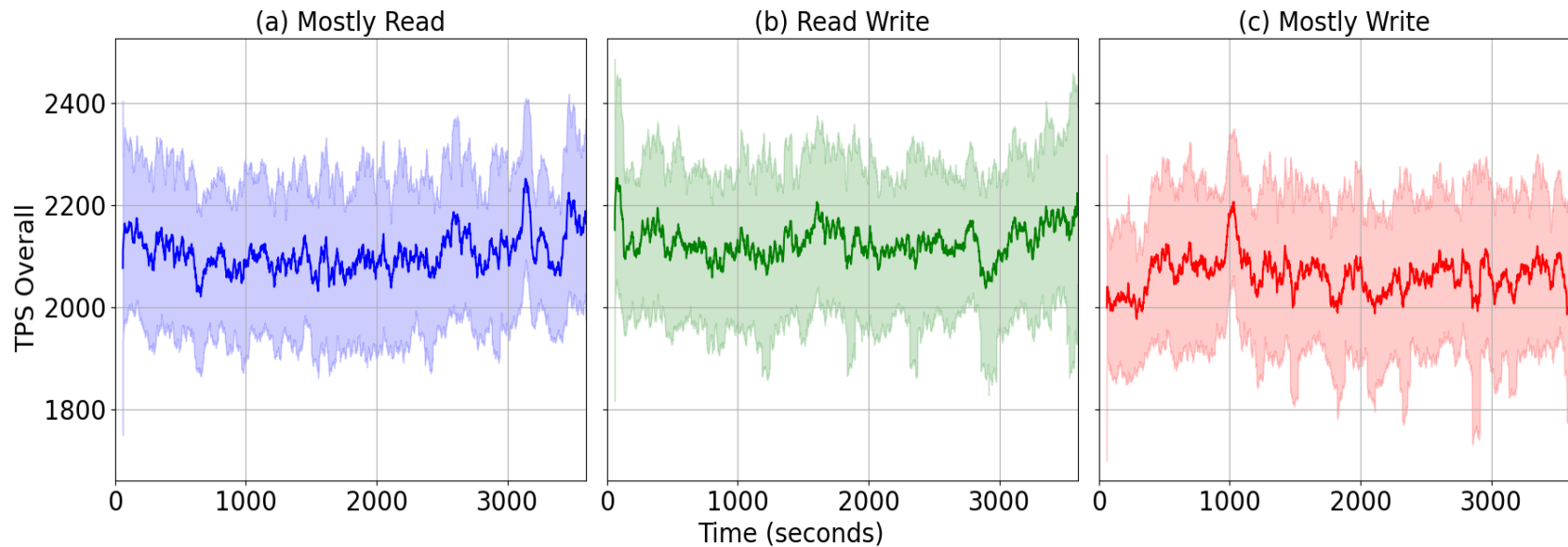
- Komplexe Anforderungen an die Pseudonymisierung (u.a. primär, sekundär, Datenarten, Audit Trail, Gültigkeiten)
- Hohe Skalierbarkeit benötigt (bspw. im DIZ der Charité 4 Mio. Patienten, 16 Mio. Fälle)

# Design und Features





- Hierarchisches Domänenmodell mit Gültigkeitszeiträumen und Vererbung
- Docker, PostgreSQL, Spring Boot, Keycloak
- Berechtigungsmodell: Nutzer x Domäne x Operation
- Konfigurierbar: Pseudonymisierungsalgorithmus, Alphabet, Checksums
- Vollständiger Audit-Trail

# Performance




# Weiterführende Informationen

**THS-MED**  **TrustDeck**




## Open-Source-Werkzeuge für Treuhandstellen in der Translationalen Medizin




**Advanced Confidentiality Engine**

ACE bietet eine robuste Lösung für Pseudonymisierungsdienste und zeichnet sich durch seine hohe Skalierbarkeit und einen integrierten Audit-Trail aus. Seine Architektur ermöglicht die Bildung von Domänenhierarchien mit vererbaren Eigenschaften.



**Key Index of Names and General Identification Numbers**

KING bietet fortgeschrittene Funktionen im Bereich Identitätsmanagement. Es ist hochskalierbar und in hohem Maße konfigurierbar. Verschiedene Record-Linkage-Verfahren sind integriert, um die Bedürfnisse unterschiedlicher Anwendungsfälle



**Holistic Electronic Agreement Recording Tool**

HEART ist ein spezialisiertes Tool zur elektronischen Erfassung und Verwaltung von Einwilligungsdokumenten. Es ist flexibel konfigurierbar, unterstützt die semantische Annotation von Einwilligungsdokumenten und ist für den Einsatz auf Tablets und

<https://www.ths-med.de>



# Danke für Ihre Aufmerksamkeit!

mi.bihealth.org  
medic.charite.de