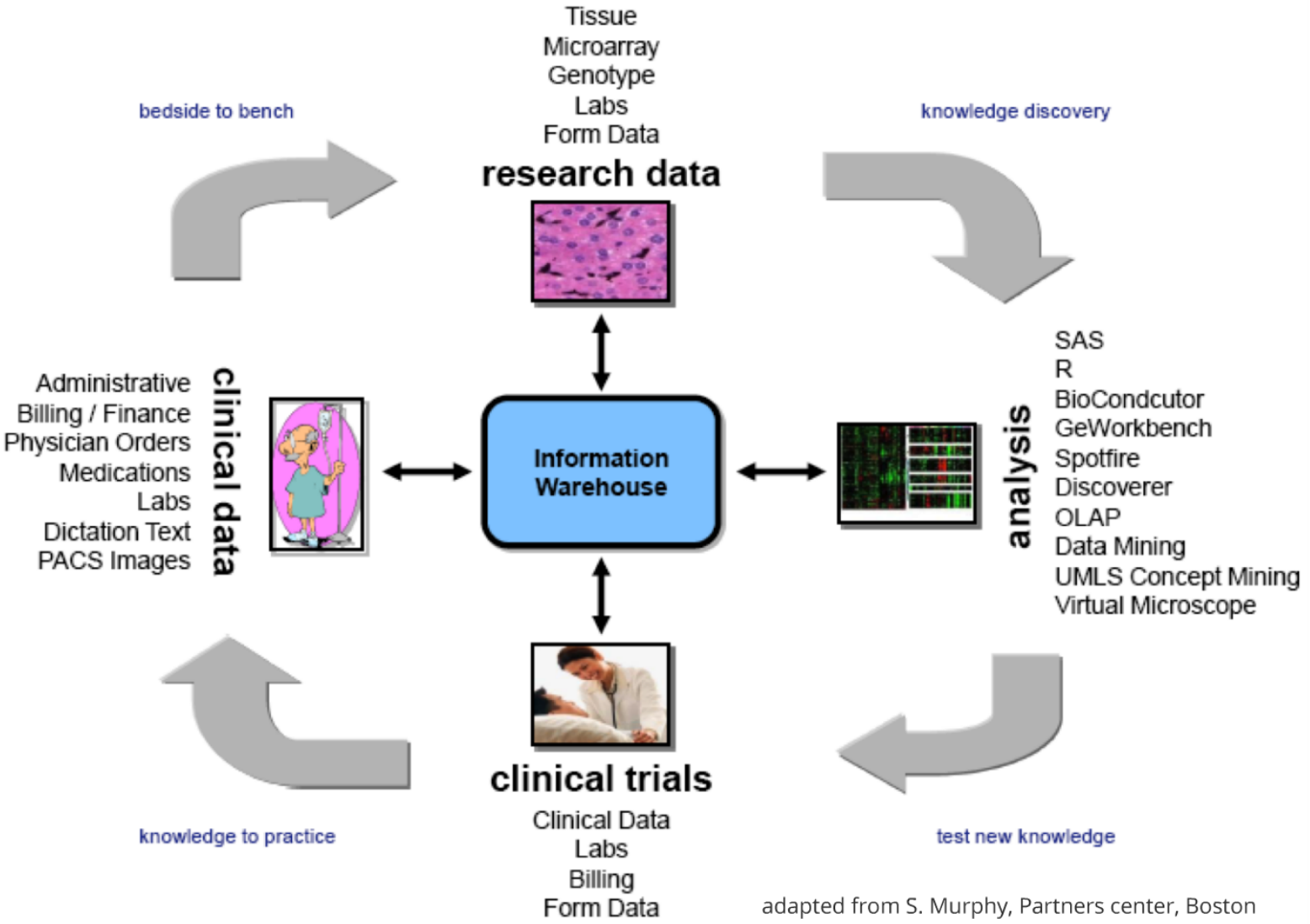
6. TMF-Jahreskongress 02.04.2014

Standardisierte Datenaufbereitung und Zusammenführung für die Forschung

TMF-Projekt V091-MI IDRT *Integrated Data Repository Toolkit*

Ulrich Sax, Thomas Ganslandt und Matthias Löbe,
Christian Bauer, Benjamin Baum, Matthias Quade,
Matthias Löbe, Igor Engel, Sebastian Stäubert,
Sebastian Mate, Jan Christoph, Alfred Winter, Hans-Ulrich Prokosch

Problemstellung
dezentrale Forschungsinfrastruktur

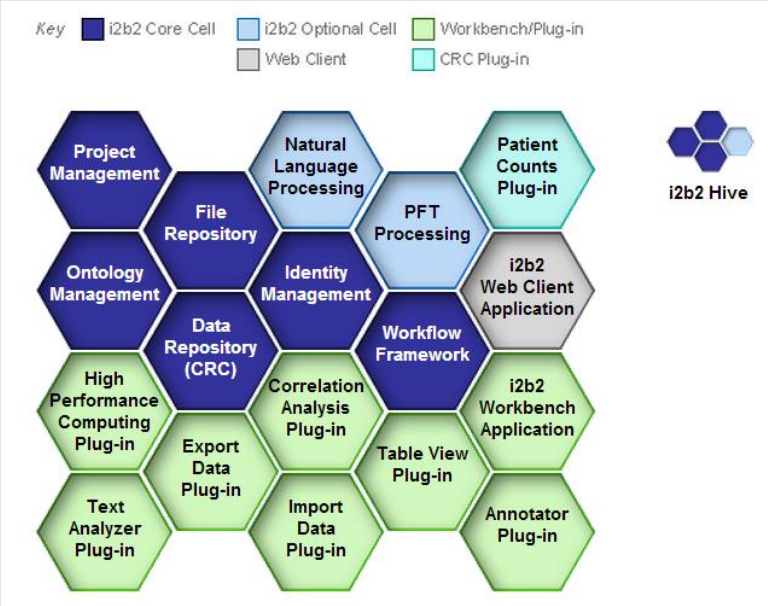


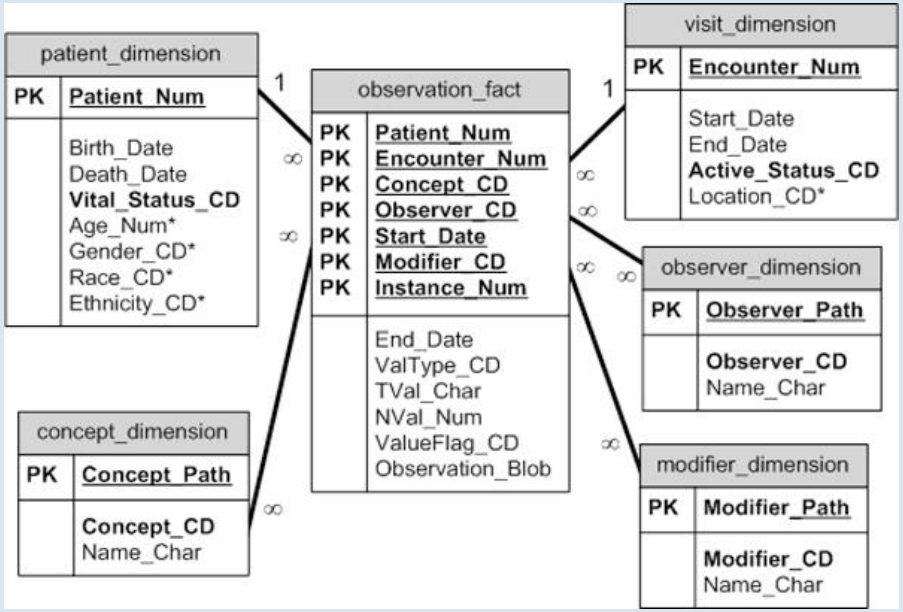
adapted from S. Murphy, Partners center, Boston

R. Bellazzi, Pavia as shown at 1st Europ. I2b2 AUG 2013

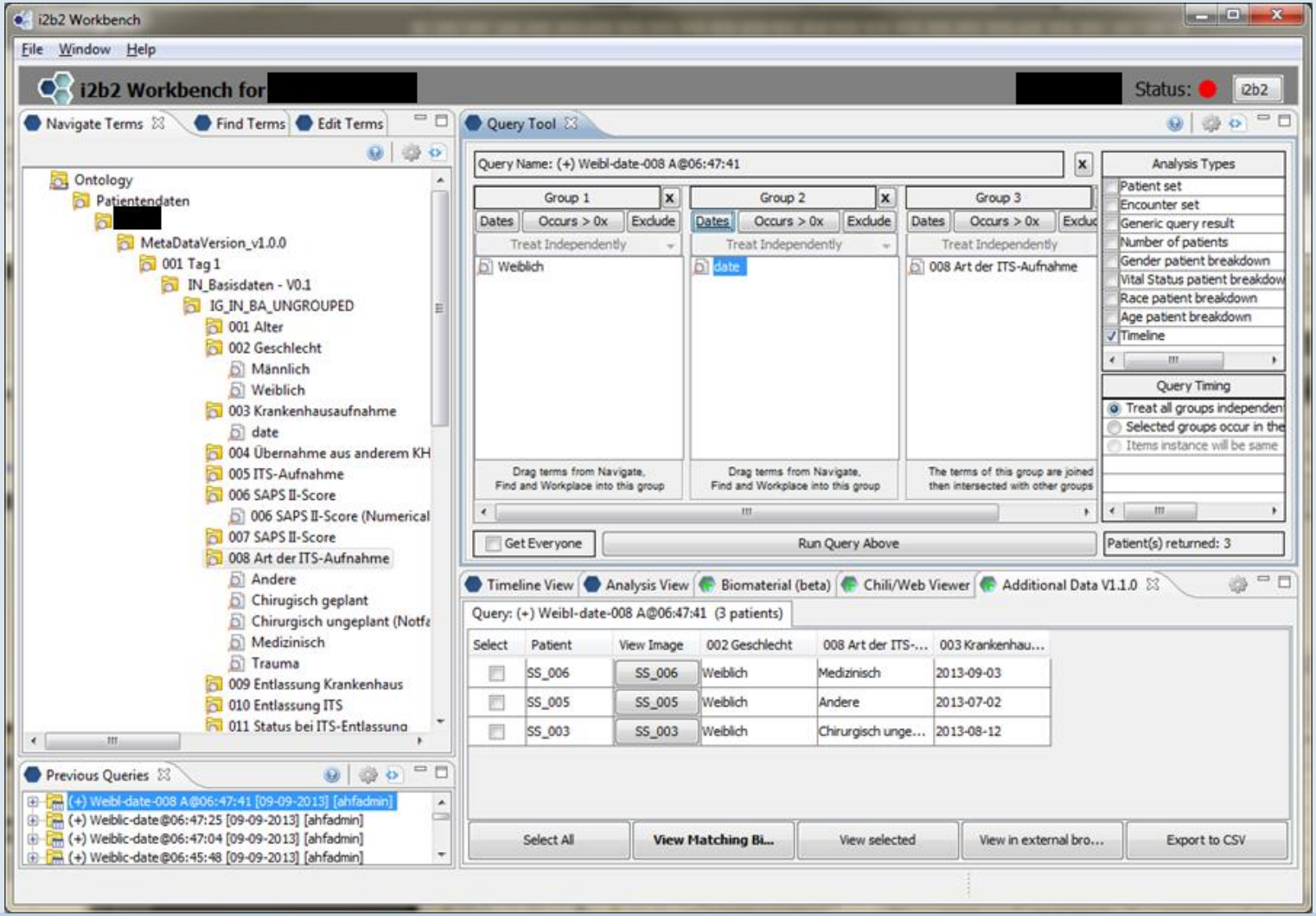
i2b2 (Informatics for Integrating Biology and the Bedside)

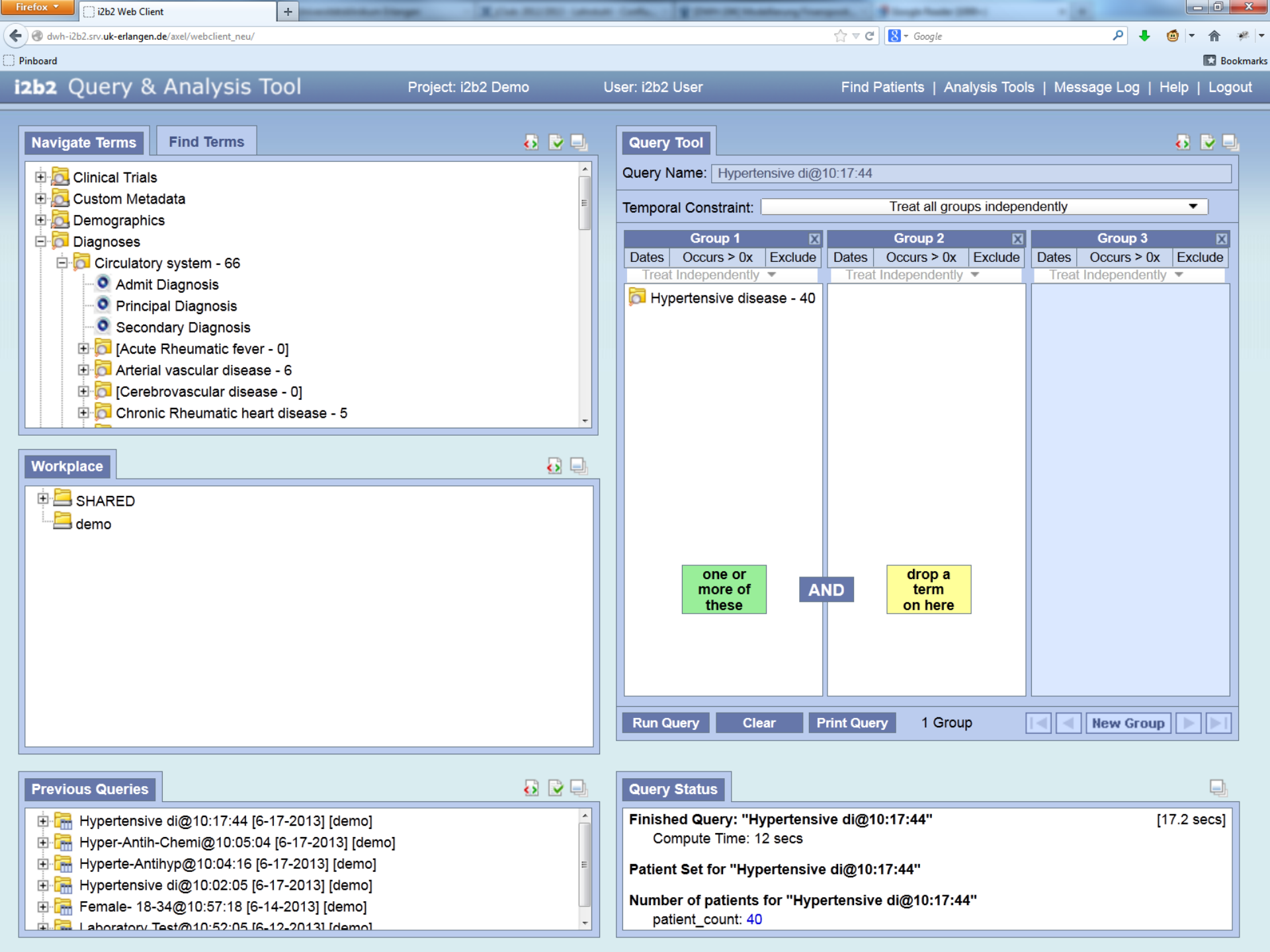
- Modulares Open Source Data Warehouse für medizinische Daten
- Design:
  - Erweiterbares Hive-Design als Anwendungsserver
  - Durch Plugins erweiterbare Abfragetools
  - Patientendaten werden in einer großen EAV Tabelle abgelegt, spezielle medizinische Daten (Visiten, …) werden in verlinkten Tabellen abgelegt
  - Patientdenaten werden durch einen hierarischen Konzeptbaum durchsuchbar gemacht (i2b2 ontology)





i2b2 (Informatics for Integrating Biology and the Bedside)




Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Google
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Clinical Trials
Custom Metadata
Demographics
Diagnoses
Circulatory system - 66
Admit Diagnosis
Principal Diagnosis
Secondary Diagnosis
[Acute Rheumatic fever - 0]
Arterial vascular disease - 6
[Cerebrovascular disease - 0]
Chronic Rheumatic heart disease - 5
Workplace
SHARED
demo
Query Tool
Query Name: Hypertensive di@10:17:44
Temporal Constraint: Treat all groups independently
Group 1
Group 2
Group 3
Dates
Occurs > 0x
Exclude
Treat Independently
Hypertensive disease - 40
one or more of these
AND
drop a term on here
Run Query
Clear
Print Query
1 Group
New Group
Previous Queries
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Hypertensive di@10:02:05 [6-17-2013] [demo]
Female- 18-34@10:57:18 [6-14-2013] [demo]
Query Status
Finished Query: "Hypertensive di@10:17:44"
[17.2 secs]
Compute Time: 12 secs
Patient Set for "Hypertensive di@10:17:44"
Number of patients for "Hypertensive di@10:17:44"
patient_count: 40

i2b2 Query & Analysis Tool

Navigate Terms | Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
  - Circulatory system - 66
    - Admit Diagnosis
    - Principal Diagnosis
    - Secondary Diagnosis
    - [Acute Rheumatic fever - 0]
    - Arterial vascular disease - 6
    - [Cerebrovascular disease - 0]
    - Chronic Rheumatic heart disease - 5
    - Disease of capillaries - 1
    - [Diseases of pulmonary circulation - 0]
    - Hypertensive disease - 40
    - Ischemic heart disease - 17
    - Other forms of heart disease - 35

Workplace

- SHARED
- demo

Previous Queries

- Hypertensive di@10:17:44 [6-17-2013] [demo]
  - Results of Hypertensive di@10:17:44 [6-17-2013] [demo]
    - Patient Set for "Hypertensive di@10:17:44" - FINISHED
    - Number of patients for "Hypertensive di@10:17:44" - FINISHED
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]

Demographics

Specify Data | View Results | Plugin Help

Patient Count: 40

Age in Years

10-20	3
20-30	4
30-40	5
40-50	8
50-60	4
60-70	10
70-80	3
80-90	3

Sex

F	17
M	23

Race

asian	7
black	16
hispanic	10
indian	3
white	4

Language

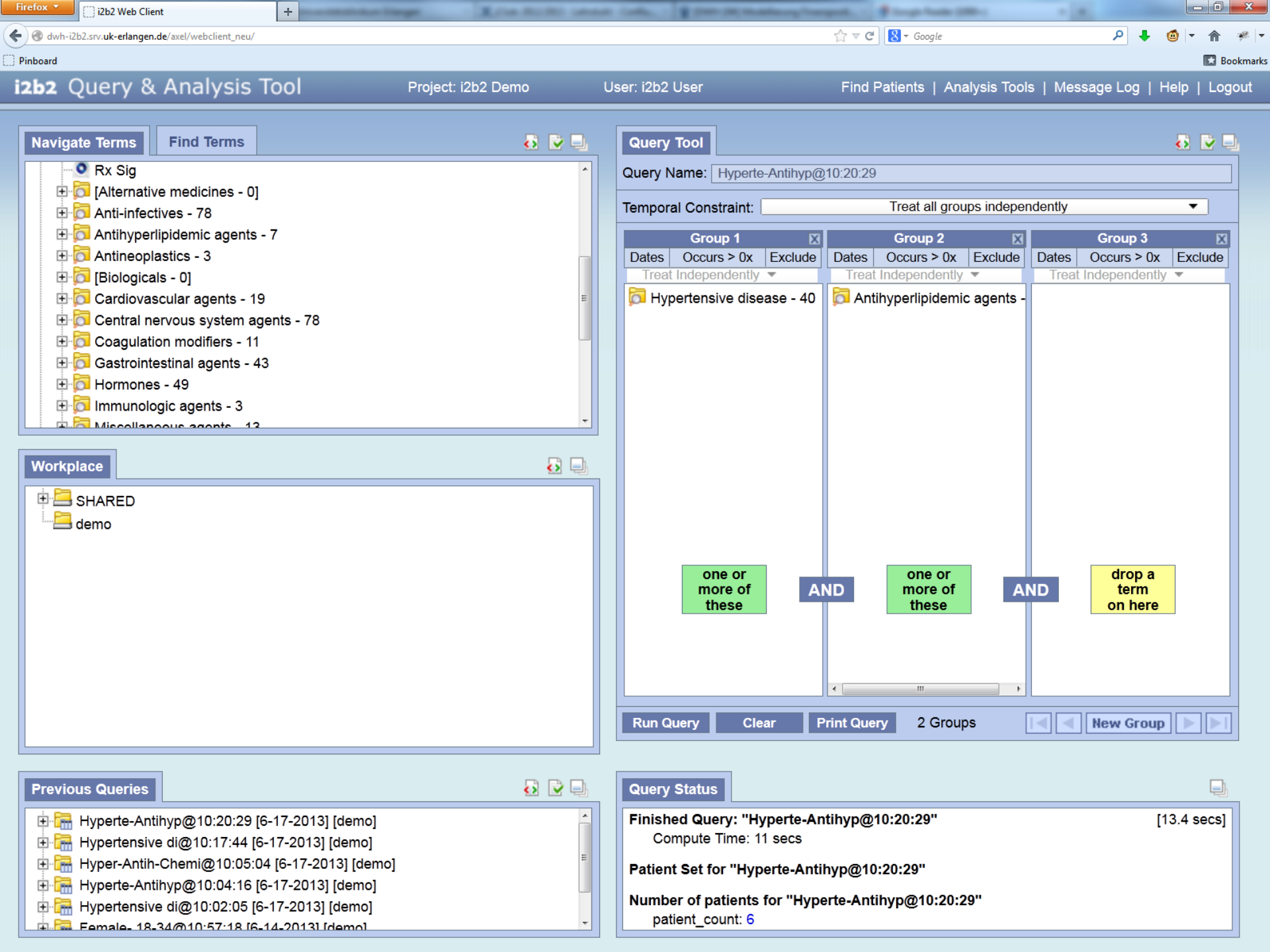
english	14
german	17
spanish	9

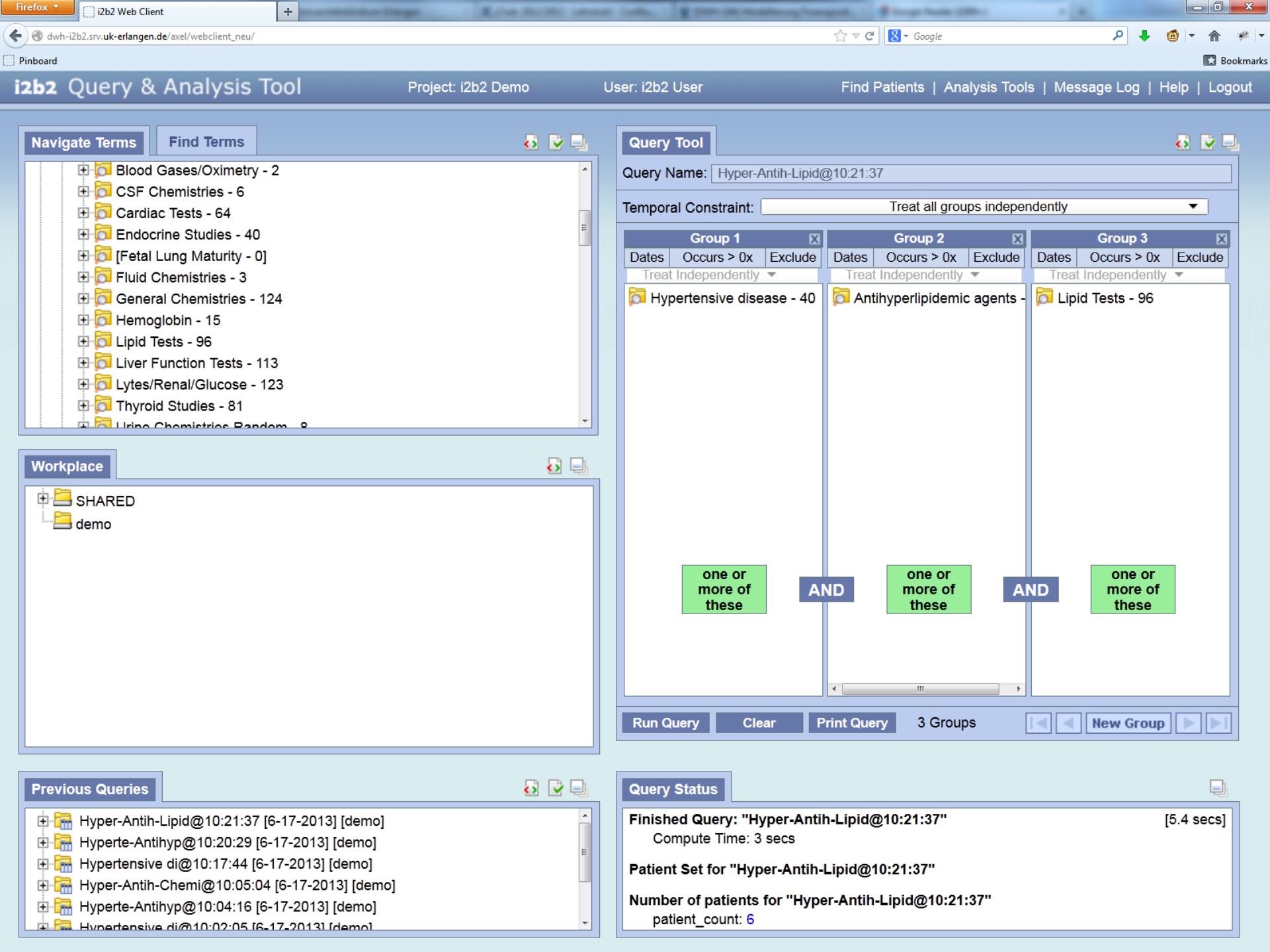
Plugins

Detailed List View Category: ALL

ExportXLS
This plugin populates a table of selectable concepts/observations and provides the possibility to download the result as a file.

Demographics (1 Patient Set) - Simple Counts
This plugin displays a demographic break-down of a Patient Set.


Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Rx Sig
[Alternative medicines - 0]
Anti-infectives - 78
Antihyperlipidemic agents - 7
Antineoplastics - 3
[Biologicals - 0]
Cardiovascular agents - 19
Central nervous system agents - 78
Coagulation modifiers - 11
Gastrointestinal agents - 43
Hormones - 49
Immunologic agents - 3
Workplace
SHARED
demo
Query Tool
Query Name: Hyperte-Antihyp@10:20:29
Temporal Constraint: Treat all groups independently
Group 1
Group 2
Group 3
Dates
Occurs > 0x
Exclude
Treat Independently
Hypertensive disease - 40
Antihyperlipidemic agents -
one or more of these
AND
one or more of these
AND
drop a term on here
Run Query
Clear
Print Query
2 Groups
New Group
Previous Queries
Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Hypertensive di@10:02:05 [6-17-2013] [demo]
Query Status
Finished Query: "Hyperte-Antihyp@10:20:29"
[13.4 secs]
Compute Time: 11 secs
Patient Set for "Hyperte-Antihyp@10:20:29"
Number of patients for "Hyperte-Antihyp@10:20:29"
patient_count: 6


Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Google
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Blood Gases/Oximetry - 2
CSF Chemistries - 6
Cardiac Tests - 64
Endocrine Studies - 40
[Fetal Lung Maturity - 0]
Fluid Chemistries - 3
General Chemistries - 124
Hemoglobin - 15
Lipid Tests - 96
Liver Function Tests - 113
Lytes/Renal/Glucose - 123
Thyroid Studies - 81
Workplace
SHARED
demo
Query Tool
Query Name: Hyper-Antih-Lipid@10:21:37
Temporal Constraint: Treat all groups independently
Group 1
Group 2
Group 3
Dates
Occurs > 0x
Exclude
Treat Independently
Hypertensive disease - 40
Antihyperlipidemic agents -
Lipid Tests - 96
one or more of these
AND
one or more of these
AND
one or more of these
Run Query
Clear
Print Query
3 Groups
New Group
Previous Queries
Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Query Status
Finished Query: "Hyper-Antih-Lipid@10:21:37"
[5.4 secs]
Compute Time: 3 secs
Patient Set for "Hyper-Antih-Lipid@10:21:37"
Number of patients for "Hyper-Antih-Lipid@10:21:37"
patient_count: 6

Navigate Terms | Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Providers
- Reports
- Visit Details

Workplace

- SHARED
- demo

Timeline

Specify Data | View Results | Plugin Help

Drop a Patient Set and one or more Concepts (Ontology Terms) into the input boxes below, and then click the "View Results" tab to view a timeline showing when those concepts were observed in the selected patient set.

Patient Set: Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo] [PATIENT

Concept(s):

- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Reports

Click a concept to remove it from the list.

Previous Queries

- Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
- Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
- Hypertensive di@10:17:44 [6-17-2013] [demo]
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
- Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
- Hypertensive di@10:02:05 [6-17-2013] [demo]

Plugins

Detailed List View Category: ALL

Timeline
This plugin creates a visual representation of when selected observations occur within a given patient set.

Project Request
This plugin is uesed to generate a request for a new project based on concepts, patients and sets..

i2b2 Query & Analysis Tool

Navigate Terms / Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Providers
- Reports
- Visit Details

Workplace

- SHARED
- demo

Timeline

Specify Data | View Results | Plugin Help

<<< start: 1 size: 10 go >>> zoom: - + pan: < >

4/3/1997 — 11/11/2003 — 6/22/2010

Person_#1000000025__Female__76yrold__Hispanic
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Person_#1000000031__Male__50yrold__Hispanic
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Person_#1000000054__Female__66yrold__Asian
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Person_#1000000083__Female__65yrold__Black
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Previous Queries

- Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
- Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
- Hypertensive di@10:17:44 [6-17-2013] [demo]
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
- Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
- Hypertensive di@10:02:05 [6-17-2013] [demo]

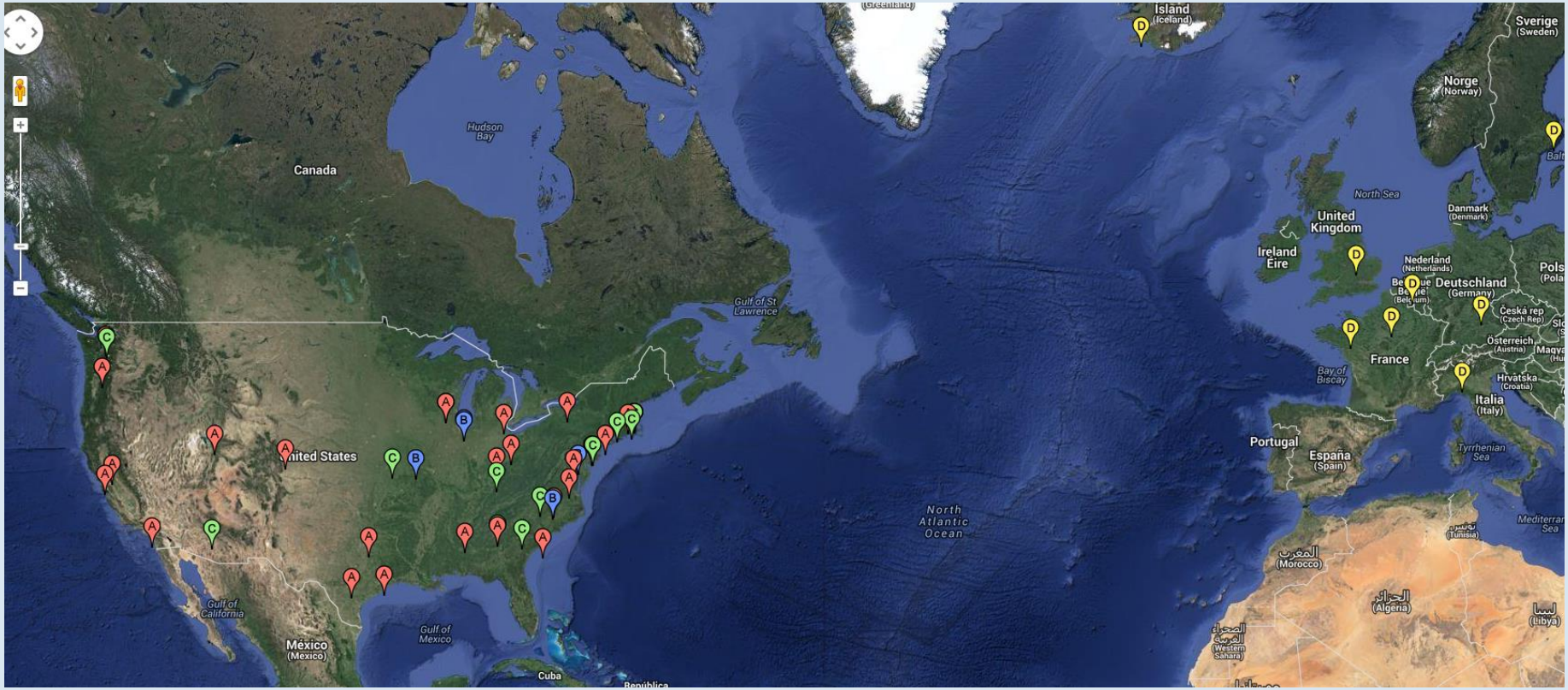
Plugins

Detailed List View — Category: ALL

Timeline
This plugin creates a visual representation of when selected observations occur within a given patient set.

Project Request
This plugin is uesed to generate a request for a new project based on concepts, patients and sets..

i2b2
Installationen weltweit



i2b2 shortcomings

- i2b2 Plattform fehlt …
  - einfache Installation
  - einfaches Laden von Daten typischer Formate
  - Erstellung von (komplexen) i2b2-Hirachien
  - Semantische Integration

- Lösung:
  - Werzeug-Set, dass i2b2 in eine einfach einzusetzende Lösung für Datenzusammen-führung, Filterung und Abfrage und Weitergabe an statistische Analyse macht.
  - Integrated Data Repository Toolkit (IDRT)
    - TMF gefördertes Projekt der Universitäten Erlangen, Leipzig und Göttingen

Extraktion

ODM

CSV/SQL

§21

Biomaterial

Aufbereitung

Standard-Terminologien

MDR-Anbindung

Ontologie-Editor

PIDgen-Anbindung

Eigenst. Objekte

Support

Deliverables

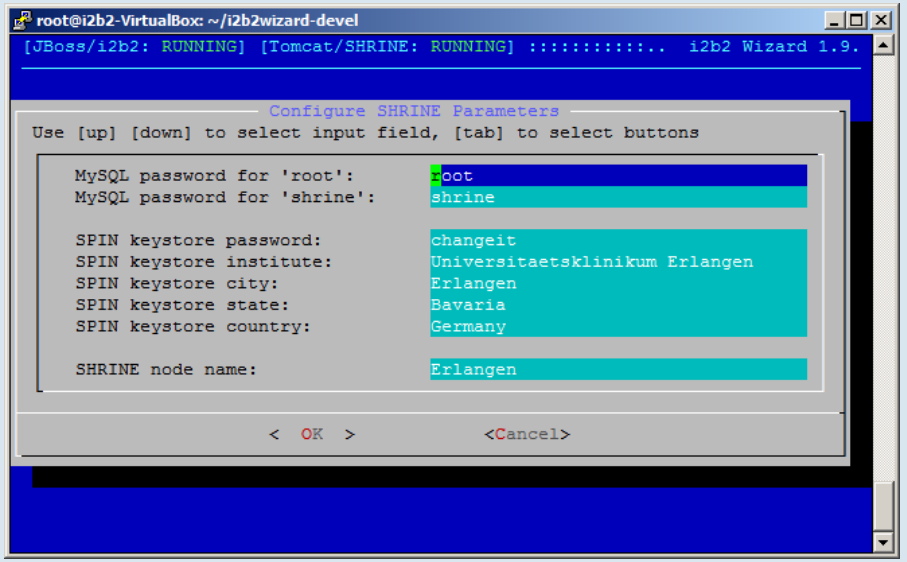
Best Practice-Guides

Dissemination

Setup & Konfiguration

Ziele

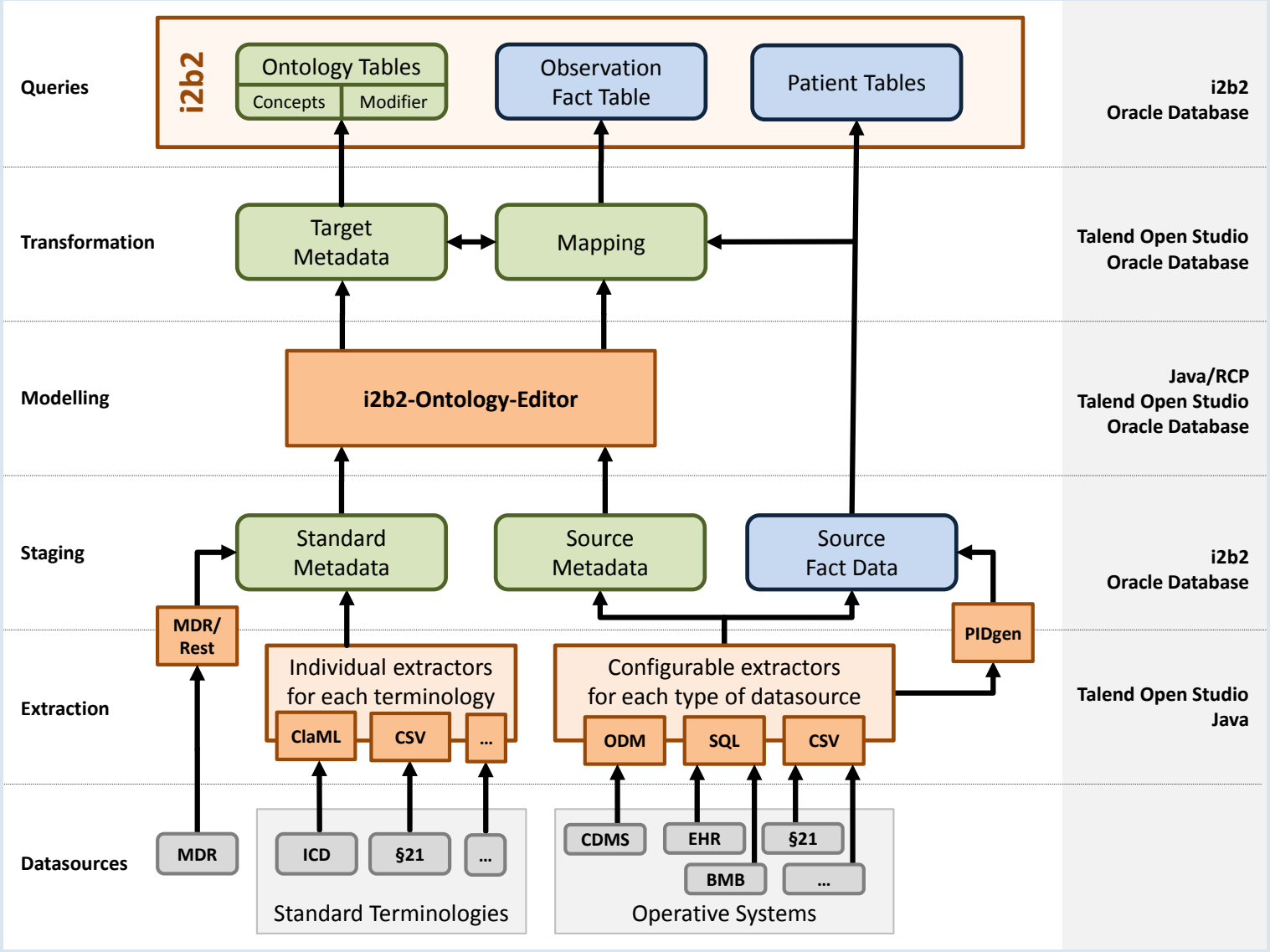
- Setup automatisieren
- Konfiguration von Projekten und Nutzern erleichtern



Umsetzung

- Shellscript für Linux & Windows
  - modulares Konzept
  - unterstützt verschiedene Betriebssystem- und Datenbanktypen
- automatische Installation der Plattform in ~30m
- einfache Verwaltung von Projekten, Nutzern & Rechten
- Unterstützung von SHRINE (föderierte Datamarts)
- stark nachgefragtes Werkzeug (auch international)

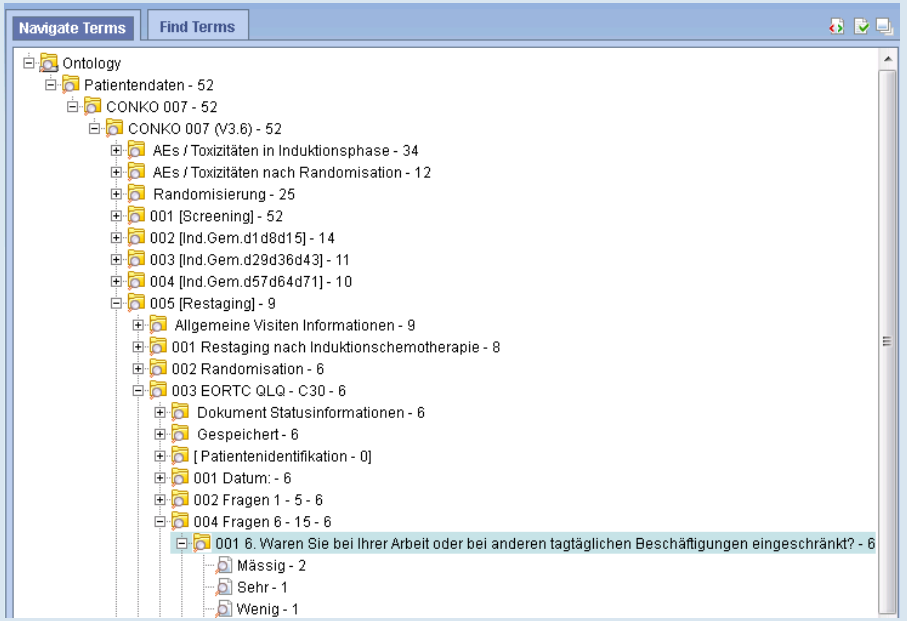
Ergebnisse
ETL-Strecke: Architektur



Ergebnisse
ETL-Strecke: ODM-Extraktor

Ziele

- Import von Studiendaten im CDISC ODM-Format
  - weit verbreitet
  - strukturiert



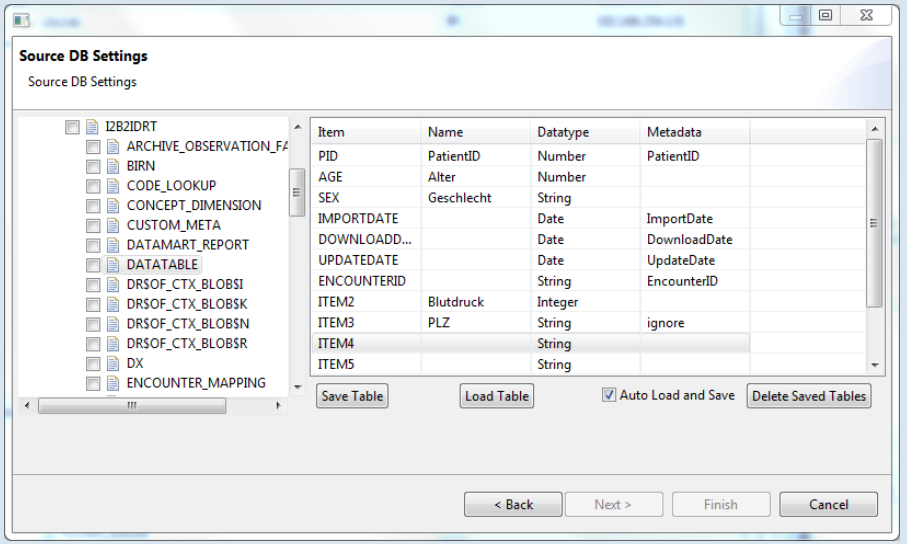
Umsetzung

- Talend Open Studio-Job
  - konfigurierbar
  - wiederverwendbar in eigenen Jobstreams
  - Integration in "Standalone" IDRT-Frontend
- generiert hierarchische i2b2-Ontologie
- Unterstützung für ODM v1.3
  - getestet mit Exporten aus secuTrial, OpenClinica

Ergebnisse
ETL-Strecke: CSV/SQL-Extraktor

Ziele

- Import von Routinedaten aus CSV-Dateien und SQL-Datenbanken
  - unstrukturiert
  - flexibel



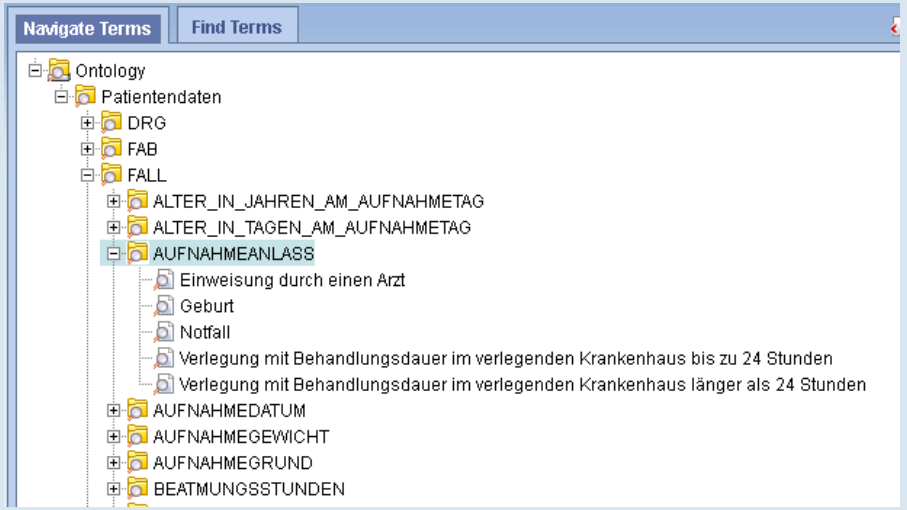
Umsetzung

- Talend Open Studio-Job
  - wie ODM-Extraktor
- generiert flache i2b2-Ontologie
  - ein Ast pro Quelltabelle + Knoten für Felder & Ausprägungen
  - Ableitung aus Konfigurationsdatei, Feldnamen und -inhalten
- Unterstützung für beliebige Quelldatenbanken (JDBC-Connector)

Ergebnisse
ETL-Strecke: §21-Extraktor

Ziele

- Import des Benchmarking-Datensatz nach KHEntgG §21
  - weit verbreitet
  - limitierter Datenumfang
  - ideal für "quick win" bei Projektstart

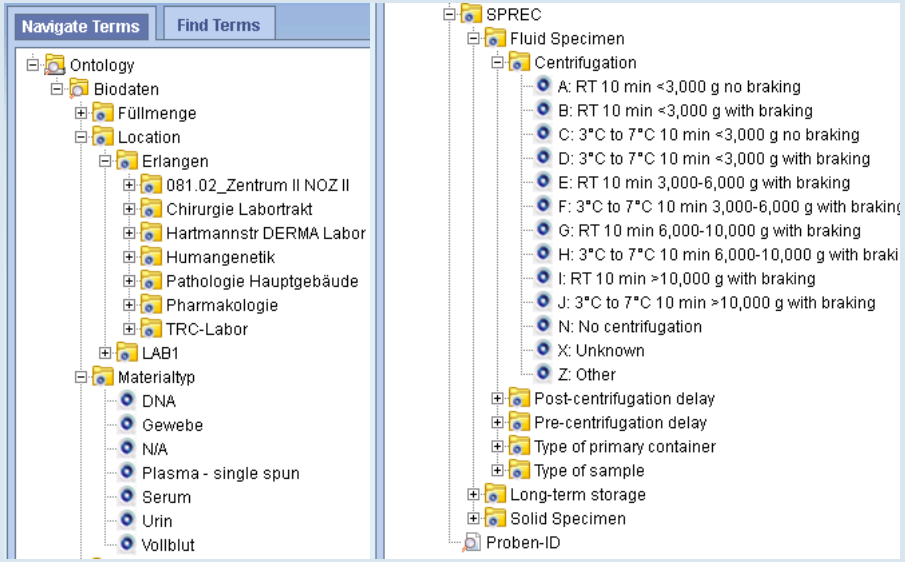


Umsetzung

- Talend Open Studio-Job
  - wie ODM-Extraktor
- verknüpft mehrere IDRT-Teilpakete
  - Vorparametrierung für CSV-Extraktor
  - Einbindung relevanter Standardterminologien (ICD, OPS, DRG)
- Unterstützung für verschiedene Jahresversionen des §21

Ziele

- Import von Biomaterialdaten
  - generischer Ansatz
  - Umsetzung am Beispiel Starlims



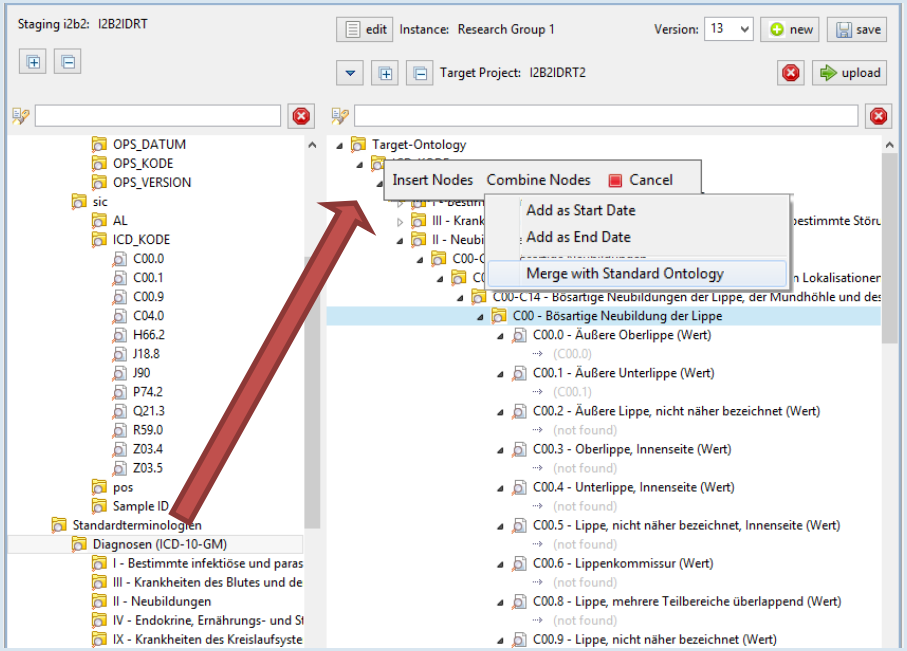
Umsetzung

- Talend Open Studio-Job
  - wie ODM-Extraktor
- importiert relevante Quelltabellen aus Starlims Biorepository
- generiert hierarchische i2b2-Ontologie
  - Abbildung flexibler Lagerhierarchien
  - strukturierte Darstellung des SPREC-Codes (präanalytisches Probenhandling)
  - Übernahme projektspezifischer Probenmetadaten
- Unterstützung von Modifiern und eigenständigen i2b2-Objekten

Ergebnisse
i2b2-Ontologieeditor

Ziele

- Nachbearbeitung importierter i2b2-Daten
  - Strukturierung der i2b2-Ontologie
  - Zusammenführung von Datenelementen
  - Zuordnung von Attributen (z.B. Datum)



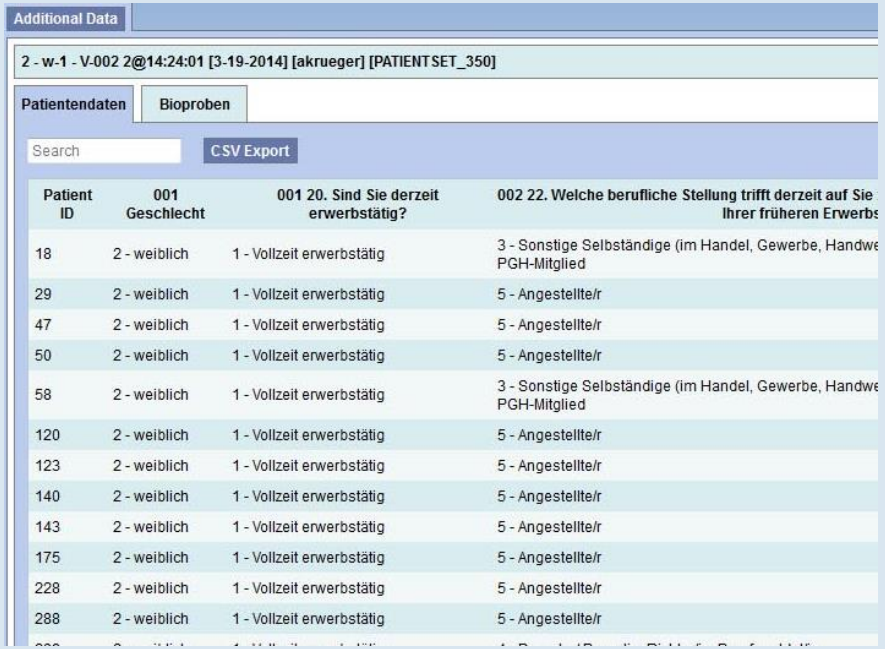
Umsetzung

- Java Eclipse-Frontend
- grafische Benutzeroberfläche
  - Integration der IDRT-Extraktoren & Standardterminologie-Loader
  - Laden von Quell- und Zielontologien
  - Erstellung von Mappings per Drag & Drop
- Überführung von Quelldaten in die Zielontologie

Ergebnisse
Unterstützung eigenständiger Objekte

Ziele

- "Nachrüsten" komplexer Datenelemente in i2b2
  - "Modifier-Codes" erlauben ab i2b2 v1.6 multiple Attribute für ein Objekt
  - begrenzte Unterstützung für Anzeige und Export in i2b2



Patient ID	001 Geschlecht	001 20. Sind Sie derzeit erwerbstätig?	002 22. Welche berufliche Stellung trifft derzeit auf Sie Ihrer früheren Erwerbs
18	2 - weiblich	1 - Vollzeit erwerbstätig	3 - Sonstige Selbständige (im Handel, Gewerbe, Handwe PGH-Mitglied
29	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
47	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
50	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
58	2 - weiblich	1 - Vollzeit erwerbstätig	3 - Sonstige Selbständige (im Handel, Gewerbe, Handwe PGH-Mitglied
120	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
123	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
140	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
143	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
175	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
228	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r
288	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellte/r

Umsetzung

- Erweiterung der IDRT-Komponenten
  - Unterstützung von Modifiern
- Plugin für i2b2-Webclient
  - getrennte Darstellung von Konzepten und Modifiern
  - Bündelung in Karteireitern für die Anzeige
  - CSV-Export

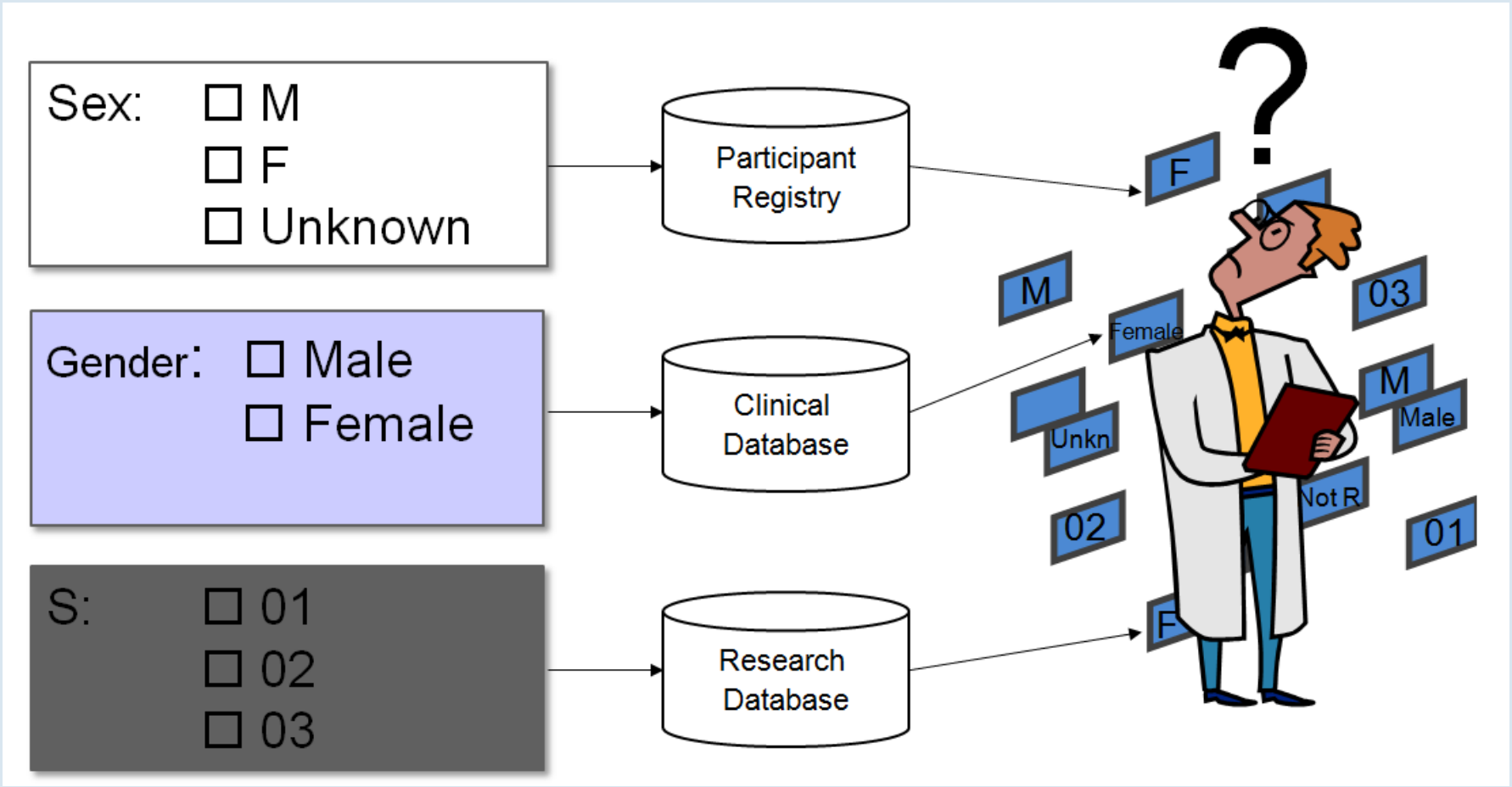
Standardterminologien

Was ist das?

- Standardterminologien sind kontrollierte Ordnungssysteme für medizinische Konzepte
- Variablen bzw. Datenfelder in Studien-CRFs, Registerbögen oder KIS-Formularen werden (zum Teil) durch harmonisierte Konzepte oder Codes beschrieben
- Diagnosen
- Prozeduren
- Laborwerte
- Medikation
- Leistungsverrechnung
- Biomaterial
- Bilddaten
- Psycho-soziale Instrumente

Eine Welt OHNE Metadaten…

- Interpretation der Datenfelder nur durch Erhebenden

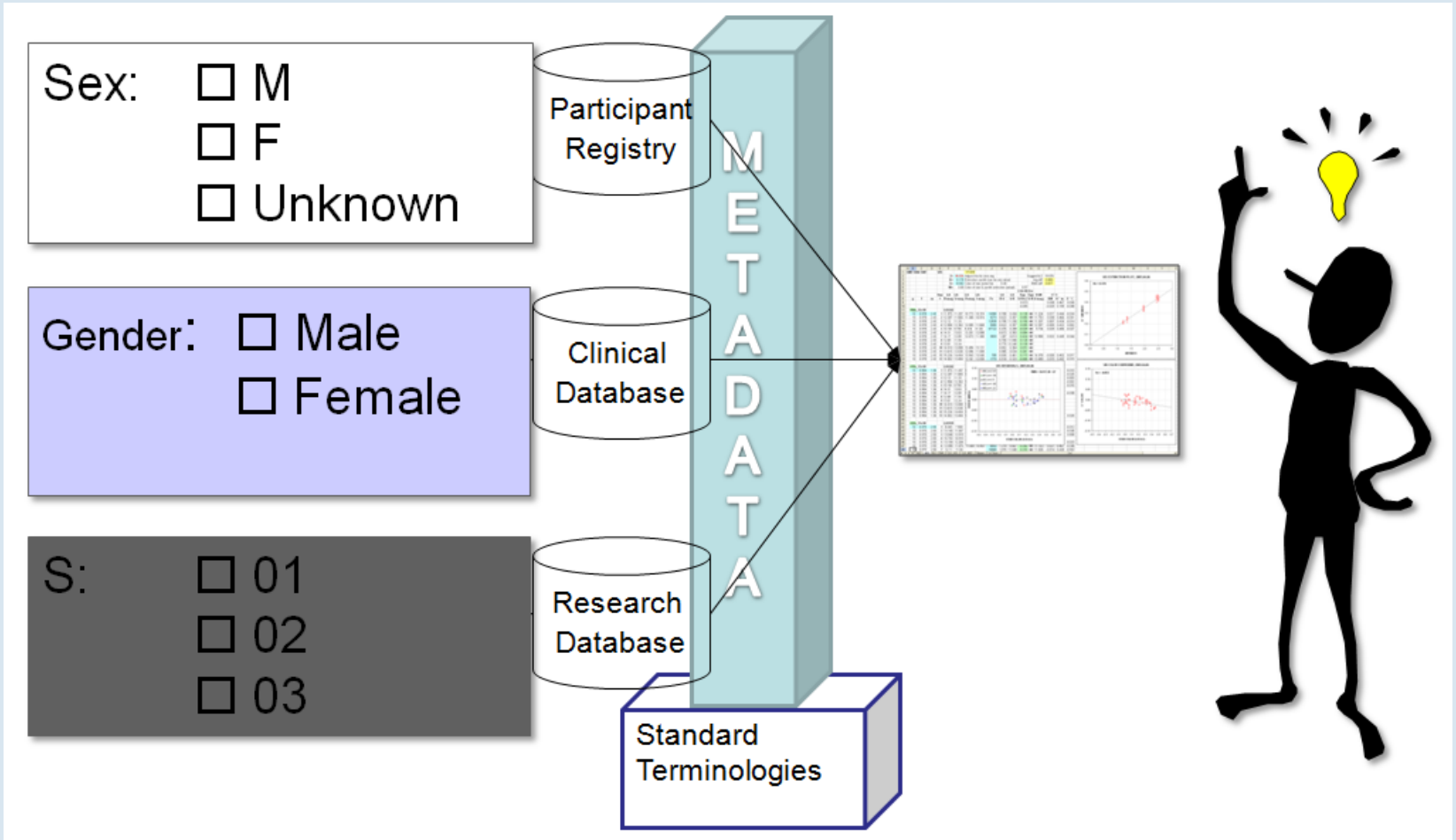


Quelle: NCI caBIG Training 1000: Introduction to caCORE

Eine Welt MIT Metadaten…

Terminologien sollen Mehrdeutigkeiten beseitigen

- Interpretation der Datenfelder ist jedermann möglich

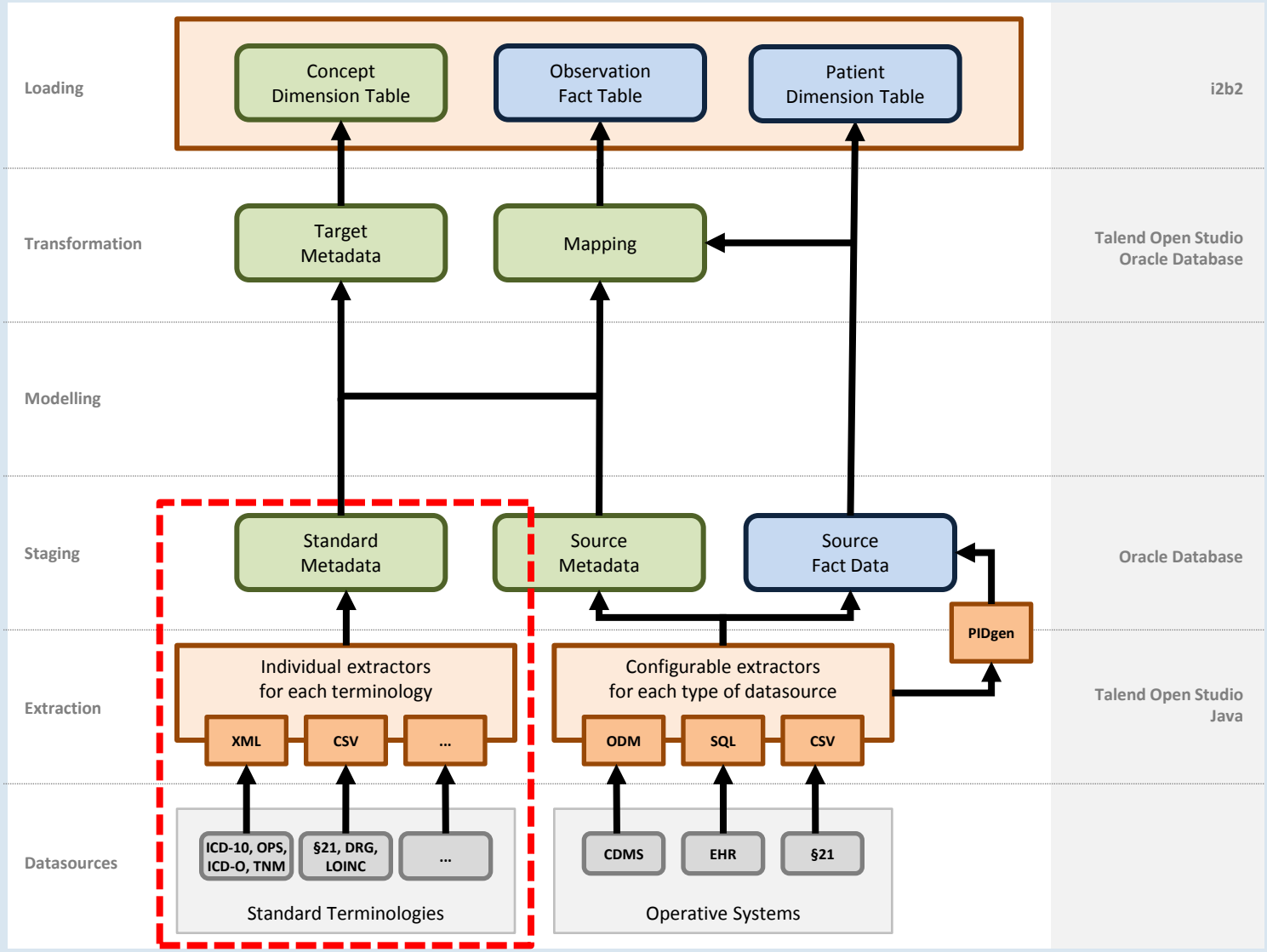


Quelle: NCI caBIG Training 1000: Introduction to caCORE

Standardterminologien in i2b2

- i2b2 enthält nur „amerikanische“ Terminologien (ICD-9, National Drug File, RadLex, Zip-Codes)
- Deutschland verwendet eigene Terminologien
  - Unterschiedliche T. in Versorgung und Forschung
- Ziel: Verfügbarmachung dieser Terminologien in i2b2
  - **ICD-10-GM** zur Kodierung von Krankheitsdiagnosen und **OPS** zur Kodierung von Operationen und Prozeduren als die in Deutschland am häufigsten verwendeten Systeme in der Krankenversorgung
  - **DRG-Schlüssel** und **§21-Codelisten** (Aufnahmegrund etc.) als Systeme zur Leistungsabrechnung, die in jedem Krankenhaus zur Verfügung stehen
  - **LOINC** für die Kodierung von Laborwerten und Dokumenten
  - **TNM** und **ICD-O** für onkologische Datensammlungen
  - **MedDRA** zur Kodierung von unerwünschten Ereignissen bzw. Nebenwirkungen besonders in klinischen Studien mit Menschen
- Weitere i.V.: SNOMED CT, ATC, SPREC, CDISC SDTM …
  - Nur lohnend, wenn entsprechende Daten vorhanden sind

Einbettung in IDRT Architektur



Beispiel OPS

Schritt 1: Download von DIMDI-Homepage

- OPS – Operationen- und Prozedurenschlüssel
- ClaML – Classification Markup Language
  - ↳XML-basierter Standard für medizinische Terminologien

Beispiel OPS

Schritt 2: XSL-Transformation von ClaML in einfaches XML

```
<xsl:template match="/">

  <!-- xsl:text instructions provide nice formatting for generated code -->
  <xsl:text>&#xA;</xsl:text><ops><xsl:text>&#xA;</xsl:text>

  <!-- Only <Class> elements of kind 'category' are parsed into concepts -->
    <xsl:for-each select="//Class[@kind='category']"><!-- //Class[not(./SubClas

    <!-- Add surrounding <class> element -->
    <xsl:text>&#9;</xsl:text><class><xsl:text>&#xA;</xsl:text>

      <!-- Add <code> element-->
      <xsl:text>&#9;&#9;</xsl:text><code><xsl:value-of select="@code"/></co><xsl:te

      <!-- Add <label> element-->
      <xsl:text>&#9;&#9;</xsl:text><label><xsl:value-of select="Rubric[@kind='preferr
      </xsl:text>

      <!-- Add <path> element-->
      <xsl:text>&#9;&#9;</xsl:text><path>

        <!-- Write path prefix and call template "i2b2path"-->
        <xsl:text>\i2b2\ST\OPS</xsl:text>
        <xsl:call-template name="i2b2path">
          <xsl:with-param name="code" select="@code"/>
```

- XSL-Stylesheets für Ordner und Blätter der i2b2-Ontologie

```
54509    <Class code="5-351" kind="category">
54510      <Meta name="P17b-d" value="2"/>
54511      <Meta name="DRG" value="J"/>
54512      <SuperClass code="5-35"/>
54513      <SubClass code="5-351.0"/>
54514      <SubClass code="5-351.1"/>
54515      <SubClass code="5-351.2"/>
54516      <SubClass code="5-351.3"/>
54517      <SubClass code="5-351.4"/>
54518      <SubClass code="5-351.x"/>
54519      <SubClass code="5-351.y"/>
54520      <Rubric kind="preferred">
54521        <Label xml:lang="de" xml:space="default">Ersatz von H
54522      </Rubric>
54523      <Rubric kind="note">
54524        <Label xml:lang="de" xml:space="default">Die Anwendur
54525      </Rubric>
```

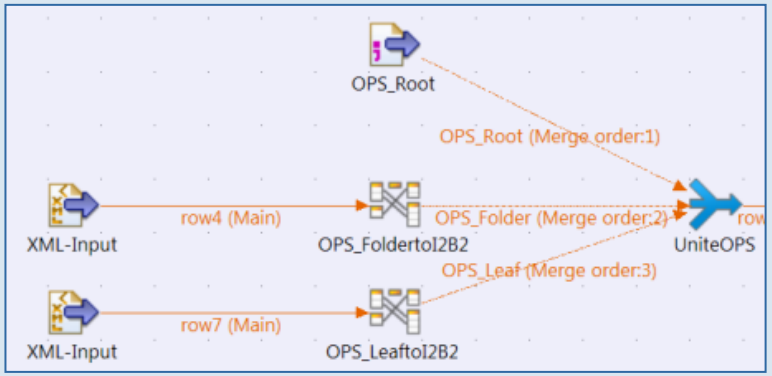
- Normatives ClaML als Eingabedatei

```
28419    <class>
28420      <code>5-351</code>
28421      <label>Ersatz von Herzklappen durch Prothese
28422      <path>\i2b2\ST\OPS\5\5-35...5-37\5-35\5-351\
28423      <level>7</level>
28424    </class>
28425    <class>
28426      <code>5-351.0</code>
28427      <label>Aortenklappe</label>
28428      <path>\i2b2\ST\OPS\5\5-35...5-37\5-35\5-351\
28429      <level>8</level>
28430    </class>
28431    <class>
28432      <code>5-351.1</code>
```

- XML-Zwischenformat

Beispiel OPS (3)

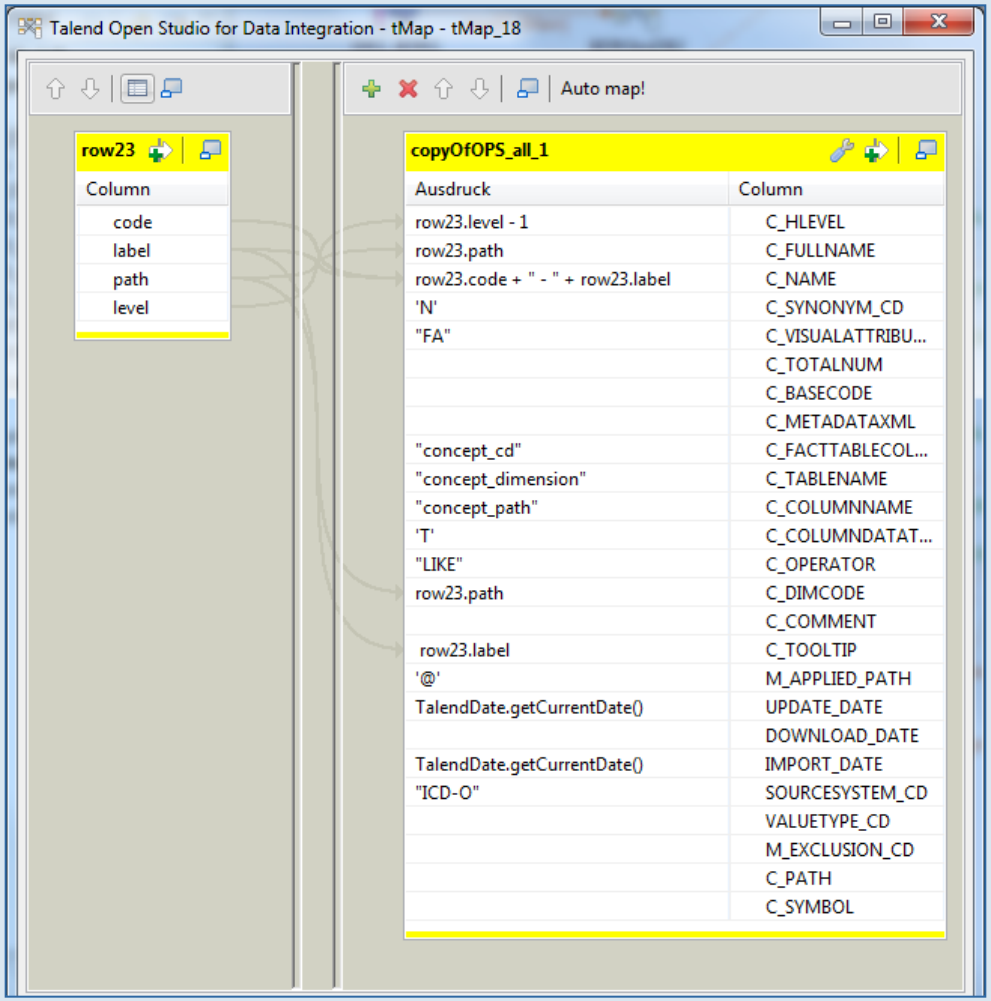
Schritt 3: Abbildung in i2b2 Databankschema



- Talend Open Studio

```
28419    <class>
28420        <code>5-351</code>
28421        <label>Ersatz von Herzklappen durch Prothese
28422        <path>\i2b2\ST\OPS\5\5-35...5-37\5-35\5-351\
28423        <level>7</level>
28424    </class>
28425    <class>
28426        <code>5-351.0</code>
28427        <label>Aortenklappe</label>
28428        <path>\i2b2\ST\OPS\5\5-35...5-37\5-35\5-351\
28429        <level>8</level>
28430    </class>
28431    <class>
28432        <code>5-351.1</code>
```

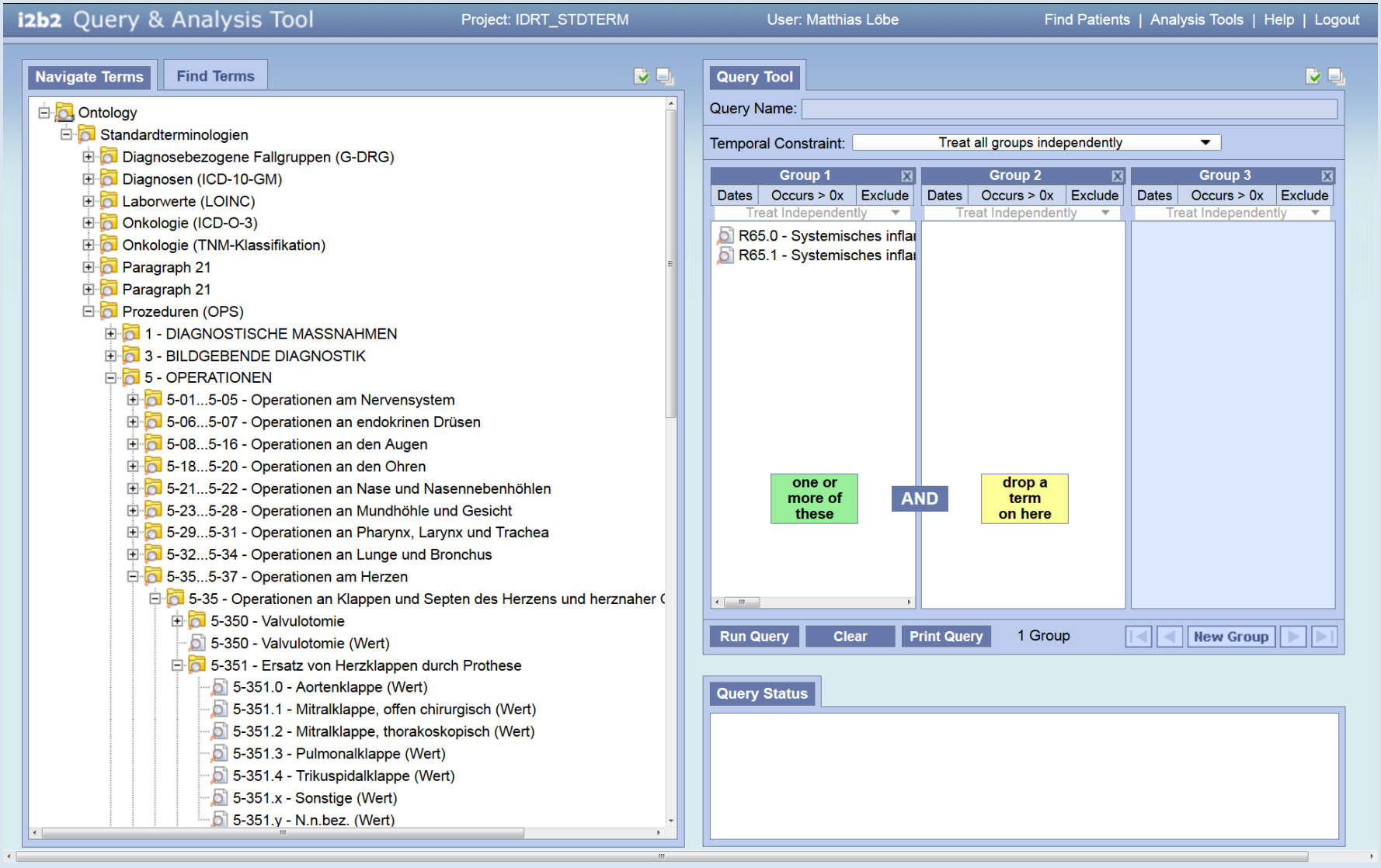
- XML-Zwischenformat



- Schemamapping nach i2b2

Beispiel OPS

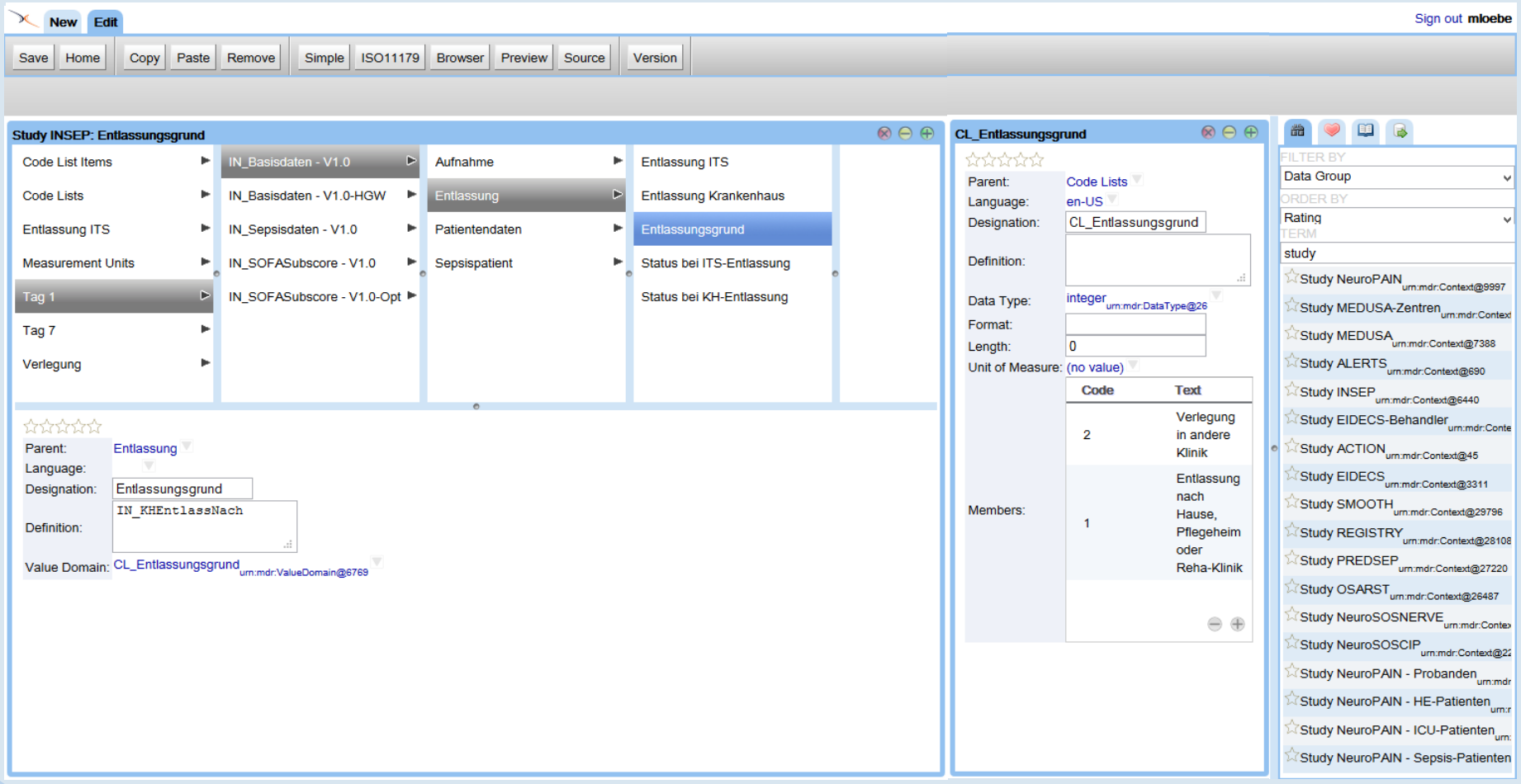
Schritt 4: Verfügbarkeit in i2b2



Nationales Metadata Repository (MDR)

Standardisierung und semantische Fundierung

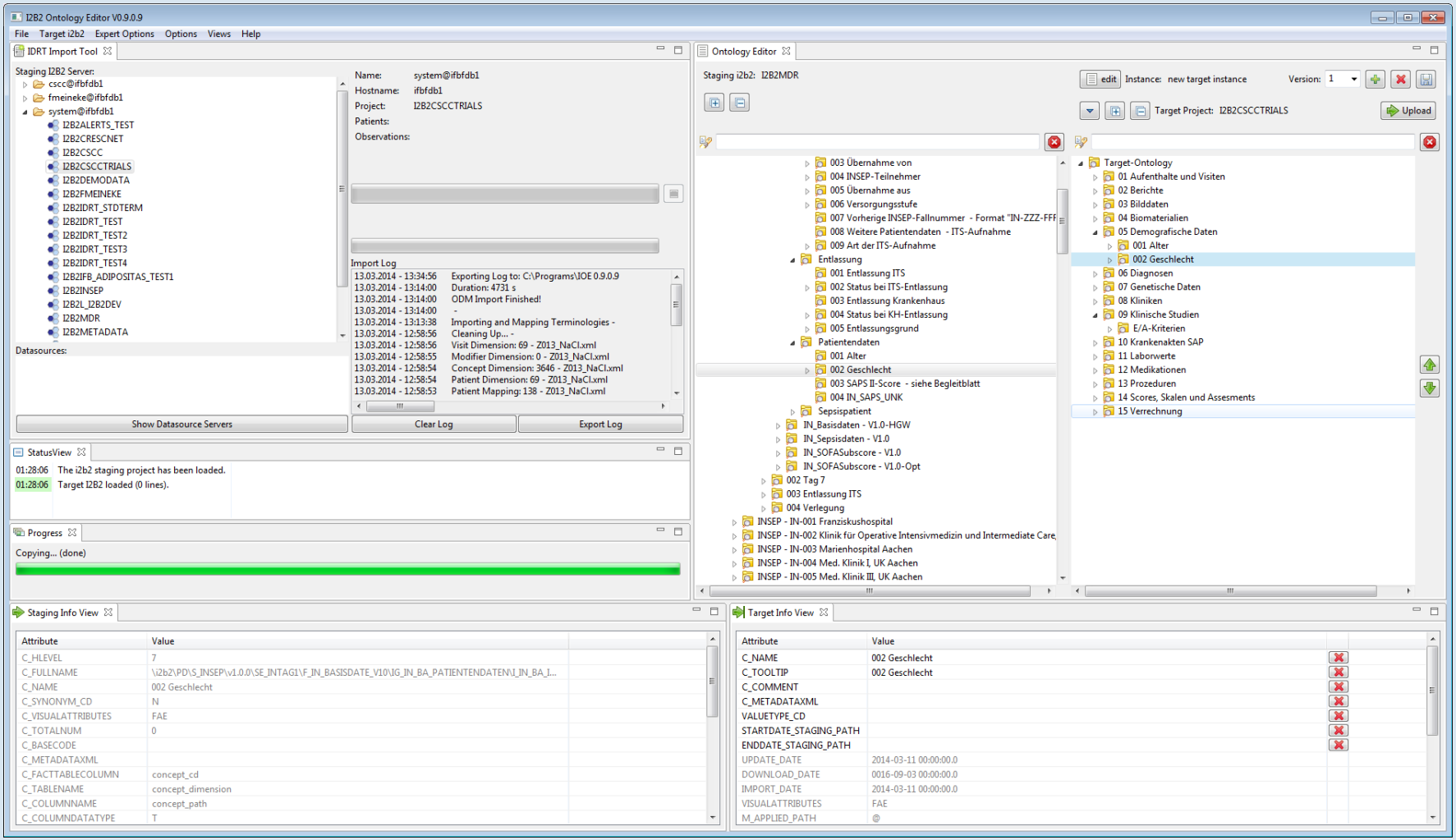
- Datenelemente, Wertelisten und Formulare zentral speichern
- Qualität, Effizienz, Vergleichbarkeit und Zugreifbarkeit steigern



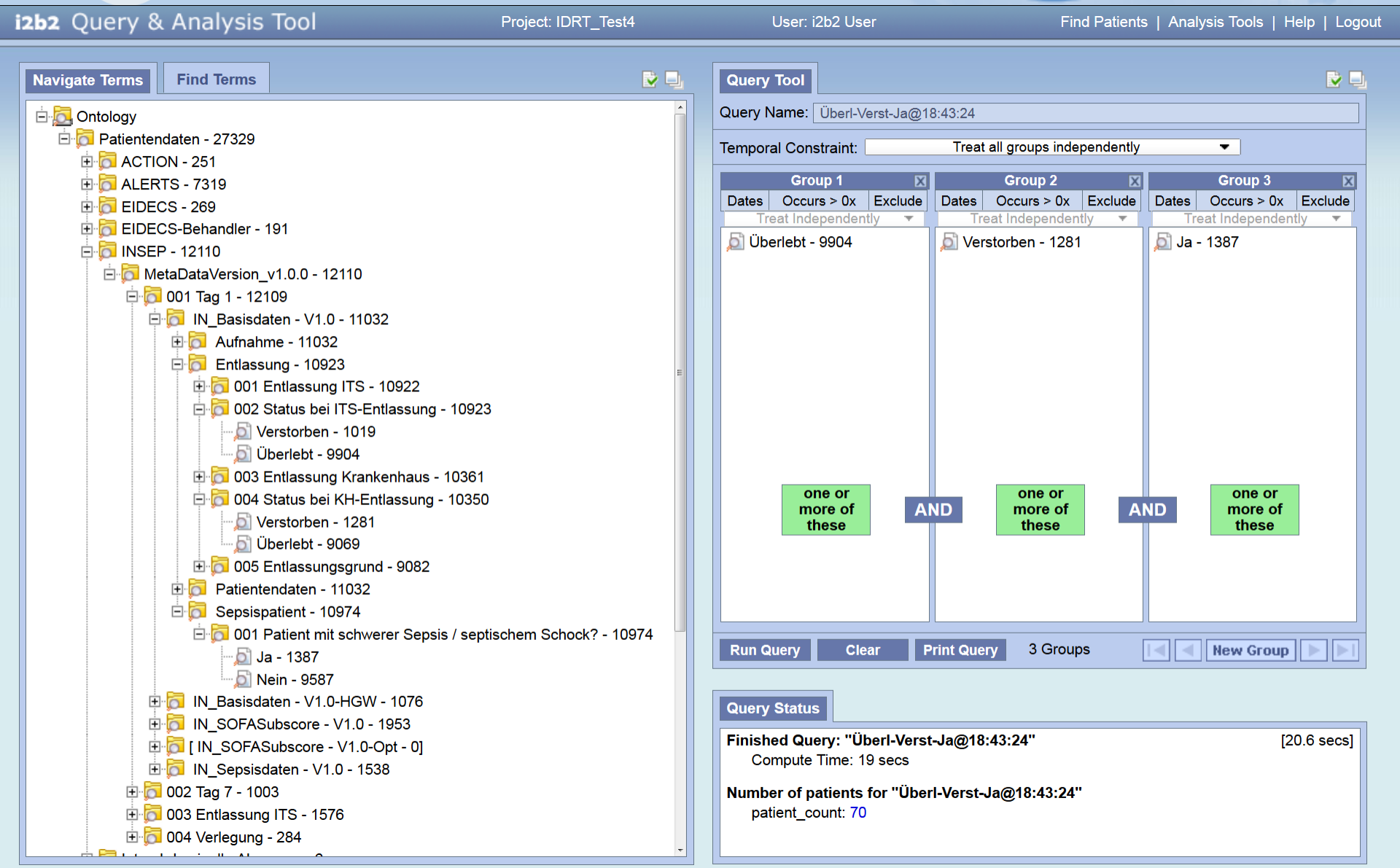
MDR-Anbindung im IDRT Import Tool

Aktuell in Arbeit

- MDR-Server als Metadatenquelle, Mapping auf klinische Fakten

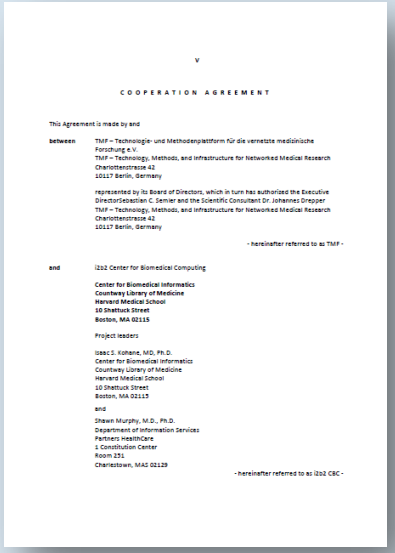


Studien mit MDR-Beschreibung in i2b2



Ausblick
Dissemination

- Kooperationsvertrag zwischen TMF und i2b2 Center for Biomedical Computing
- IDRT Präsentationen
  - ↳i2b2 AUG Konferenz Boston 10.07.2014
  - ↳IDRT2-Abschlussworkshop 10.09.2014
  - ↳2. EU i2b2 AUG Konferenz 11.09.2014



COOPERATION AGREEMENT

This Agreement is made by and

between TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.
TMF – Technology, Methods, and Infrastructure for Networked Medical Research
Charlottenstrasse 42
10117 Berlin, Germany

represented by its Board of Directors, which in turn has authorized the Executive Director Sebastian C. Semler and the Scientific Consultant Dr. Johannes Drepper
TMF – Technology, Methods, and Infrastructure for Networked Medical Research
Charlottenstrasse 42
10117 Berlin, Germany

- hereinafter referred to as TMF -

and i2b2 Center for Biomedical Computing

Center for Biomedical Informatics
Countway Library of Medicine
Harvard Medical School
10 Shattuck Street
Boston, MA 02115

Project leaders

Isaac S. Kohane, MD, Ph.D.
Center for Biomedical Informatics
Countway Library of Medicine
Harvard Medical School
10 Shattuck Street
Boston, MA 02115

and

Shawn Murphy, M.D., Ph.D.
Department of Information Services
Partners HealthCare
1 Constitution Center
Room 231
Charlestown, MAS 02129

- hereinafter referred to as i2b2 CBC -

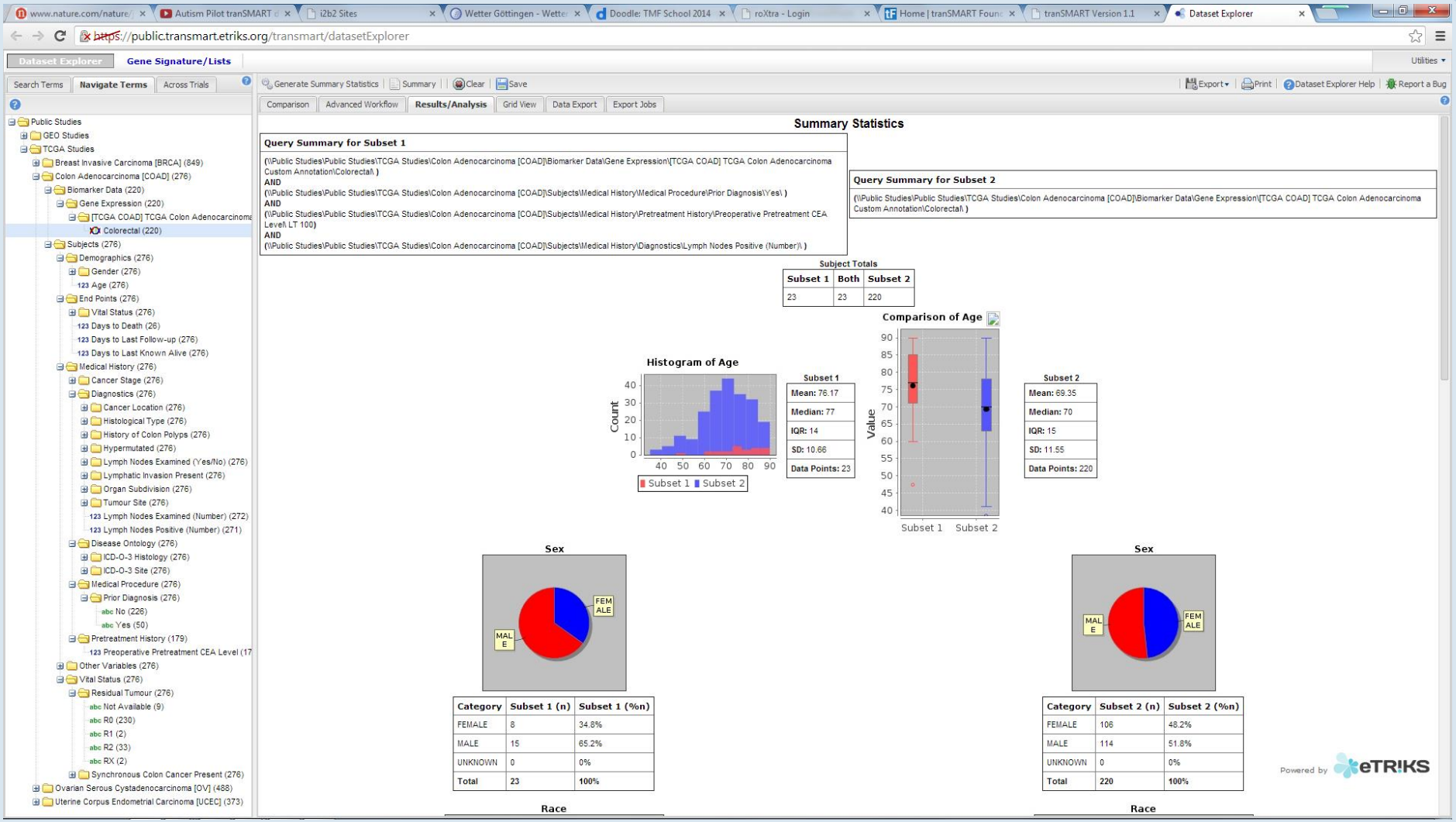
Ausblick
Weiterentwicklung von i2b2

- NIH-funded National Center for Biomedical Computing (NCBC)
- Erste Förderphase ausgelaufen
- Verstetigung in den Harvard Medical-Häusern
- Große Community weltweit
- Jetzt: Center for Biomedical Computing

- Derzeit Version 1.7.01, Version 2 in Vorbereitung

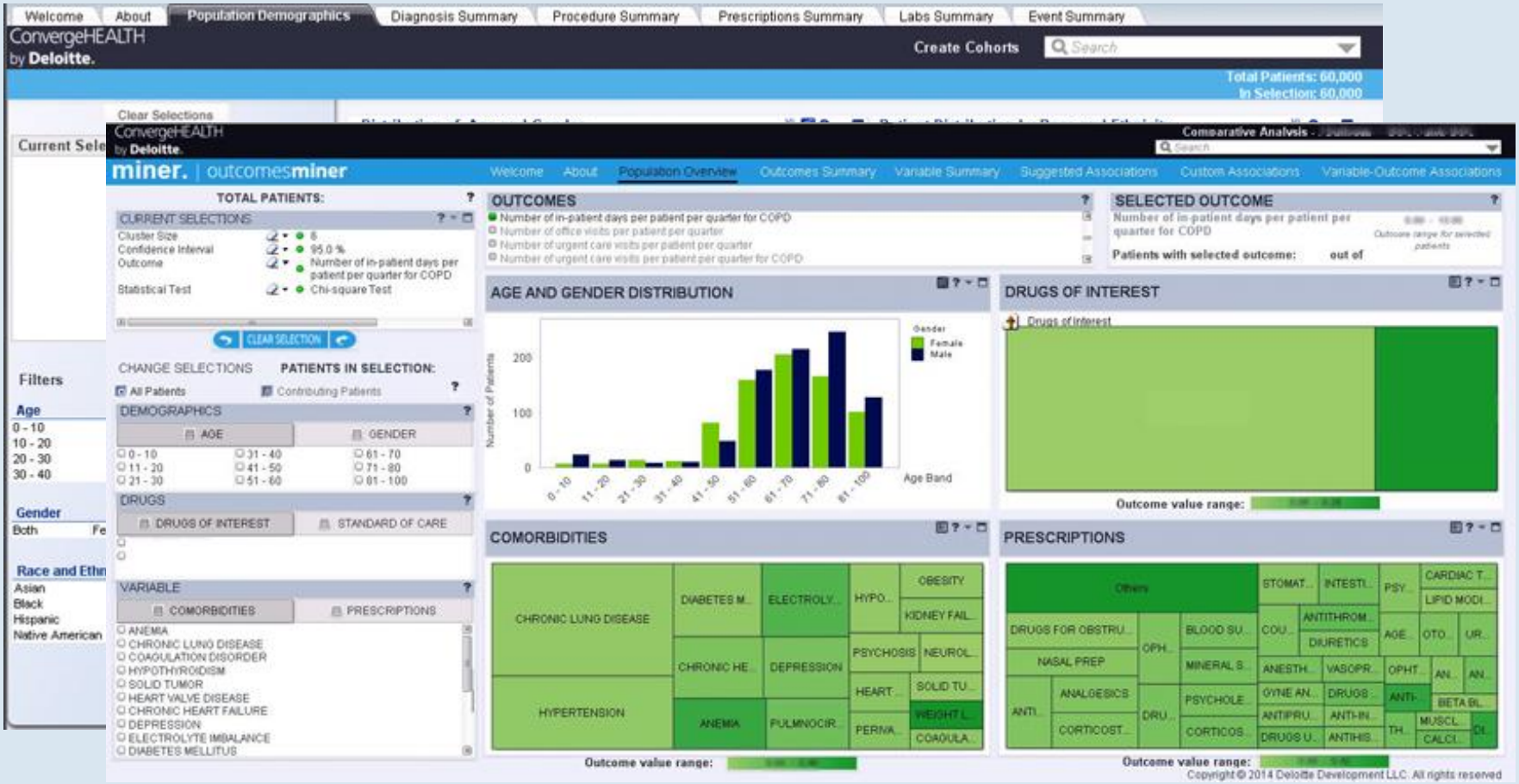
- Details:
  - I2b2-Website https://www.i2b2.org/
  - Jährliche AUG meetings Ostküste/Westküste USA, EU
  - Website IDRT http://idrt.imise.uni-leipzig.de/
  - MDR-Website https://mdr.imise.uni-leipzig.de/

Ausblick
Beobachten: TranSMART Foundation



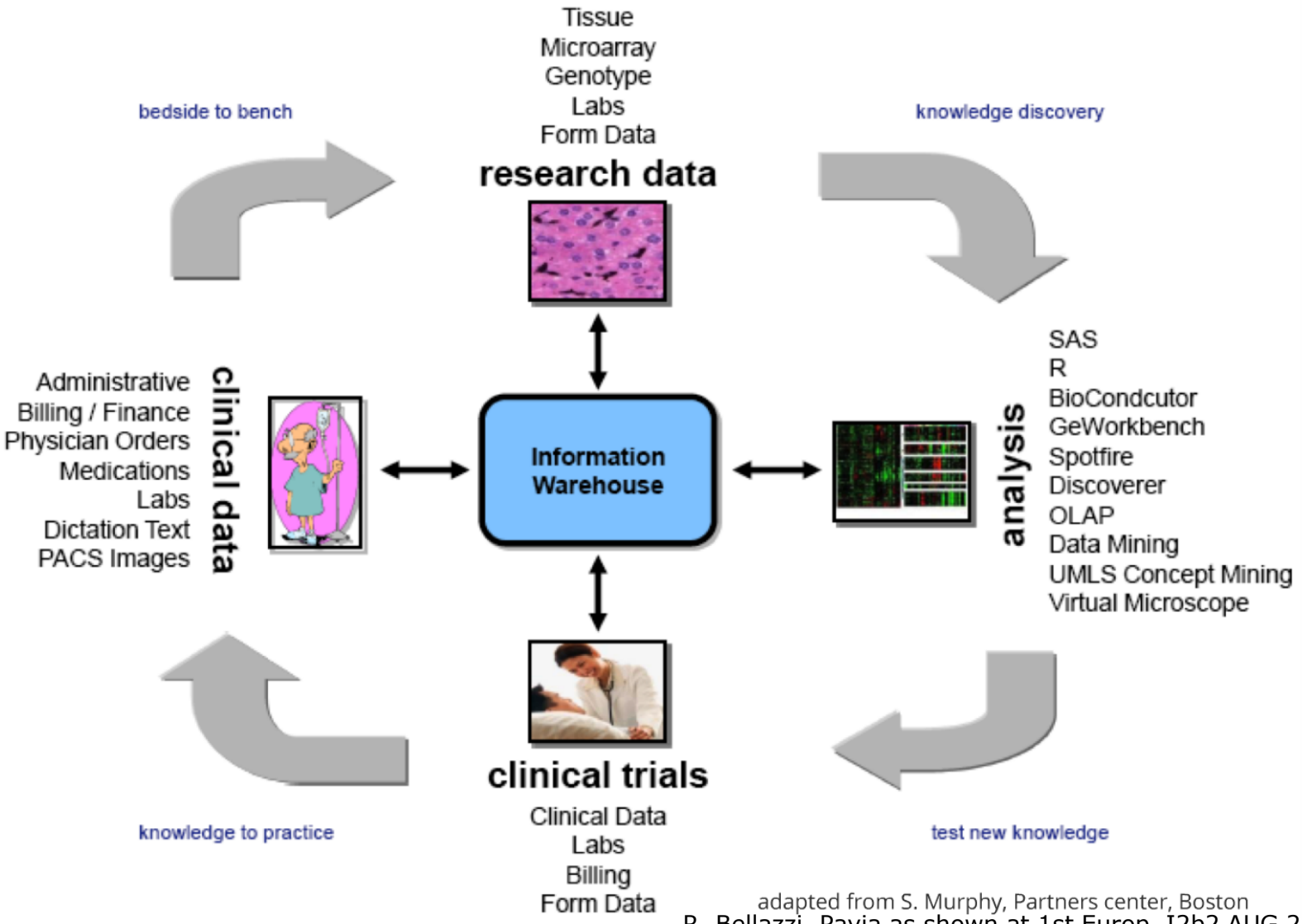
https://public.transmart.etriks.org/transmart/datasetExplorer

Ausblick
Beobachten: Deloitte Miner-Werkzeuge



http://www.converge-health.com/solutions/miner/outcomes-miner/

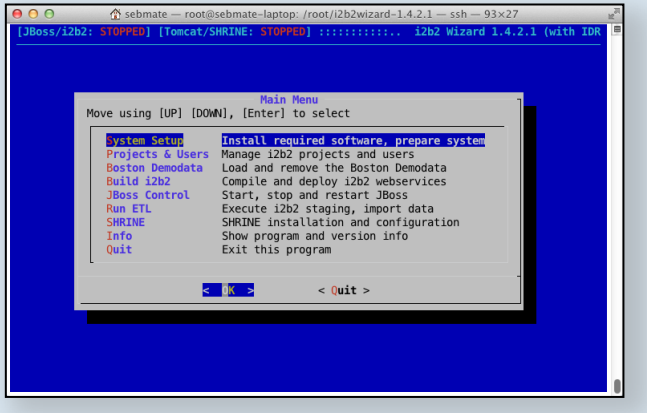
Lösungsansatz
dezentrale Forschungsinfrastruktur



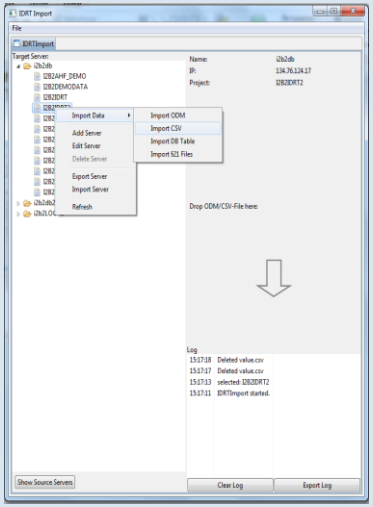
adapted from S. Murphy, Partners center, Boston
R. Bellazzi, Pavia as shown at 1st Europ. I2b2 AUG 2013

IDRT Produkte

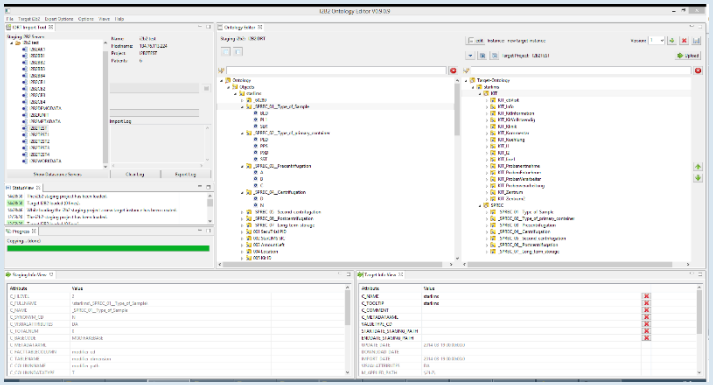
i2b2Wizard



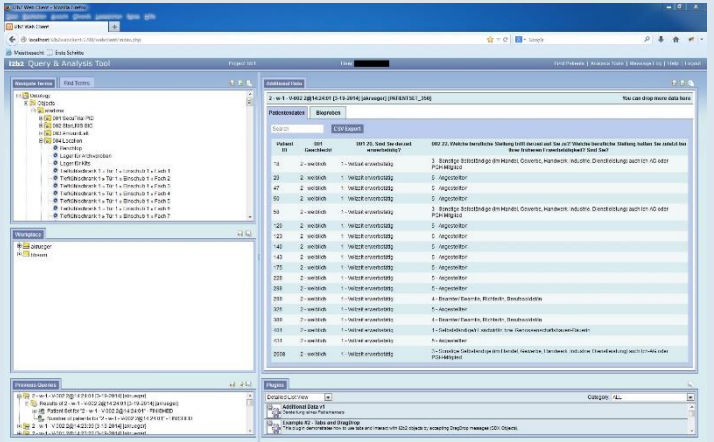
IDRT Import Tool



i2b2 Ontology Editor



i2b2 Web Client PlugIn



6. TMF-Jahreskongress 02.04.2014

Standardisierte Datenaufbereitung und Zusammenführung für die Forschung

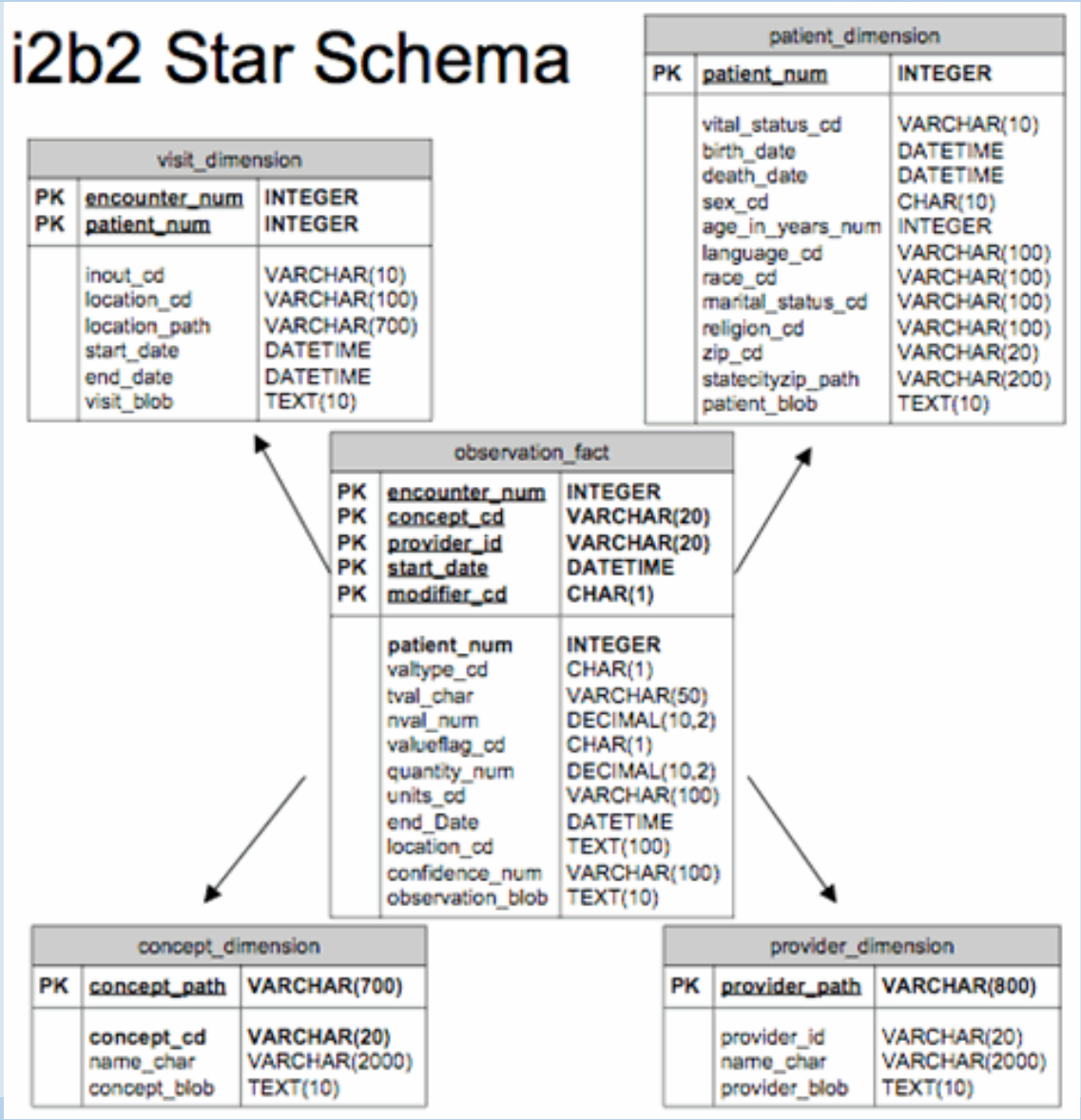
TMF-Projekt V091-MI IDRT *Integrated Data Repository Toolkit Integrated*

Ulrich Sax, Thomas Ganslandt und Matthias Löbe,
Christian Bauer, Benjamin Baum, Matthias Quade,
Matthias Löbe, Igor Engel, Sebastian Stäubert,
Sebastian Mate, Jan Christoph, Alfred Winter, Hans-Ulrich Prokosch

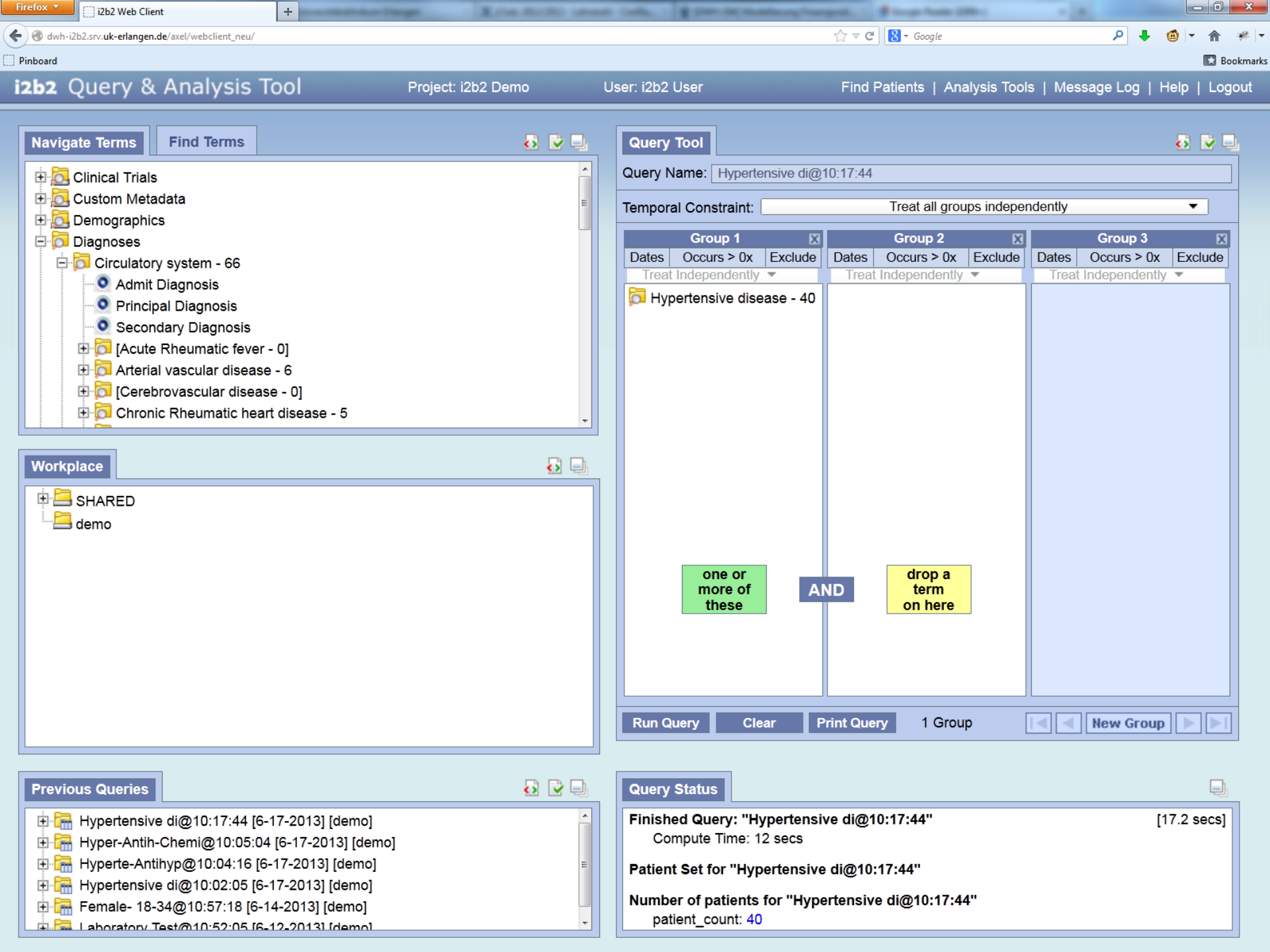
Additional Slides

Cranland T, Baumke P et al. eCF Focus Webinar "Natural 4

i2b2: Data Model



Transland I, Daumke P et al. eCF Focus Webinar "Natural 4


Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Google
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Clinical Trials
Custom Metadata
Demographics
Diagnoses
Circulatory system - 66
Admit Diagnosis
Principal Diagnosis
Secondary Diagnosis
[Acute Rheumatic fever - 0]
Arterial vascular disease - 6
[Cerebrovascular disease - 0]
Chronic Rheumatic heart disease - 5
Workplace
SHARED
demo
Query Tool
Query Name: Hypertensive di@10:17:44
Temporal Constraint: Treat all groups independently
Group 1
Group 2
Group 3
Dates
Occurs > 0x
Exclude
Treat Independently
Hypertensive disease - 40
one or more of these
AND
drop a term on here
Run Query
Clear
Print Query
1 Group
New Group
Previous Queries
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Hypertensive di@10:02:05 [6-17-2013] [demo]
Female- 18-34@10:57:18 [6-14-2013] [demo]
Query Status
Finished Query: "Hypertensive di@10:17:44"
[17.2 secs]
Compute Time: 12 secs
Patient Set for "Hypertensive di@10:17:44"
Number of patients for "Hypertensive di@10:17:44"
patient_count: 40

i2b2 Query & Analysis Tool

Navigate Terms | Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
  - Circulatory system - 66
    - Admit Diagnosis
    - Principal Diagnosis
    - Secondary Diagnosis
    - [Acute Rheumatic fever - 0]
    - Arterial vascular disease - 6
    - [Cerebrovascular disease - 0]
    - Chronic Rheumatic heart disease - 5
    - Disease of capillaries - 1
    - [Diseases of pulmonary circulation - 0]
    - Hypertensive disease - 40
    - Ischemic heart disease - 17
    - Other forms of heart disease - 35

Workplace

- SHARED
- demo

Previous Queries

- Hypertensive di@10:17:44 [6-17-2013] [demo]
  - Results of Hypertensive di@10:17:44 [6-17-2013] [demo]
    - Patient Set for "Hypertensive di@10:17:44" - FINISHED
    - Number of patients for "Hypertensive di@10:17:44" - FINISHED
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]

Demographics

Specify Data | View Results | Plugin Help

Patient Count: 40

Age in Years

10-20	3
20-30	4
30-40	5
40-50	8
50-60	4
60-70	10
70-80	3
80-90	3

Sex

F	17
M	23

Race

asian	7
black	16
hispanic	10
indian	3
white	4

Language

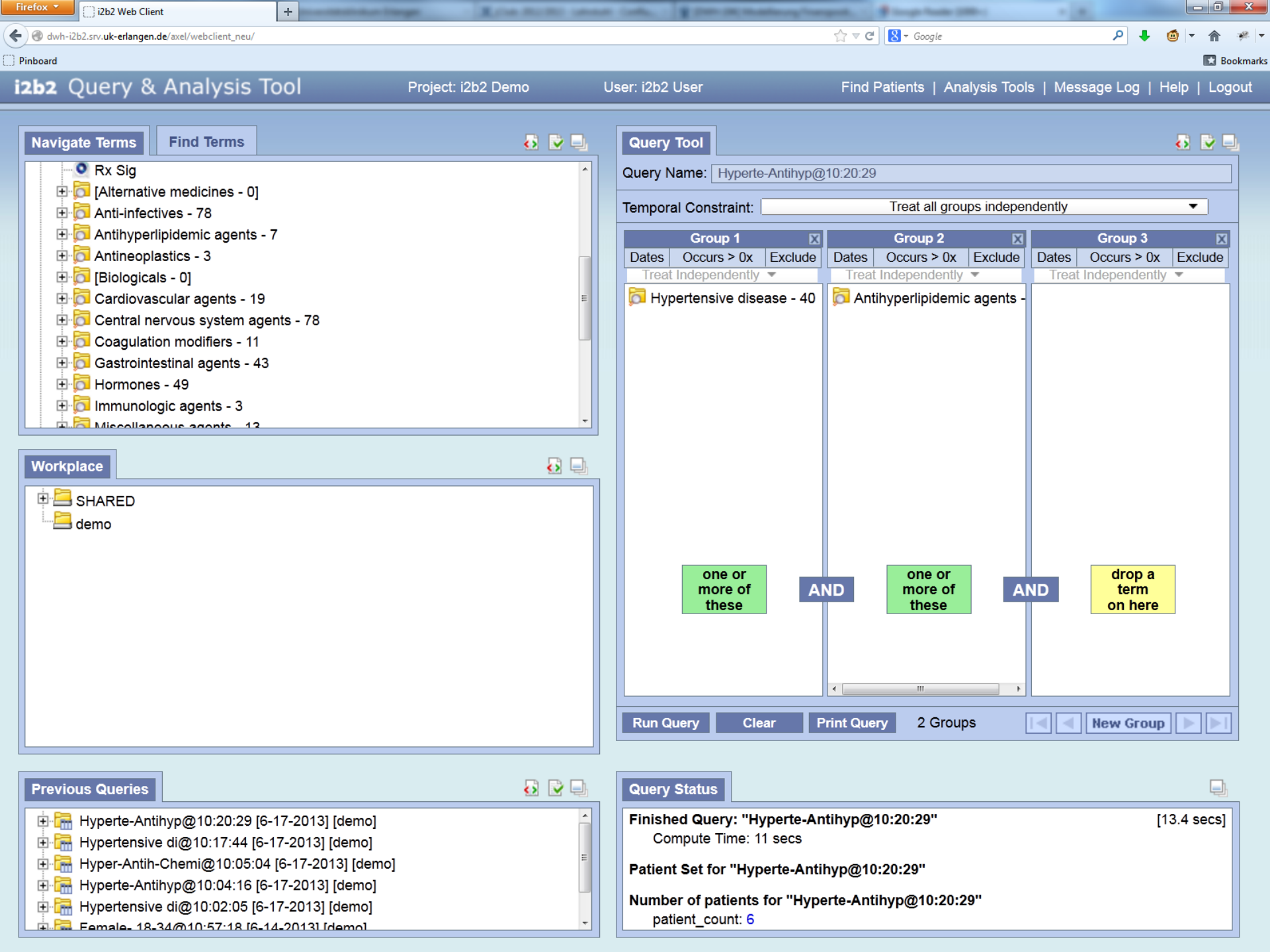
english	14
german	17
spanish	9

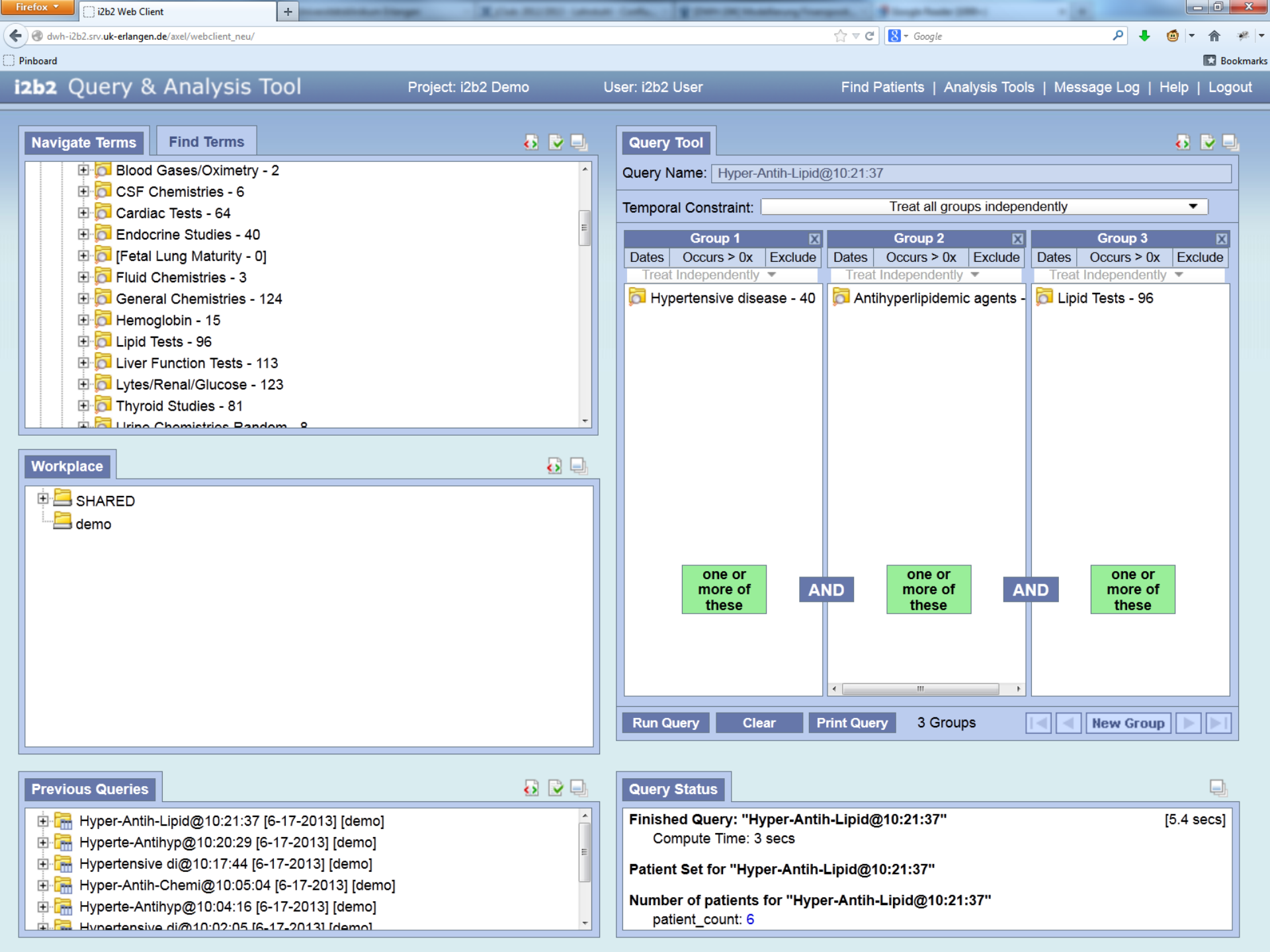
Plugins

Detailed List View Category: ALL

ExportXLS
This plugin populates a table of selectable concepts/observations and provides the possibility to download the result as a file.

Demographics (1 Patient Set) - Simple Counts
This plugin displays a demographic break-down of a Patient Set.


Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Rx Sig
[Alternative medicines - 0]
Anti-infectives - 78
Antihyperlipidemic agents - 7
Antineoplastics - 3
[Biologicals - 0]
Cardiovascular agents - 19
Central nervous system agents - 78
Coagulation modifiers - 11
Gastrointestinal agents - 43
Hormones - 49
Immunologic agents - 3
Workplace
SHARED
demo
Query Tool
Query Name: Hyperte-Antihyp@10:20:29
Temporal Constraint: Treat all groups independently
Group 1
Group 2
Group 3
Dates
Occurs > 0x
Exclude
Treat Independently
Hypertensive disease - 40
Antihyperlipidemic agents -
one or more of these
AND
one or more of these
AND
drop a term on here
Run Query
Clear
Print Query
2 Groups
New Group
Previous Queries
Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Hypertensive di@10:02:05 [6-17-2013] [demo]
Query Status
Finished Query: "Hyperte-Antihyp@10:20:29"
[13.4 secs]
Compute Time: 11 secs
Patient Set for "Hyperte-Antihyp@10:20:29"
Number of patients for "Hyperte-Antihyp@10:20:29"
patient_count: 6


Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Google
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Blood Gases/Oximetry - 2
CSF Chemistries - 6
Cardiac Tests - 64
Endocrine Studies - 40
[Fetal Lung Maturity - 0]
Fluid Chemistries - 3
General Chemistries - 124
Hemoglobin - 15
Lipid Tests - 96
Liver Function Tests - 113
Lytes/Renal/Glucose - 123
Thyroid Studies - 81
Workplace
SHARED
demo
Query Tool
Query Name: Hyper-Antih-Lipid@10:21:37
Temporal Constraint: Treat all groups independently
Group 1
Group 2
Group 3
Dates
Occurs > 0x
Exclude
Treat Independently
Hypertensive disease - 40
Antihyperlipidemic agents -
Lipid Tests - 96
one or more of these
AND
one or more of these
AND
one or more of these
Run Query
Clear
Print Query
3 Groups
New Group
Previous Queries
Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Query Status
Finished Query: "Hyper-Antih-Lipid@10:21:37"
[5.4 secs]
Compute Time: 3 secs
Patient Set for "Hyper-Antih-Lipid@10:21:37"
Number of patients for "Hyper-Antih-Lipid@10:21:37"
patient_count: 6

Navigate Terms | Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Providers
- Reports
- Visit Details

Workplace

- SHARED
- demo

Timeline

Specify Data | View Results | Plugin Help

Drop a Patient Set and one or more Concepts (Ontology Terms) into the input boxes below, and then click the "View Results" tab to view a timeline showing when those concepts were observed in the selected patient set.

Patient Set: Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo] [PATIENT

Concept(s):

- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Reports

Click a concept to remove it from the list.

Previous Queries

- Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
- Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
- Hypertensive di@10:17:44 [6-17-2013] [demo]
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
- Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
- Hypertensive di@10:02:05 [6-17-2013] [demo]

Plugins

Detailed List View Category: ALL

Timeline
This plugin creates a visual representation of when selected observations occur within a given patient set.

Project Request
This plugin is uesed to generate a request for a new project based on concepts, patients and sets..

i2b2 Query & Analysis Tool

Navigate Terms / Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Providers
- Reports
- Visit Details

Workplace

- SHARED
- demo

Timeline

Specify Data | View Results | Plugin Help

<<< start: 1 size: 10 go >>> zoom: - + pan: < >

4/3/1997 — 11/11/2003 — 6/22/2010

Person_#1000000025__Female__76yrold__Hispanic
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Person_#1000000031__Male__50yrold__Hispanic
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Person_#1000000054__Female__66yrold__Asian
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Person_#1000000083__Female__65yrold__Black
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports

Previous Queries

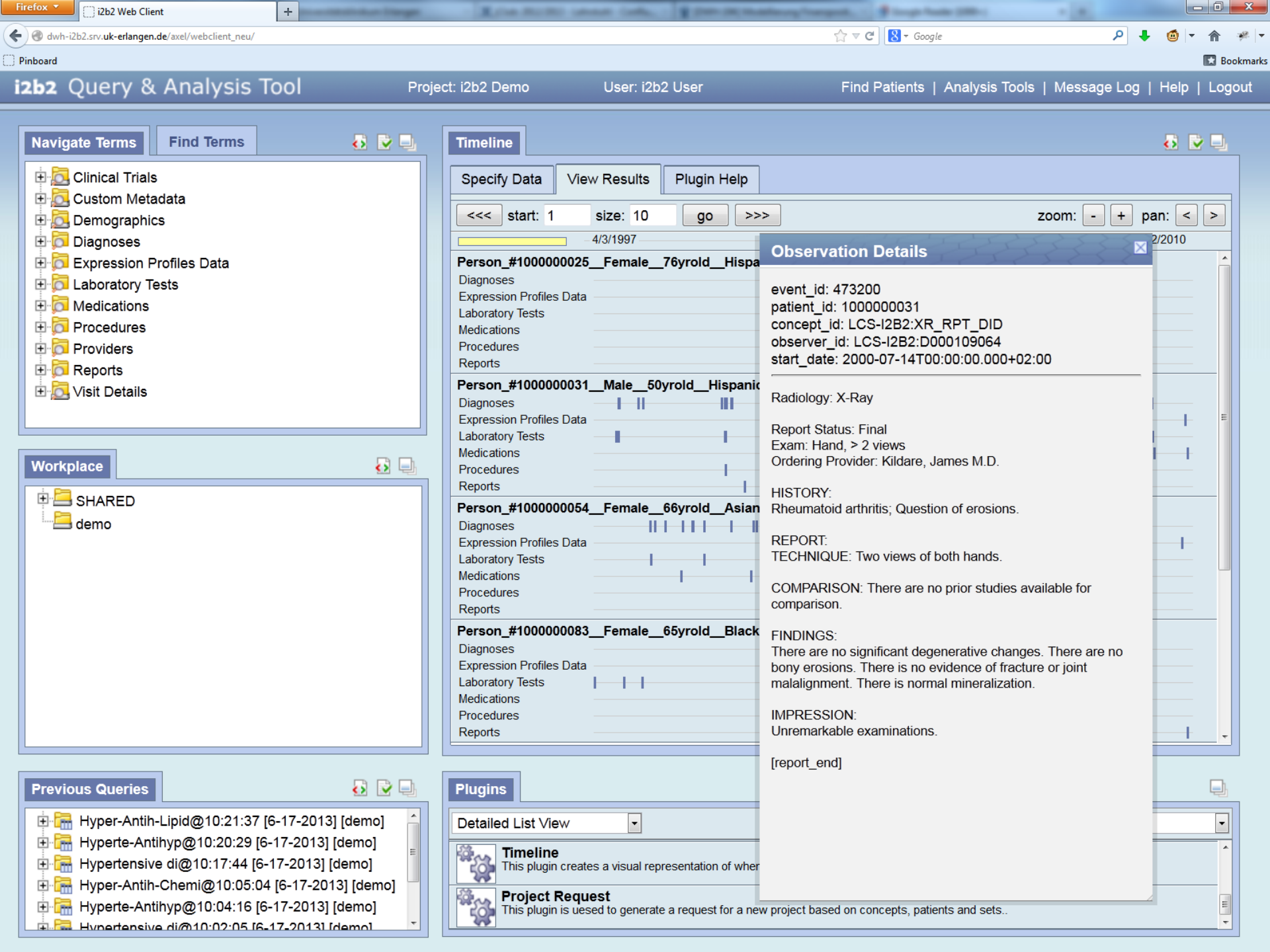
- Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
- Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
- Hypertensive di@10:17:44 [6-17-2013] [demo]
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
- Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
- Hypertensive di@10:02:05 [6-17-2013] [demo]

Plugins

Detailed List View — Category: ALL

Timeline
This plugin creates a visual representation of when selected observations occur within a given patient set.

Project Request
This plugin is uesed to generate a request for a new project based on concepts, patients and sets..


Firefox
i2b2 Web Client
dwh-i2b2.srv.uk-erlangen.de/axel/webclient_neu/
Google
Pinboard
Bookmarks
i2b2 Query & Analysis Tool
Project: i2b2 Demo
User: i2b2 User
Find Patients | Analysis Tools | Message Log | Help | Logout
Navigate Terms
Find Terms
Clinical Trials
Custom Metadata
Demographics
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Providers
Reports
Visit Details
Workplace
SHARED
demo
Timeline
Specify Data
View Results
Plugin Help
<<< start: 1 size: 10 go >>>
zoom: - + pan: < >
4/3/1997
2/2010
Person_#1000000025__Female__76yrold__Hispa
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports
Person_#1000000031__Male__50yrold__Hispanic
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports
Person_#1000000054__Female__66yrold__Asian
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports
Person_#1000000083__Female__65yrold__Black
Diagnoses
Expression Profiles Data
Laboratory Tests
Medications
Procedures
Reports
Observation Details
event_id: 473200
patient_id: 1000000031
concept_id: LCS-I2B2:XR_RPT_DID
observer_id: LCS-I2B2:D000109064
start_date: 2000-07-14T00:00:00.000+02:00
Radiology: X-Ray
Report Status: Final
Exam: Hand, > 2 views
Ordering Provider: Kildare, James M.D.
HISTORY:
Rheumatoid arthritis; Question of erosions.
REPORT:
TECHNIQUE: Two views of both hands.
COMPARISON: There are no prior studies available for comparison.
FINDINGS:
There are no significant degenerative changes. There are no bony erosions. There is no evidence of fracture or joint malalignment. There is normal mineralization.
IMPRESSION:
Unremarkable examinations.
[report_end]
Previous Queries
Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
Hypertensive di@10:17:44 [6-17-2013] [demo]
Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
Plugins
Detailed List View
Timeline
This plugin creates a visual representation of wher
Project Request
This plugin is uesed to generate a request for a new project based on concepts, patients and sets..

i2b2 Query & Analysis Tool

Project: i2b2 Demo | User: i2b2 User | Find Patients | Analysis Tools | Message Log | Help | Logout

Navigate Terms | Find Terms

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Providers
- Reports
- Visit Details

Workplace

- SHARED
- demo

Previous Queries

- Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
- Hyperte-Antihyp@10:20:29 [6-17-2013] [demo]
- Hypertensive di@10:17:44 [6-17-2013] [demo]
- Hyper-Antih-Chemi@10:05:04 [6-17-2013] [demo]
- Hyperte-Antihyp@10:04:16 [6-17-2013] [demo]
- Hypertensive di@10:02:05 [6-17-2013] [demo]

ExportXLS

Specify Data | View Results | Plugin Help | Settings

Drop a patient set and at least one concept (ontology term) onto the approriate input boxes below, then click the "View Results" tab to retrieve the respectiv observations in the selected patient set.

For more info refer to the "Plugin Help" tab.

Patient Set: Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo] [PATIENT

Concept(s):

- Diagnoses
- Laboratory Tests
- Medications
- Procedures

Click a Concept to remove it from the list.

Output Options:

Formatting: 1 row per observation (detailed, 1 column per observation detail)

- [ ] Replace Patient IDs with Ascending Numbers starting at 1

Include Patient Demographic Data:

- [x] Sex
- [x] Age
- [ ] Birth Date
- [ ] Birth Year
- [ ] Vital Status
- [ ] Language
- [ ] Race
- [ ] Religion
- [ ] Income
- [ ] State/City/ZIP

Plugins

Detailed List View | Category: ALL

ExportXLS
This plugin populates a table of selectable concepts/observations and provides the possibility to download the result as a file.

Demographics (1 Patient Set) - Simple Counts
This plugin displays a demographic break-down of a Patient Set.

Navigate Terms | **Find Terms**

- Clinical Trials
- Custom Metadata
- Demographics
- Diagnoses
- Expression Profiles Data
- Laboratory Tests
- Medications
- Procedures
- Providers
- Reports
- Visit Details

Workplace

- SHARED
- demo

Previous Queries

- Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
  - Results of Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
    - Patient Set for "Hyper-Antih-Lipid@10:21:37" - FINISH
    - Number of patients for "Hyper-Antih-Lipid@10:21:37" -
  - Results of Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]
  - Results of Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo]

ExportXLS

Specify Data | View Results | Plugin Help | Settings

Click the one of the buttons on the right to download the following table in the appropriate format.

CSV Export HTML/XLS Export

Patient Information
for Patient Set 'Hyper-Antih-Lipid@10:21:37 [6-17-2013] [demo] [PATIENTSET_14]'

	Patient ID	Sex	Age in Years	Timestamp Start	Timestamp End	Observation Set	Observation Concept Code	Observation Value	Observati Unit
1	1000000025	F	76	2001-12-05	2001-12-05	Diagnoses	ICD9:786.50		
2	1000000025	F	76	2001-12-07	2001-12-07	Diagnoses	ICD9:493.90		
3	1000000025	F	76	2001-12-07	2001-12-07	Diagnoses	ICD9:780.8		
4	1000000025	F	76	2001-12-07	2001-12-07	Diagnoses	ICD9:782.0		
5	1000000025	F	76	2001-12-07	2001-12-07	Diagnoses	ICD9:786.59		
6	1000000025	F	76	2001-12-10	2001-12-10	Diagnoses	ICD9:354.0		
7	1000000025	F	76	2001-12-10	2001-12-10	Diagnoses	ICD9:493.90		
8	1000000025	F	76	2001-12-10	2001-12-10	Diagnoses	ICD9:723.4		
9	1000000025	F	76	2001-12-10	2001-12-10	Diagnoses	ICD9:786.50		
10	1000000025	F	76	2001-12-10	2001-12-10	Diagnoses	ICD9:786.59		
11	1000000025	F	76	2001-12-11	2001-12-11	Diagnoses	ICD9:786.59		
12	1000000025	F	76	2003-07-15	2003-07-15	Diagnoses	ICD9:494		
13	1000000025	F	76	2003-07-16	2003-07-16	Diagnoses	ICD9:273.2		
14	1000000025	F	76	2003-07-16	2003-07-16	Diagnoses	ICD9:493.90		
15	1000000025	F	76	2003-12-23	2003-12-23	Diagnoses	ICD9:V16.0		

Plugins

Detailed List View Category: ALL

ExportXLS
This plugin populates a table of selectable concepts/observations and provides the possibility to download the result as a file.

Demographics (1 Patient Set) - Simple Counts
This plugin displays a demographic break-down of a Patient Set.