

From Hospitals to Researchers: A Data-Trustee Infrastructure to Search and Use FHIR-Data for Retrospective Medical Research

Carolin Poschen, Joscha Grüger, Britta Berens,
Helene Christ, Lukas Meyer, Konstantin Knorr

Presenting: Carolin Poschen
Trier University of Applied Sciences, c.poschen@hochschule-trier.de

HEALTHINFO 2025
26. – 30. October 2025



Informatik
Hauptcampus

H O C H
S C H U L E
T R I E R

- B.Sc. Medical Informatics, Trier University of Applied Sciences
- M.Sc. Data Science, University of St. Thomas, MN, USA
- 01/2021 – 05/2023 various research projects in ML and AI
- Since 09/2023 Researcher and Project Lead at project DaTreFo at Trier University of Applied Sciences





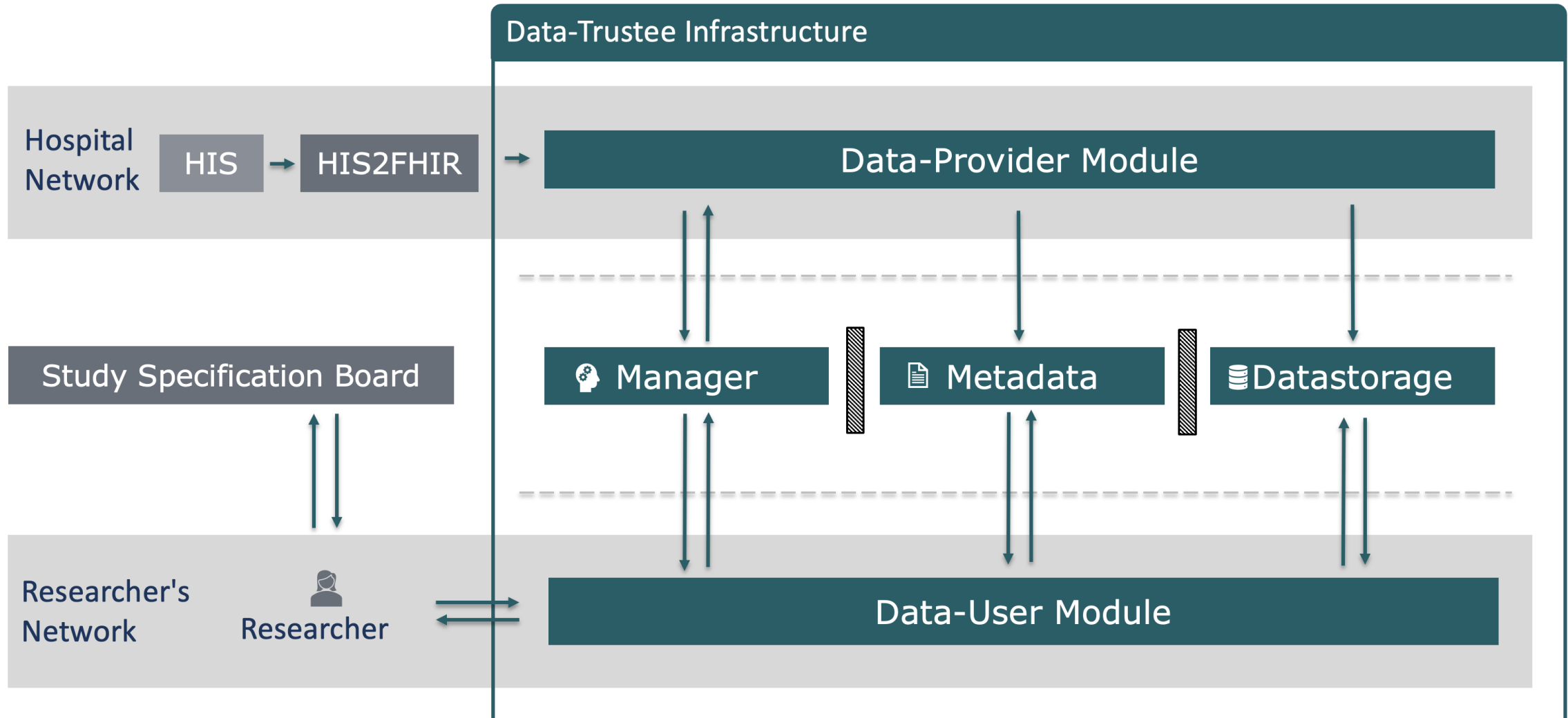
Research project to design and develop a data-trustee infrastructure to manage and provide medical data

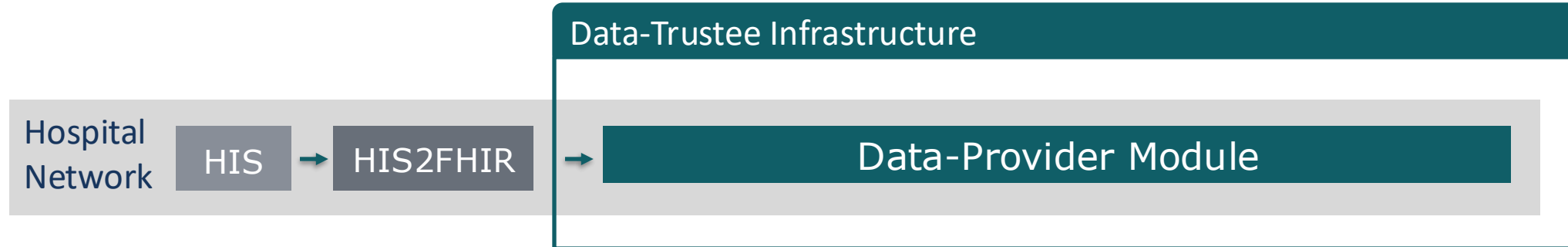


Secure data pipeline for privacy-compliant secondary use of medical data

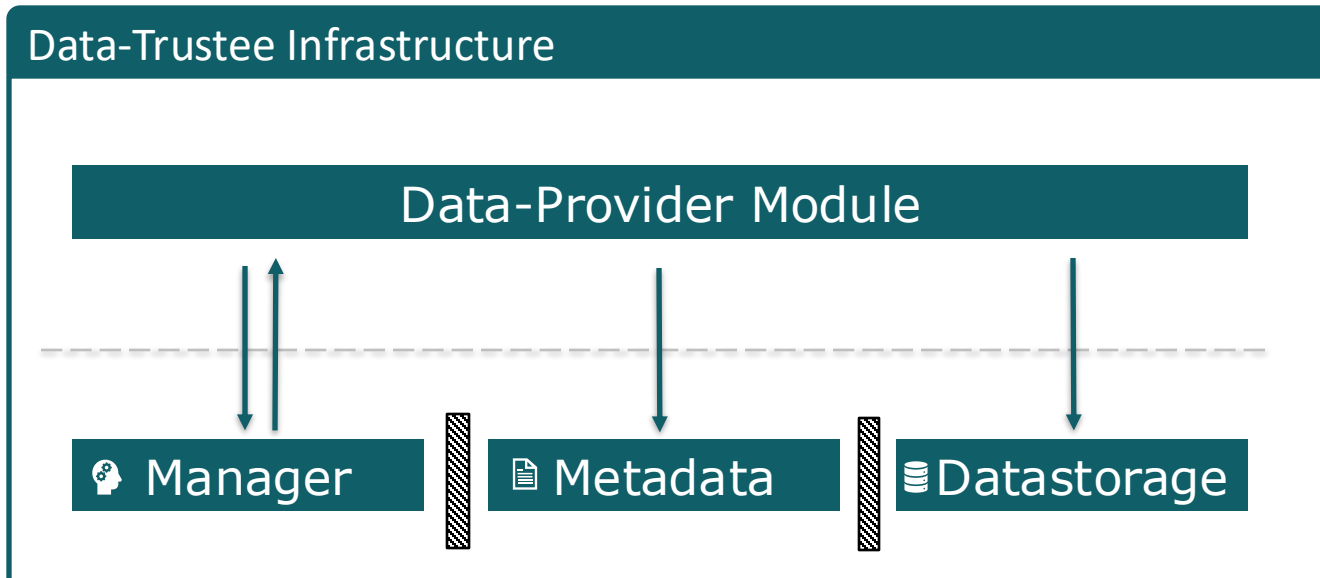


Cooperation of Trier University of Applied Sciences with Dedalus HealthCare GmbH and German Research Center for Artificial Intelligence (DFKI)





- Data (medical, demographic, and consent) from the Hospital Information System (**HIS**) are extracted, transformed into the FHIR format (**HIS2FHIR**), and send to the Data-Provider Module (**DPM**)
- DPM splits data and preprocesses data to be stored in different modules



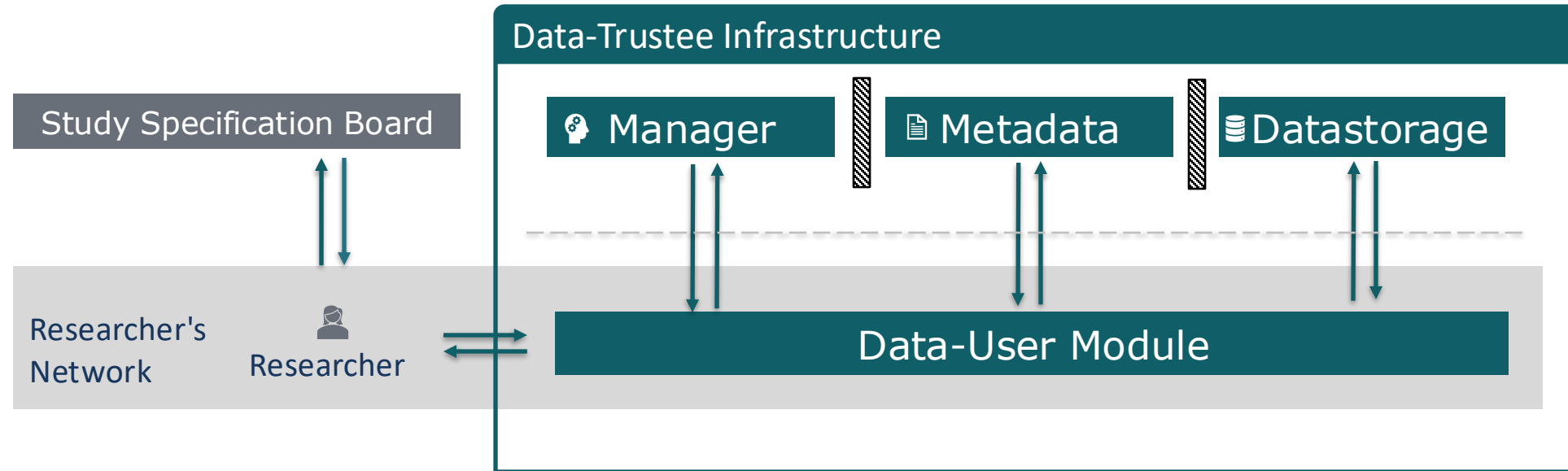
Demographic data, consent, and medical data IDs are stored in the **Manager** module



Metadata for each data entry are stored in the **Metadata** module



Raw medical data are encrypted and stored in the **Datastorage** module



- Researchers receive a customized Data-User Module (**DUM**) with their Study Specification Document
- DUM starts automated search on demographic data and metadata
- After a successful search, data are downloaded from the Datastorage module

- Researchers provide study requirements and cohort descriptions
- Data requirements based on generalized data

The image displays three screenshots of the Study Specification Board interface, illustrating the configuration of a cohort named 'Diab. Group'.

Left Screenshot: The 'Observation' tab is active. The 'Observation Type' is set to 'Laboratory'. A text box prompts the user to 'Please enter your requirements on the cohorts' laboratory values.' Below this, a checkbox for 'Related to Encounter' is unchecked. A search box for 'Glucose' is shown, with a dropdown list containing 'Glucose|MCnc|Pt|ANYBldSerPI' (LG7967-5 (LOINC)). At the bottom, a 'Value*' section has radio buttons for 'decreased', 'normal', and 'elevated', with 'elevated' selected.

Middle Screenshot: The 'Condition' tab is active. A text box prompts the user to 'Please enter your requirements on the cohorts' conditions.' Below this, a checkbox for 'Related to Encounter' is unchecked. A search box for 'Diabetes mell' is shown, with a dropdown list containing 'Diabetes mellitus' (E10-E14 (ICD-10)), 'Diabetes mellitus Type 1' (E10 (ICD-10)), and 'Diabetes mellitus Type 2' (E11 (ICD-10)). The 'Diabetes mellitus' option is selected. A 'Save' button is visible at the bottom right.

Right Screenshot: The 'Cohorts' overview is shown. It lists two cohorts: 'Diab. Group' and 'Non Diab. Group'. The 'Diab. Group' cohort is expanded, showing its requirements: 'Patients of all gender that are born before 2006', 'that have had elevated values of Glucose|MCnc|Pt|ANYBldSerPI (LG7967-5 LOINC)', and 'and have been diagnosed with Diabetes mellitus (E10-E14 ICD-10)'. The 'Non Diab. Group' cohort is also expanded, showing its requirements: 'Patients of all gender that are born before 2006', 'that have had elevated values of Glucose|MCnc|Pt|ANYBldSerPI (LG7967-5 LOINC)', 'that have not been diagnosed with Diabetes mellitus (E10-E14 ICD-10)', and 'and have been diagnosed with Other retinal disorders (H35 ICD-10)'.

Original Data

```
{  
  "resource": {  
    "resourceType": "Patient",  
    "id": "Patient-1234",  
    "birthDate": "2005",  
    "gender": "male"  
  }  
}
```

Study Specification Document

```
{  
  "resource": {  
    "resourceType": "Patient",  
    "id": "Patient-1",  
    "birthDate": {  
      "operator": "$lte",  
      "value": "2006"  
    },  
    "gender": {  
      "operator": "$in",  
      "value": [  
        "male", "female", "divers"  
      ]  
    }  
  }  
}
```

Original Data

```
{
  "resource": {
    "resourceType": "Patient",
    "id": "Patient-1234",
    "birthDate": "2005",
    "gender": "male"
  }
}
```

Study Specification Document

```
{
  "resource": {
    "resourceType": "Patient",
    "id": "Patient-1",
    "birthDate": {
      "operator": "$lte",
      "value": "2006"
    },
    "gender": {
      "operator": "$in",
      "value": [
        "male", "female", "divers"
      ]
    }
  }
}
```

“Male patients born before 1961 that have had an inpatient encounter after 2017 due to atrial fibrosis and take or have taken beta blockers.”



Medical data

- Encounter and Conditions
- Observations (vital-signs, laboratory)
- ...

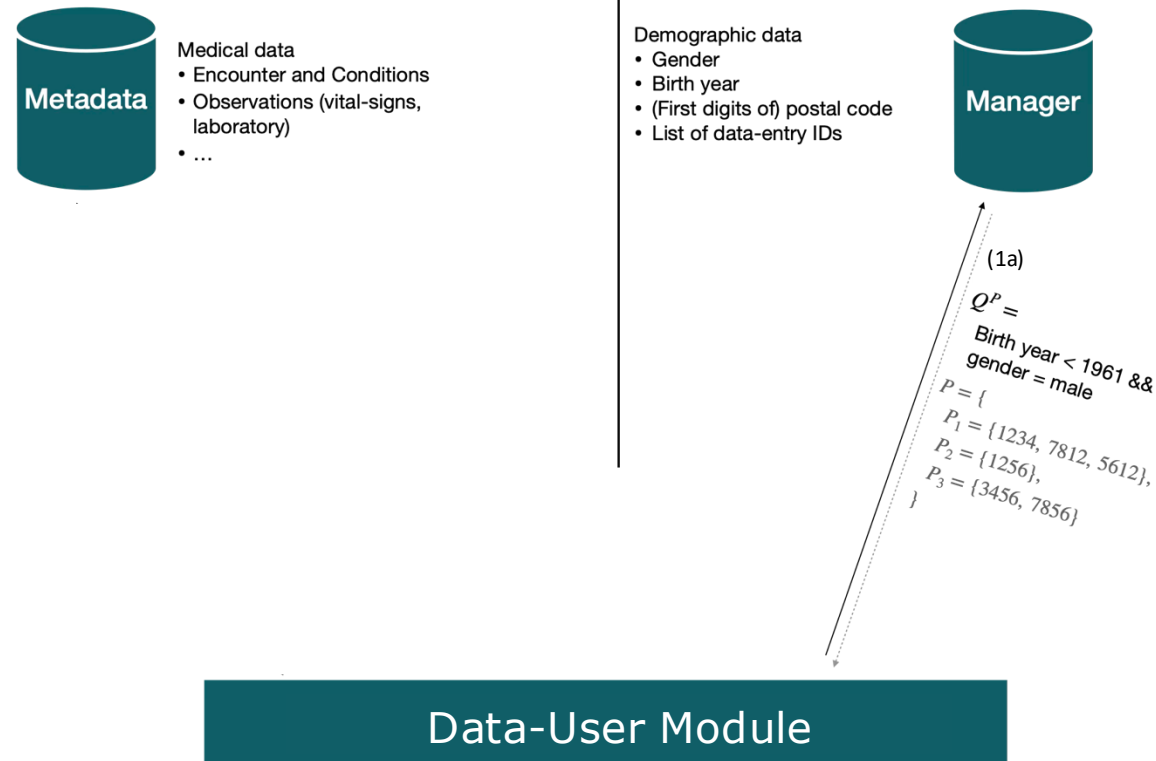
Demographic data

- Gender
- Birth year
- (First digits of) postal code
- List of data-entry IDs

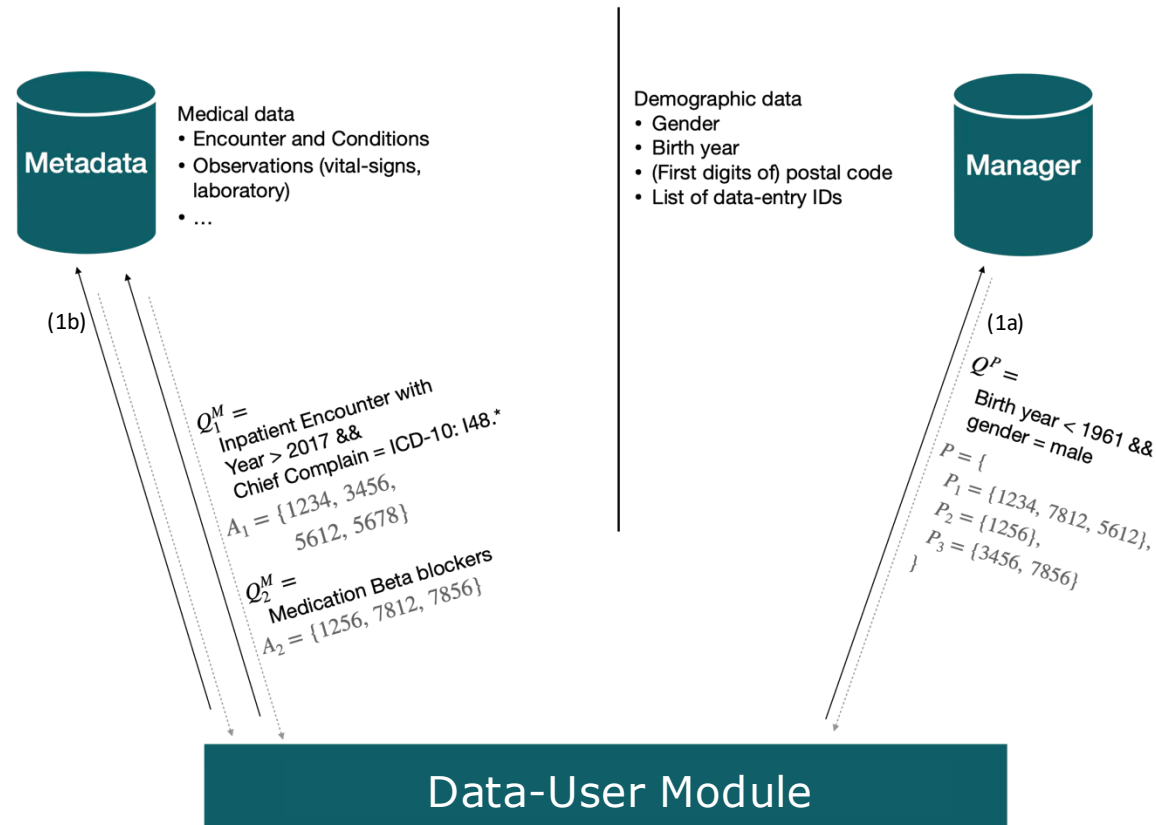


Data-User Module

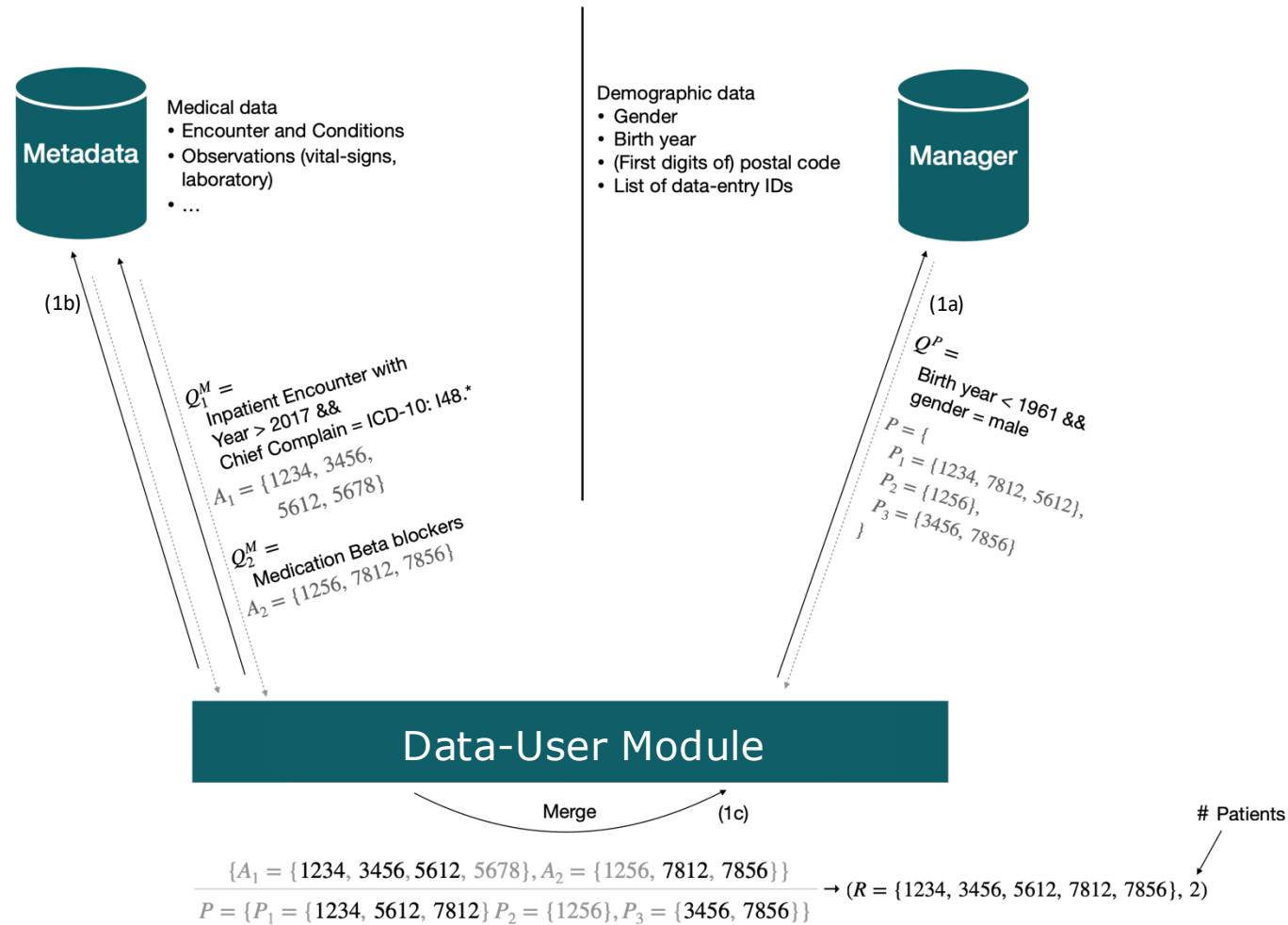
"Male patients born before 1961 that have had an inpatient encounter after 2017 due to atrial fibrosis and take or have taken beta blockers."



“Male patients born before 1961 that **have had an inpatient encounter after 2017 due to atrial fibrosis** and **take or have taken beta blockers.**”

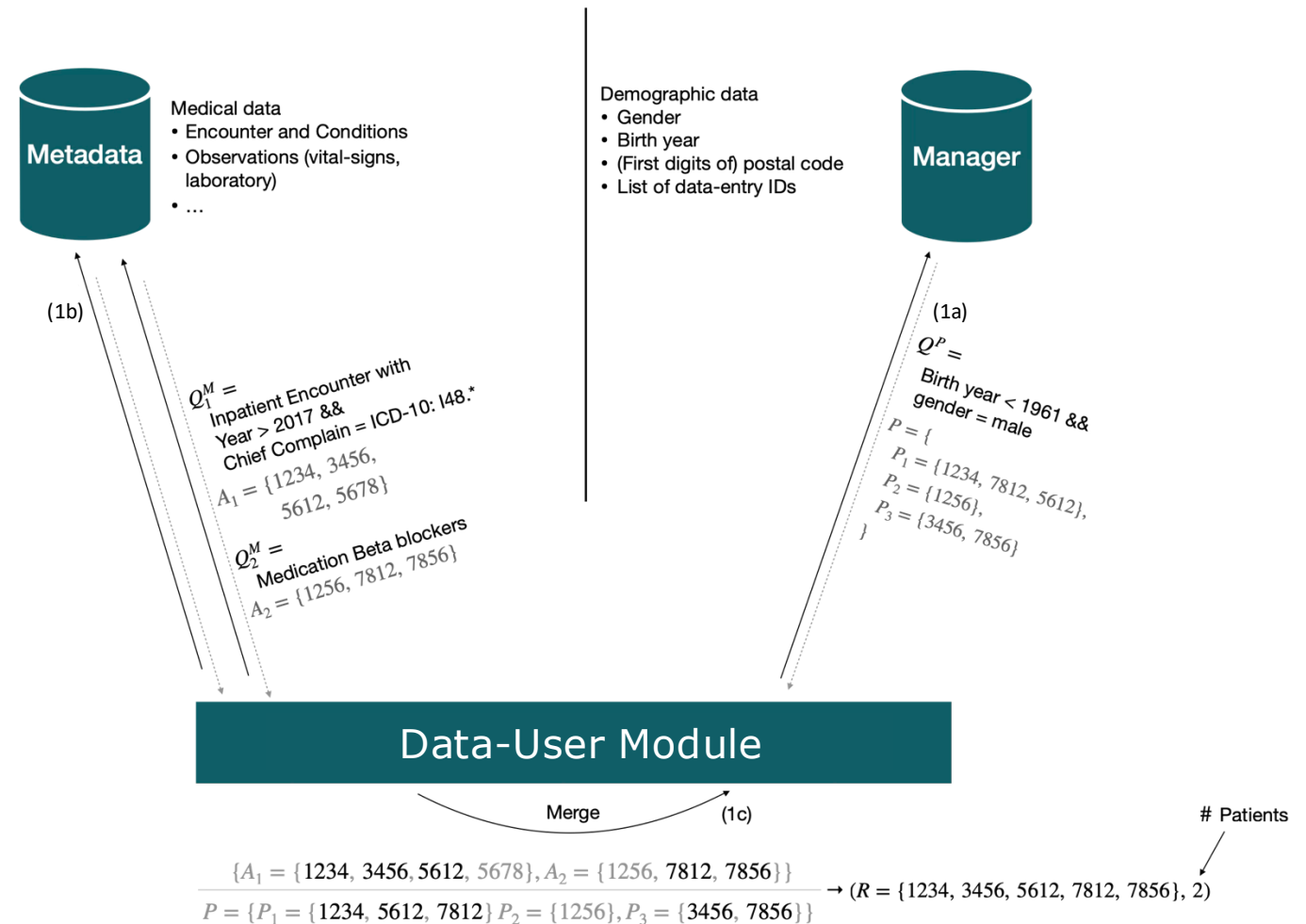


“Male patients born before 1961 that have had an inpatient encounter after 2017 due to atrial fibrrosis and take or have taken beta blockers.”



- If *enough* IDs match, in the Data-User Module:
 - Request medical data from Datastorage module
 - Request demographic data and keys from Manager module

“Male patients born before 1961 that have had an inpatient encounter after 2017 due to atrial fibrosis and take or have taken beta blockers.”



- ✓ Nearly fully-automated pipeline from hospital to researcher
- ✓ Separation of concern in core-DTI: no module has access to both patient identity and medical data
- ✓ FHIR resources enhance interoperability
- ✓ Privacy-preserving search allows access only to data meeting the study requirements
- ✗ Architecture requires own module for each participating hospital and researcher
- ✗ Generalized metadata simplifies data requirement specification but limits granularity for data queries

Carolin Poschen
Trier University of Applied Sciences
Germany

c.poschen@hochschule-trier.de

LinkedIn Profile



Project Website



DaTreFo
Datentreuhänder mit geschützten
Datenschätzen für die Forschung

With funding from the:

