

Einsatz des i2b2 Frameworks für den Secondary Use in Medizinischen Forschungsverbänden

GMDS 2010, Mannheim, 08.09.2010

CIOffice Forschungsnetze

K. Helbing, L. Reimann, S. Zeiss, U. Sax



Inhalt

- Ziele
- Ausgangssituation
- Workflow zur Übertragung von Studiendaten ins i2b2
- Diskussion
- Ausblick

Ziele

- Daten nach Abschluss einer Studien für weitere Fragestellungen zur Verfügung stellen
- Ein Tool bereitstellen, das intuitiv von den Forschern genutzt werden kann
- Die Berücksichtigung des Datenschutzes im Bezug auf die Langzeitspeicherung für medizinische Forschungsdaten
- Generischer Ansatz
- Lösung aus Open Source Produkten

3

Ausgangssituation

- Studiendaten liegen im CDISC ODM Format vor
- TMF Datenschutzkonzepte:
 - Zweistufige Pseudonymisierung der Patientendaten für Langzeitspeicherung von Forschungsdaten
 - Pseudonymisierungsdienst (PSD)
- i2b2 (Informatics for Integrating Biology and the Bedside):
 - NIH-gefördert
 - Open Source Tool
 - TMF-i2b2-Usergroup

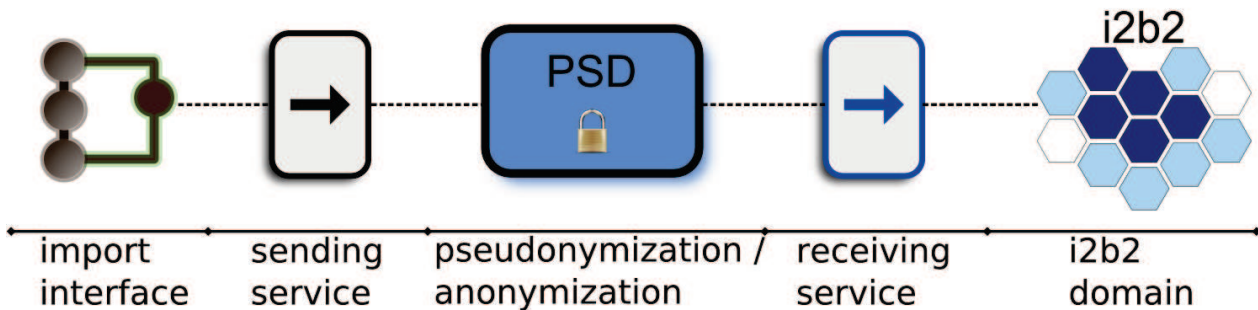
4

i2b2

The screenshot displays the i2b2 Query Tool interface. The top window, 'Query Tool', shows a query named '20-29-männlich@02:06:16' with three groups: Group 1 (20-29), Group 2 (männlich), and Group 3 (empty). The bottom window, 'Timeline View', shows a table with columns for Row #, Name of Terms, Date Constraint, Value Constraint, Shape, and Color. The 'Patient Set' is currently empty (0 patients).

5

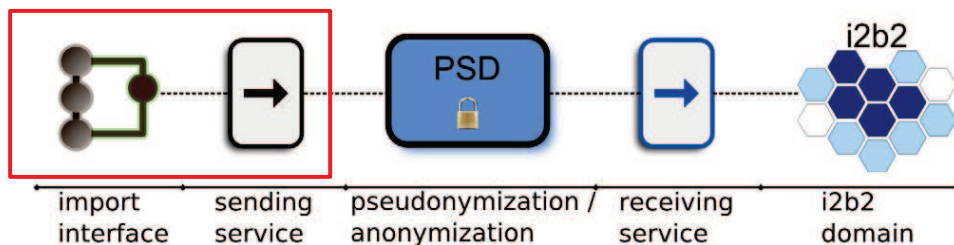
Workflow zur Übertragung von Studiendaten ins i2b2



6

Import-Schnittstelle

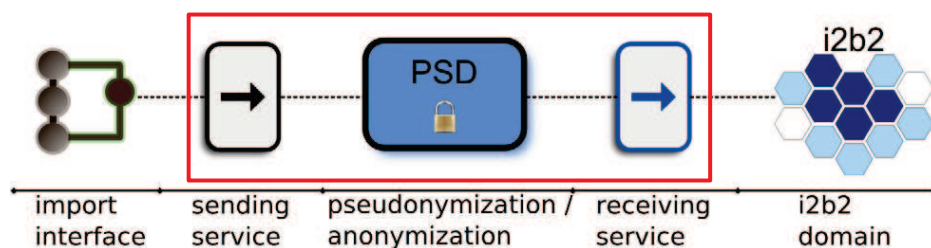
- CDISC ODM-Format 1.2 und 1.3
- Generischer Ansatz: adaptierbar durch Plug-Ins für andere Datentypen
- Datenausgabe :
 - SQL-Skripte, die direkt in das i2b2 importiert werden können
 - Pro Import-Datei eine Ontologie
 - Patientenbezogene SQL-Statements
- Patientenbezogene Daten dienen als Input für den PSD



7

Pseudonymisierung der Daten

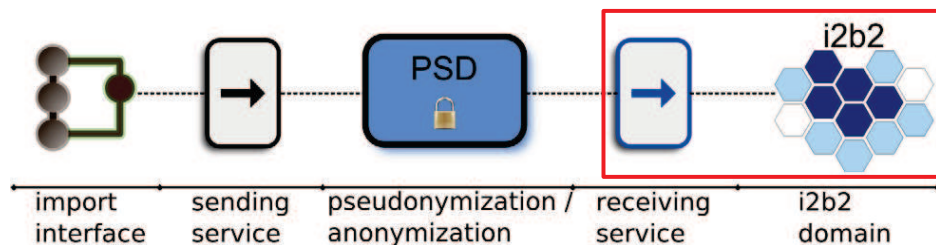
1. Python Script: Erstellt XML Nachricht aus den SQL-Statements
2. Sending Service: Schickt die Nachricht an den PSD
3. PSD: Ersetzt Pseudonym des Patienten durch ein Pseudonym der 2. Stufe
4. PSD: Schickt die Nachricht an den Receiving Service
5. Receiving Service: Speichert die Daten in einem Verzeichnis



8

Import in i2b2

1. Python Script: Extrahiert die SQL-Statements
2. Python Script: Erstellt eine Patienten-Mapping-Tabelle (PSN <-> i2b2-Nummer)
3. Python Script: Ersetzt den Patientenidentifikator in den SQL-Statements durch die i2b2-Nummer
4. Manuell: Import der SQL-Statements in die i2b2-Datenbank



9

Diskussion

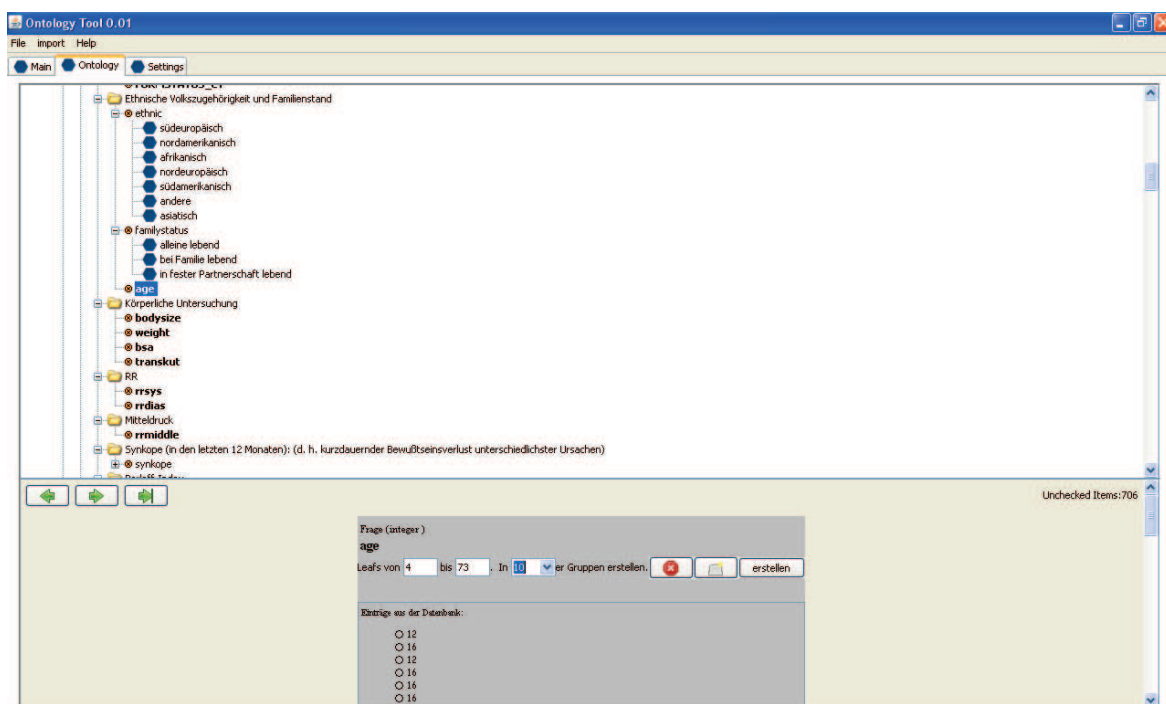
- Das Ziel, i2b2 für eine datenschutzkonforme Bereitstellung von Forschungsdaten einzusetzen, wurde erreicht.
- Importschnittstelle:
 - Nur Items mit Codelist werden automatisch erstellt
 - Funktioniert nur gut für Datenformate mit vorgegebenen Strukturen bzw. Metainformationen
- Pseudonymisierung:
 - Umsetzung aller Funktionen des PSD
 - Direkte Integration in die Services des PSD
 - Aufrufen der Funktionen aus i2b2

Ausblick

- Kombination Studiendaten mit Daten aus der Biomaterialverwaltung
- Kombination Bilddaten mit Studiendaten
- GUI für die Importschnittstelle

11

GUI Import-Schnittstelle



12