

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages



---

Cloud Computing für Big-Data-Analysen in der Medizin

---

Dr. Martin Sedlmayr  
Lehrstuhl für Medizinische Informatik, FAU



## Herausforderungen für Leistungserbringer im Gesundheitswesen:

- Verbesserung der Behandlungsqualität im Gesundheitswesen
- Erhöhen der Patientensicherheit
- Gesundheitsleistungen / medizinische Versorgung

Senkung der Kosten bei gleichzeitiger  
Verbesserung der Qualität



Innovationen aus Bereich Forschung und Entwicklung sollen helfen

## Sekundärnutzung klinischer Routinedaten:

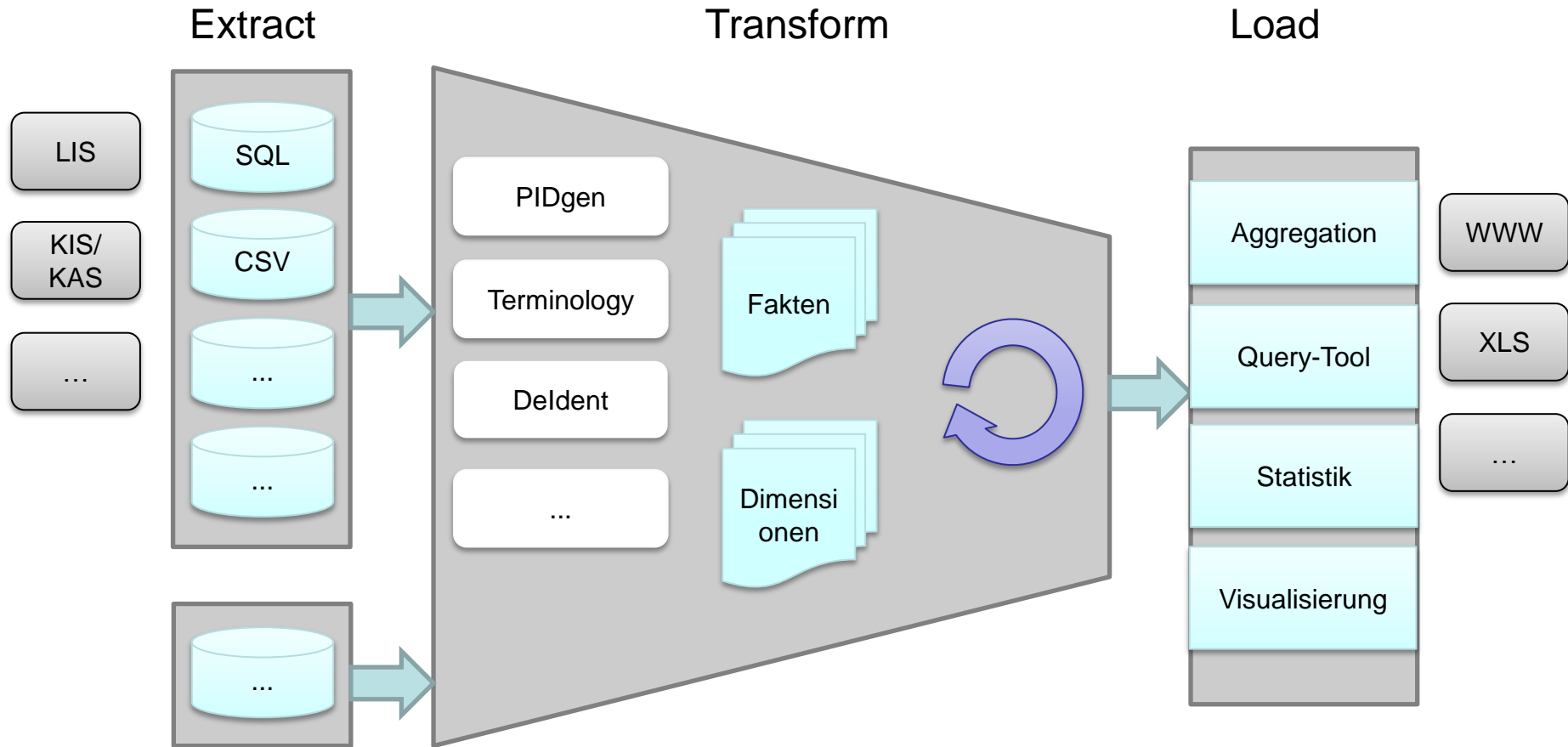
- Neudefinition von Forschung und Entwicklung im Gesundheitswesen
- Revolution der Wertschöpfungskette

## **Aggregierte Patientendaten:**

- unterstützen klinische Forschung z.B. zur Identifizierung von Krankheitsmechanismen
- ermöglichen effiziente und kostengünstige Plausibilitätsprüfung ärztlichen Handelns
- reduzieren Rekrutierungszeiten von Patienten in klinischen Studien
- unterstützen bei Verdachtsfällen die Überprüfung der Medikamentensicherheit
- ermöglichen die Überwachung der Qualität und Sicherheit medizinischer Produkte
- leisten Entbürokratisierung im Abrechnungswesen des deutschen Gesundheitssystems



lokal



# Was ist mit ... Freitext?



99.9%  
80%  
71%  
53%



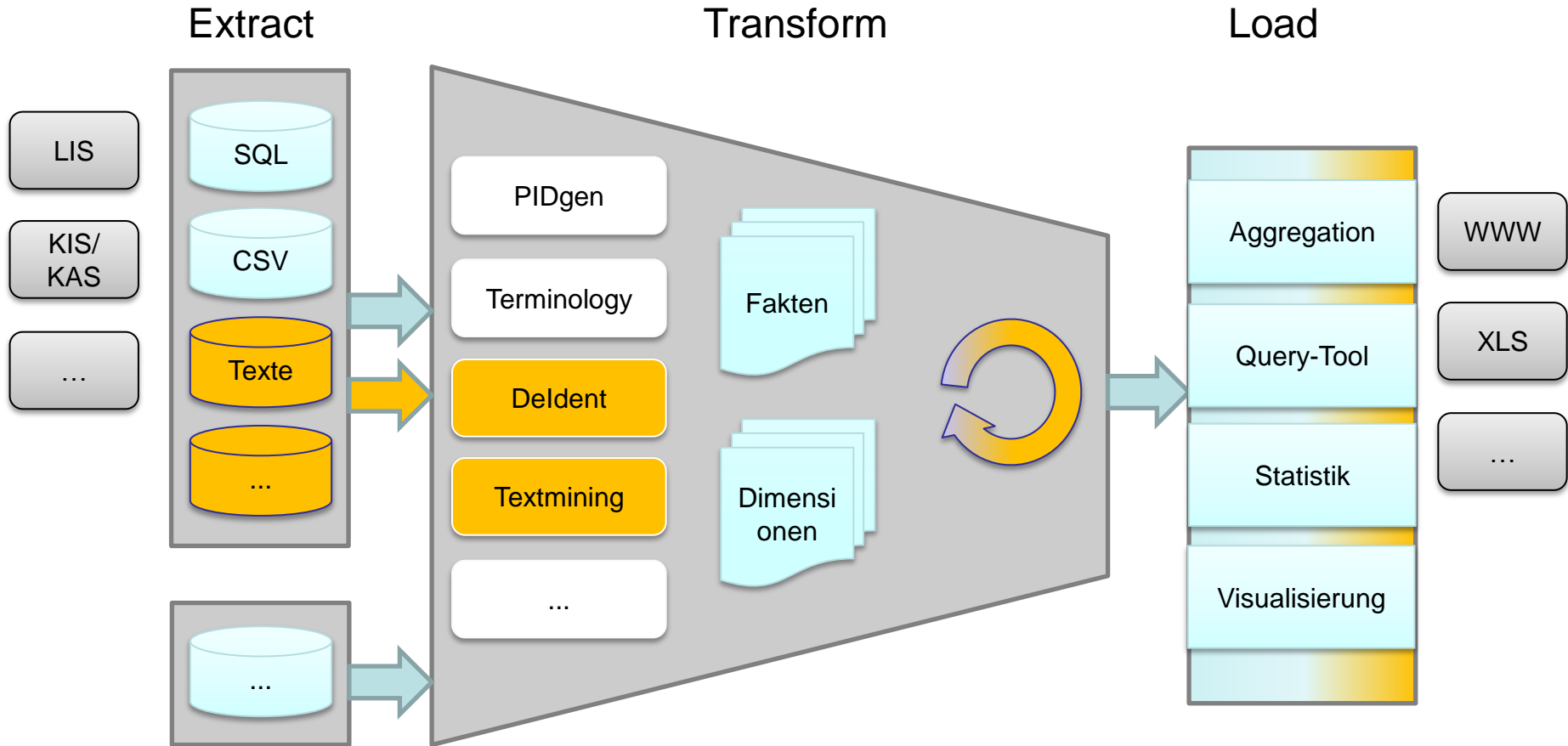


natural language  
processing

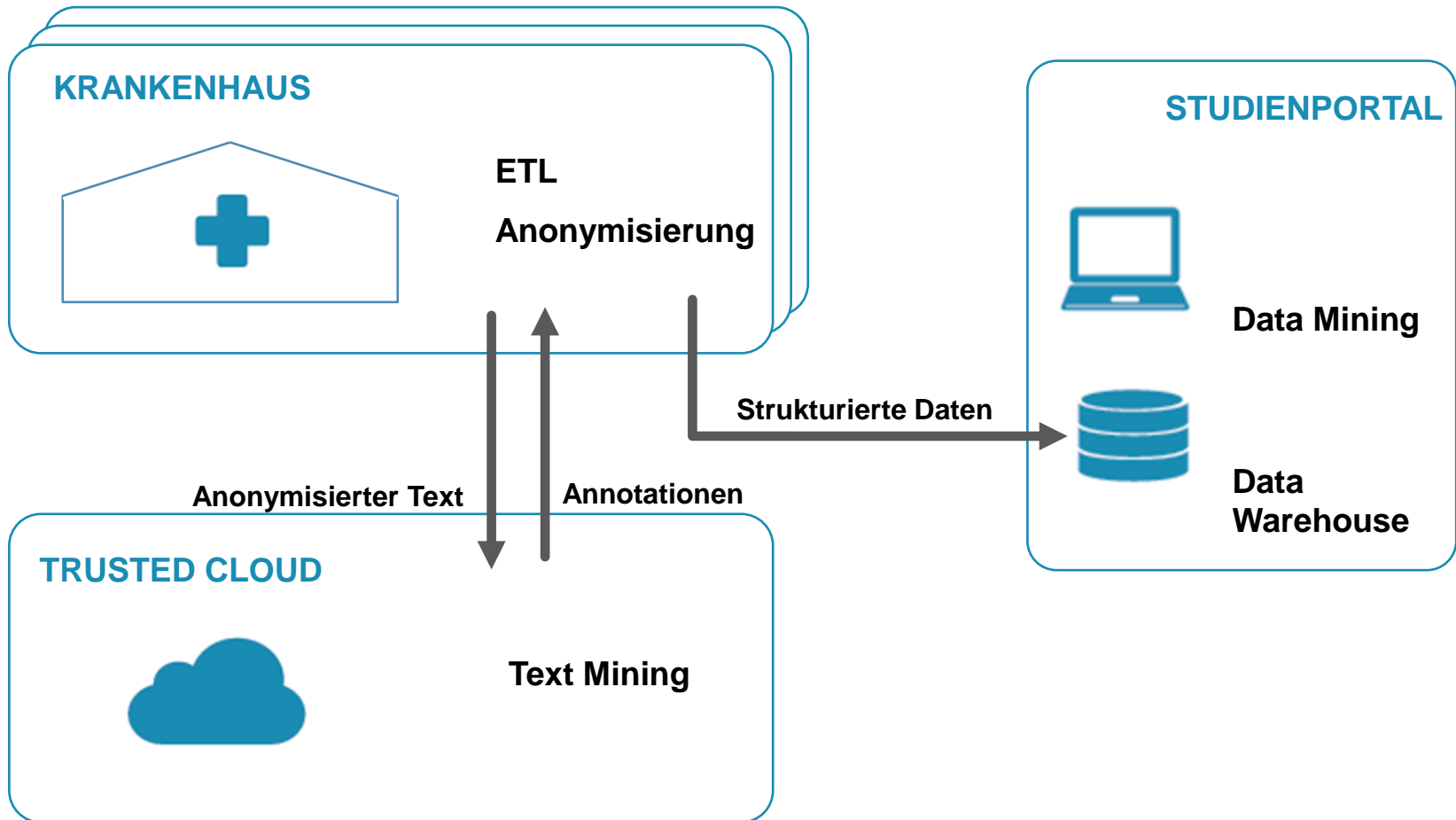


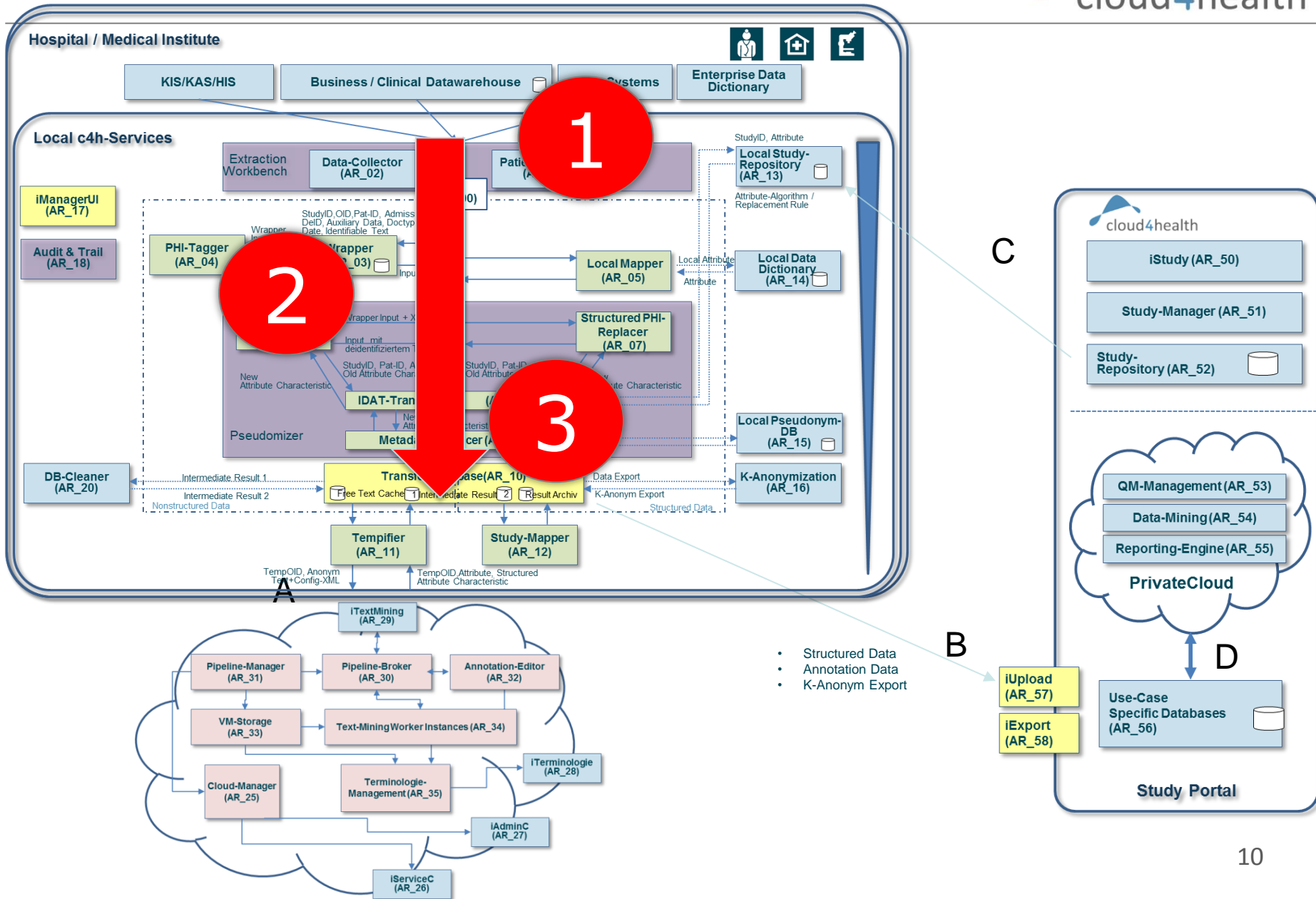
**Text Mining**  
Text Annotation

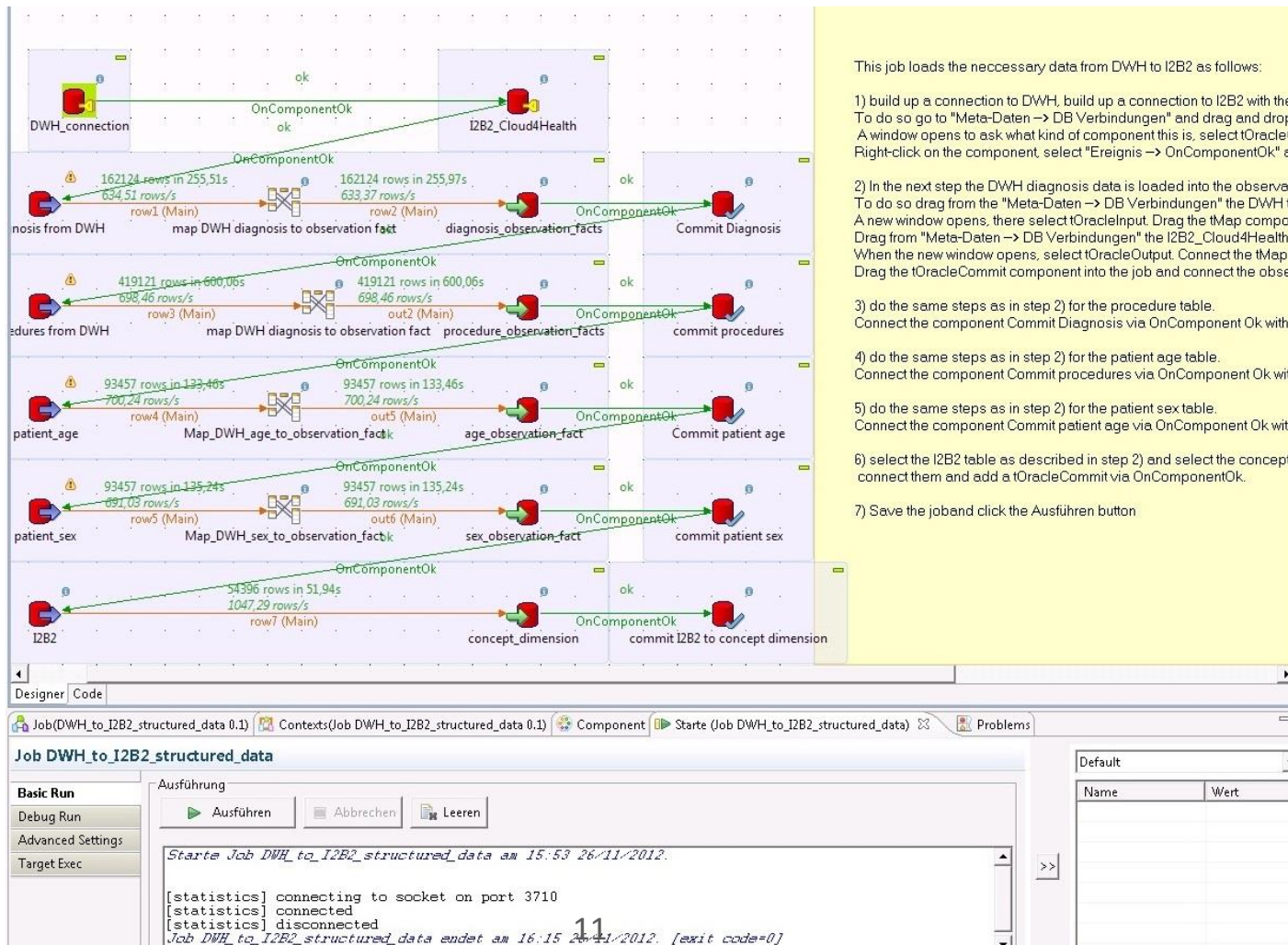
**Deidentification**











This job loads the necessary data from DWH to I2B2 as follows:

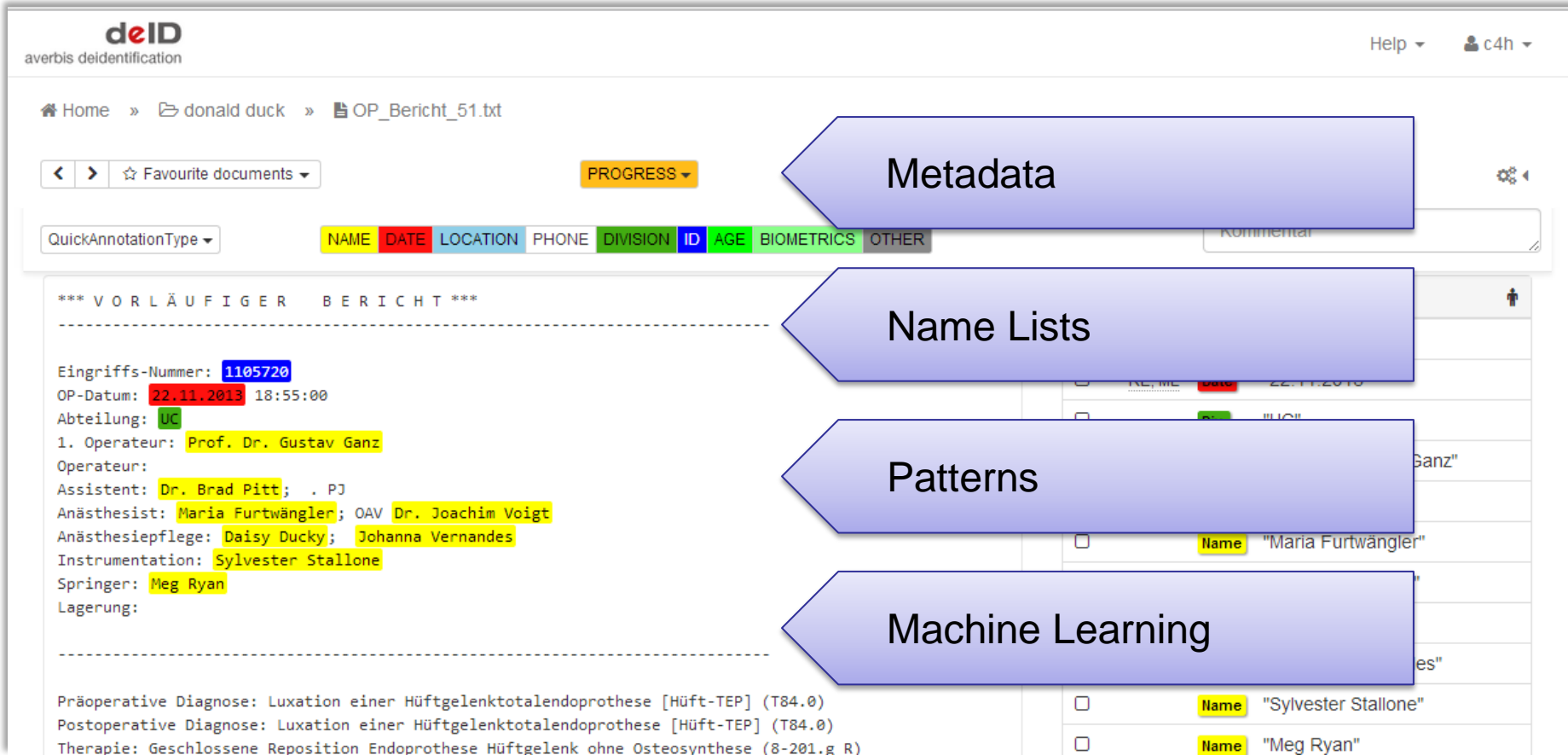
- 1) build up a connection to DWH, build up a connection to I2B2 with the To do so go to "Meta-Daten → DB Verbindungen" and drag and drop A window opens to ask what kind of component this is, select tOracleC Right-click on the component, select "Ereignis → OnComponentOk" as
- 2) In the next step the DWH diagnosis data is loaded into the observati To do so drag from the "Meta-Daten → DB Verbindungen" the DWH te A new window opens, there select tOracleInput. Drag the tMap compor Drag from "Meta-Daten → DB Verbindungen" the I2B2\_Cloud4Health ( When the new window opens, select tOracleOutput. Connect the tMap ( Drag the tOracleCommit component into the job and connect the obser
- 3) do the same steps as in step 2) for the procedure table. Connect the component Commit Diagnosis via OnComponent Ok with t
- 4) do the same steps as in step 2) for the patient age table. Connect the component Commit procedures via OnComponent Ok with
- 5) do the same steps as in step 2) for the patient sex table. Connect the component Commit patient age via OnComponent Ok with
- 6) select the I2B2 table as described in step 2) and select the concept connect them and add a tOracleCommit via OnComponentOk.
- 7) Save the job and click the Ausführen button

**Job DWH\_to\_I2B2\_structured\_data**

Ausführung

Ausführen Abbrechen Leeren

```
Starte Job DWH_to_I2B2_structured_data am 15:53 26/11/2012.  
[statistics] connecting to socket on port 3710  
[statistics] connected  
[statistics] disconnected  
Job DWH_to_I2B2_structured_data endet am 16:15 26/11/2012. [exit code=0]
```



**deID**  
averbis deidentification

Home » donald duck » OP\_Bericht\_51.txt

PROGRESS

QuickAnnotationType

NAME DATE LOCATION PHONE DIVISION ID AGE BIOMETRICS OTHER

\*\*\* VORLÄUFIGER BERICHT \*\*\*

Eingriffs-Nummer: 1105720  
OP-Datum: 22.11.2013 18:55:00  
Abteilung: UC  
1. Operateur: Prof. Dr. Gustav Ganz  
Operateur:  
Assistent: Dr. Brad Pitt; . PJ  
Anästhesist: Maria Furtwängler; OAV Dr. Joachim Voigt  
Anästhesiepflege: Daisy Ducky; Johanna Vernandes  
Instrumentation: Sylvester Stallone  
Springer: Meg Ryan  
Lagerung:

Präoperative Diagnose: Luxation einer Hüftgelenktotalendoprothese [Hüft-TEP] (T84.0)  
Postoperative Diagnose: Luxation einer Hüftgelenktotalendoprothese [Hüft-TEP] (T84.0)  
Therapie: Geschlossene Reposition Endoprothese Hüftgelenk ohne Osteosynthese (8-201.g R)

Metadata

Name Lists

Patterns

Machine Learning

## Person (entspric

- surname <string>
- familyname <string>
- affix <string>
- titel <string> (Dr., Pro
- sex [f|m] <enumeratic

## Date

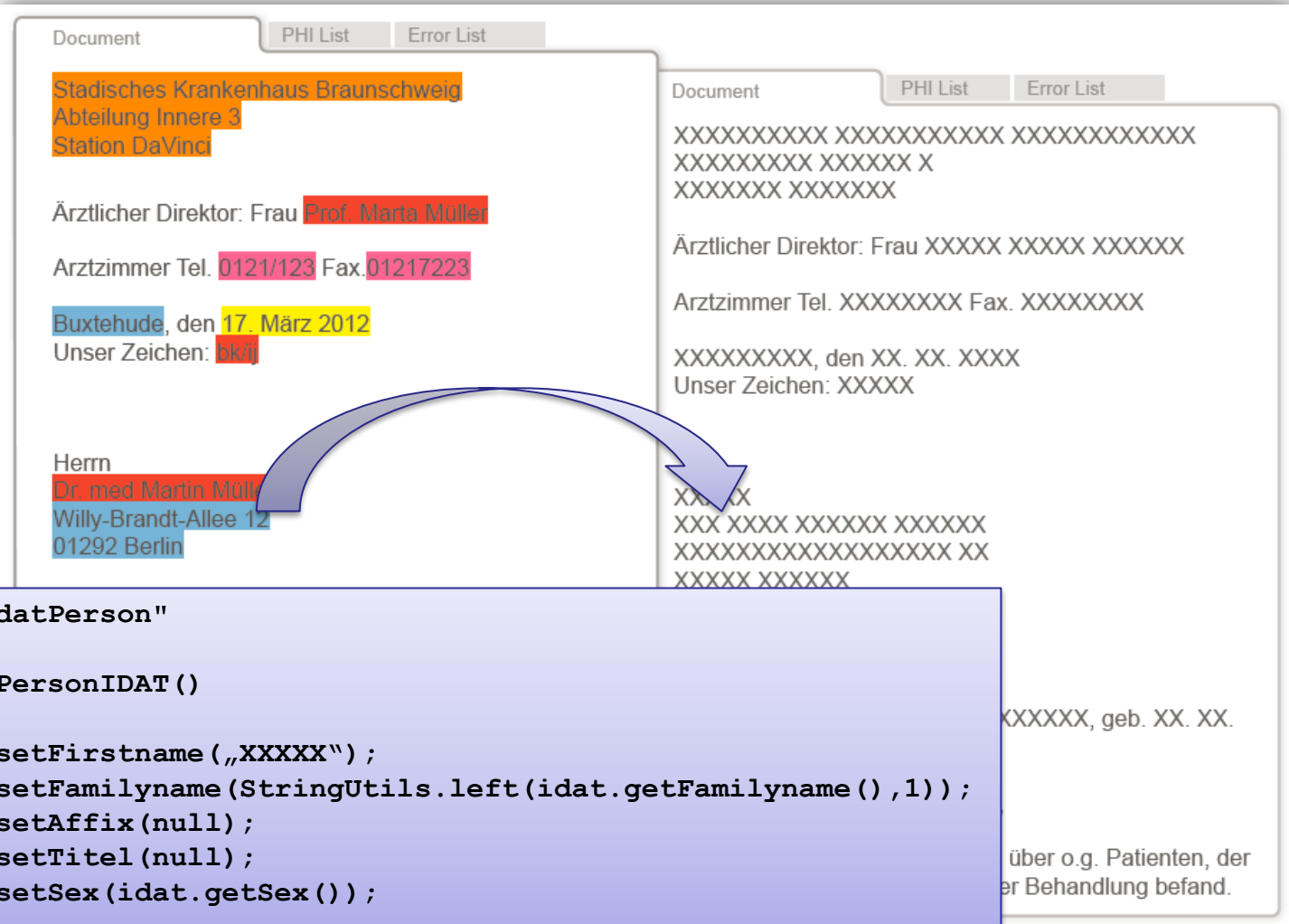
- Day <byte> 11
- Month <byte> 1..12
- Year <byte> 1921
- Weekday <byte> 1..7
- Holiday <string> (We

## Location

- street <string> (Tenn
- housenumber <string>
- city
- city
- cour
- build

## Co

- pho
- cou
- are
- pho
- er



Document | PHI List | Error List

Stadisches Krankenhaus Braunschweig  
Abteilung Innere 3  
Station DaVinci

Ärztlicher Direktor: Frau Prof. Marta Müller

Arztzimmer Tel. 0121/123 Fax. 01217223

Buxtehude, den 17. März 2012  
Unser Zeichen: bk/1

Herrn  
Dr. med Martin Müller  
Willy-Brandt-Allee 12  
01292 Berlin

Document | PHI List | Error List

XXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXXX  
XXXXXXXXXX XXXXXX X  
XXXXXXXX XXXXXXX

Ärztlicher Direktor: Frau XXXXX XXXXX XXXXXX

Arztzimmer Tel. XXXXXXXX Fax. XXXXXXXX

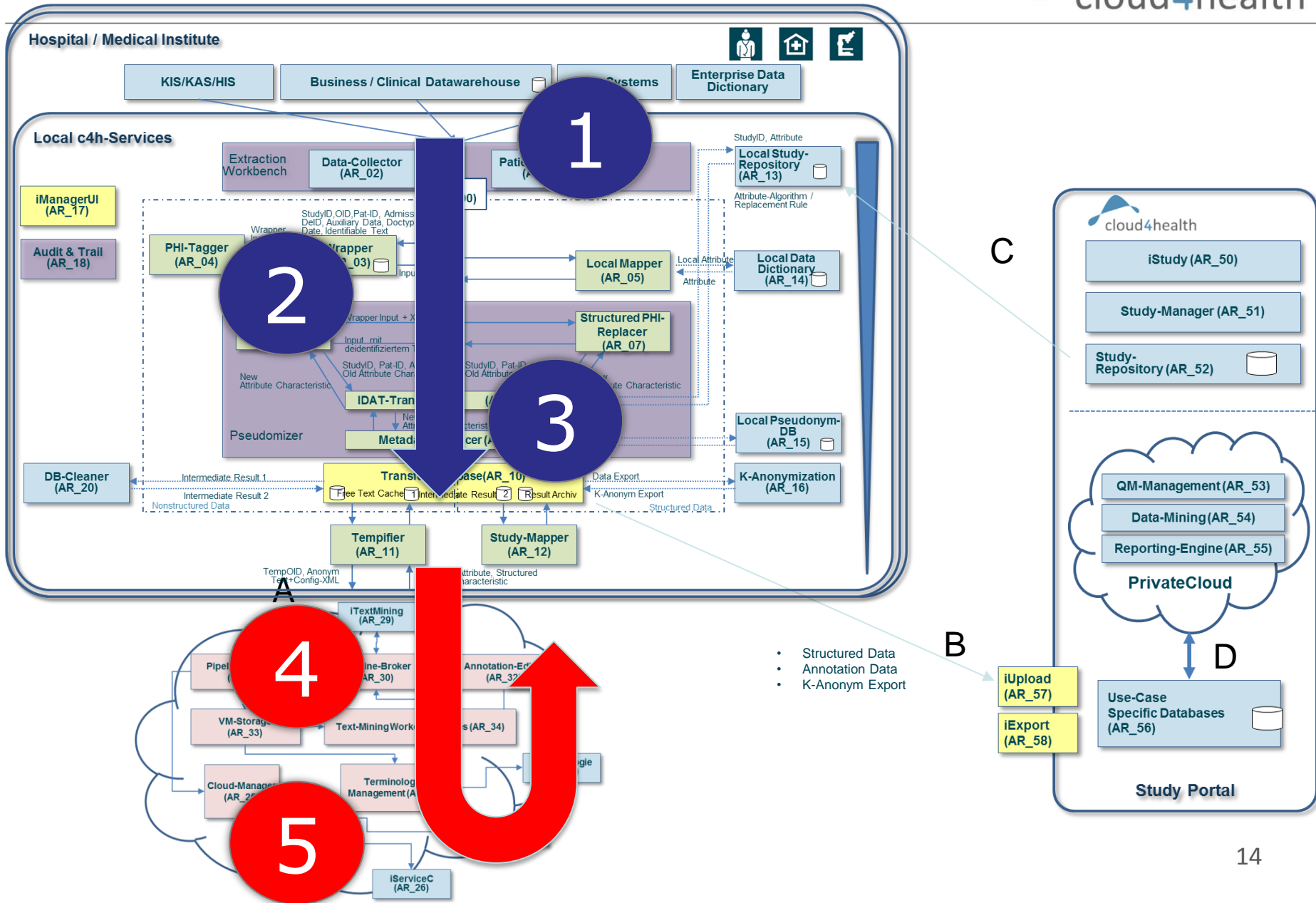
XXXXXXXXXX, den XX. XX. XXXX  
Unser Zeichen: XXXXX

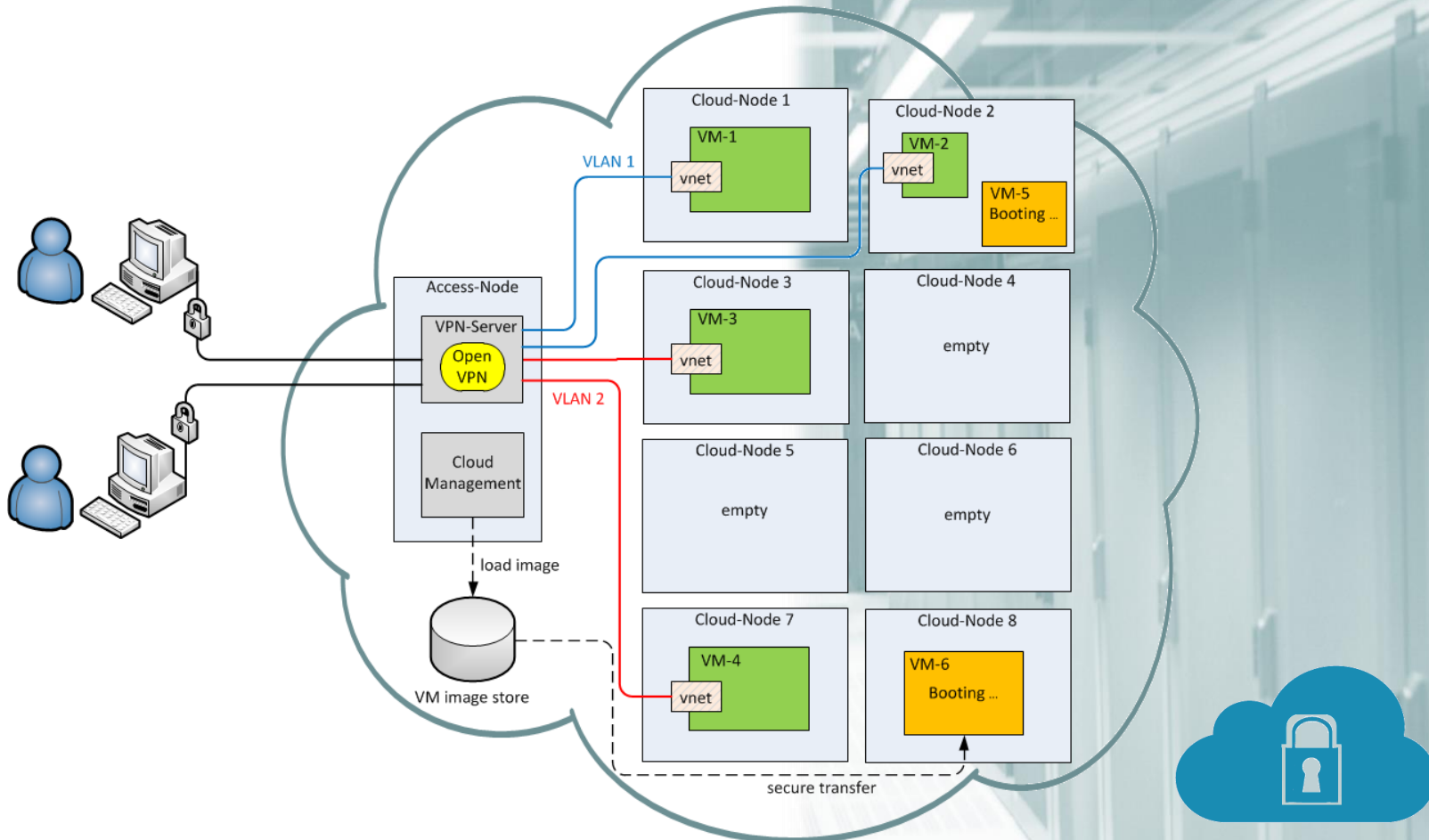
XXXXXXXX  
XXX XXXX XXXXXXX XXXXXXX  
XXXXXXXXXXXXXXXXXXXXX XX  
XXXXXX XXXXXXX

```
rule "IdatPerson"
when
  idat:PersonIDAT ()
then
  idat.setFirstname („XXXXX“);
  idat.setFamilyname (StringUtils.left(idat.getFamilyname(),1));
  idat.setAffix (null);
  idat.setTitel (null);
  idat.setSex (idat.getSex());
end
```

XXXXXX, geb. XX. XX.

über o.g. Patienten, der  
er Behandlung befand.







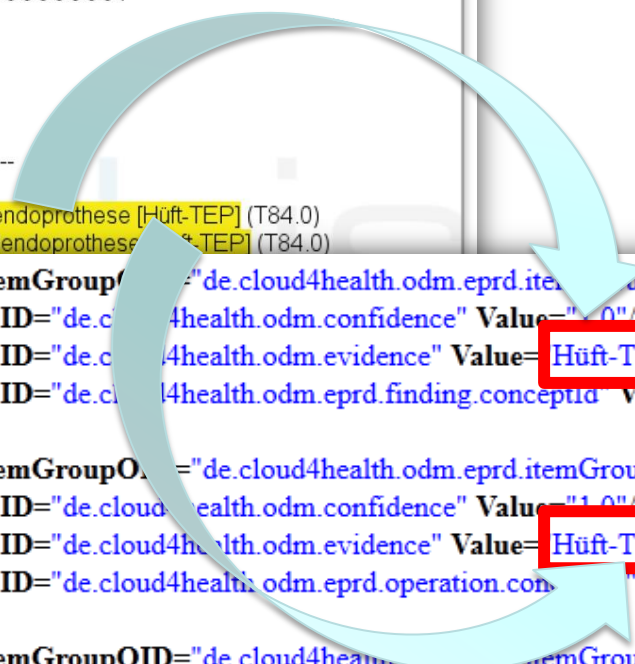
AnatomicPart  AnatomicSide  BodySide  Finding  Operation  Tense

Eingriffs-Nummer: XXXXXXXX  
OP-Datum: XXXXXXXXXXXX 18:55:00  
Abteilung: XXXXXXXXXXXXXXXXXXXX  
1. Operateur: XX;  
Assistent: XXXXXXXXXXXXXXXXXXXXXXXX; . PJ  
Anästhesist: XXXXXXXXXXXXXXXXXXXX; OAV XXXXXXXXXXXXXXXXXXXXXXXX  
Anästhesiepflege: XXXXXXXXXXXXXXXX  
Instrumentation: XXXXXXXXXXXX  
Springer: XXXXXXXXXXXXXXXX  
Lagerung:

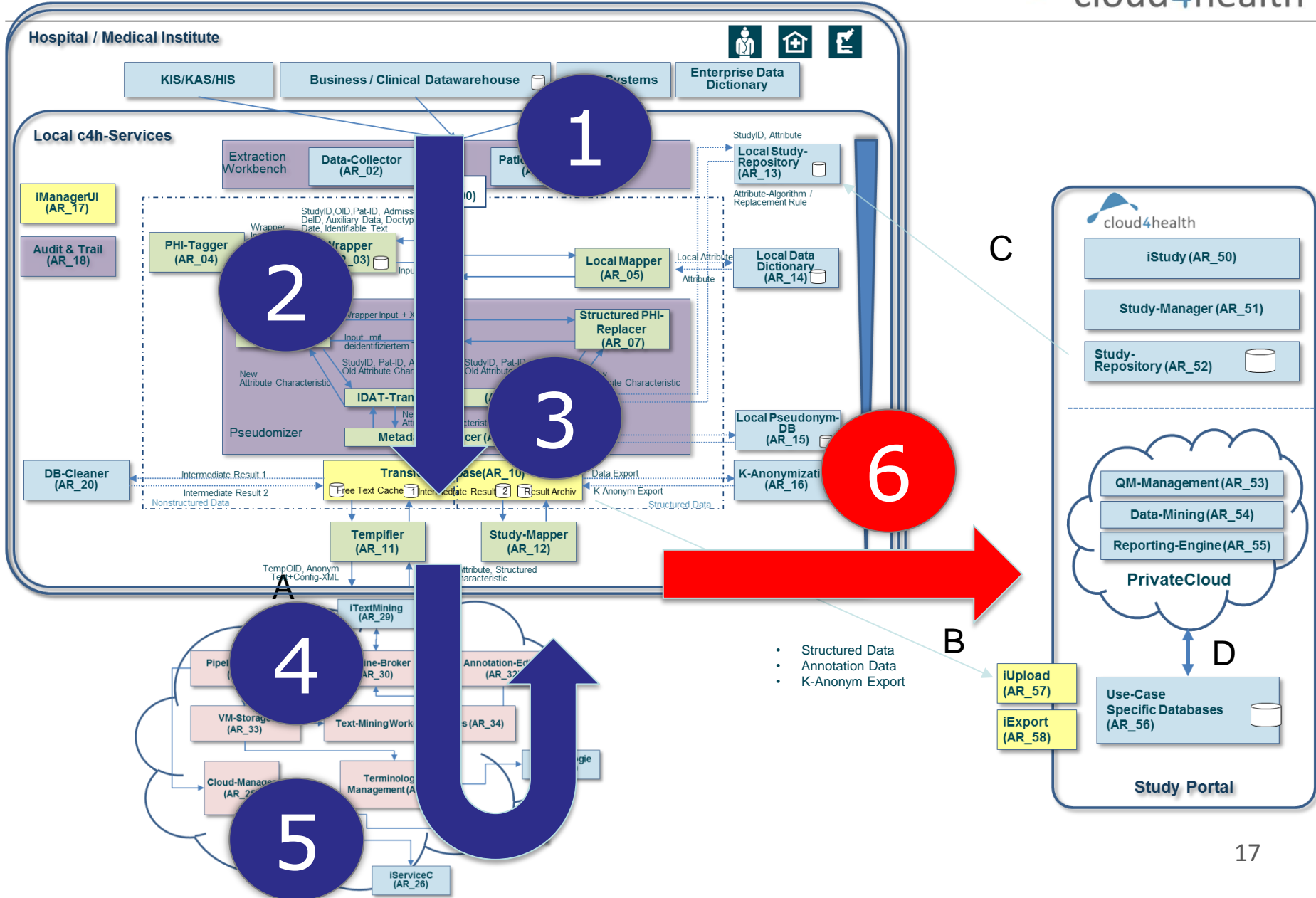
Präoperative Diagnose: Luxation einer Hüftgelenktotalendoprothese [Hüft-TEP] (T84.0)  
Postoperative Diagnose: Luxation einer Hüftgelenktotalendoprothese [Hüft-TEP] (T84.0)  
Therapie: Geschlossene Hüft-TEP  
Freitext Diagnose: Hüft-TEP  
Freitext Therapie: geschlossene Hüft-TEP  
Vorgeschichte: Der Verletzte wurde durch ein Verdrrehen des Beines ein Hüftgelenk luxiert.  
Indikation zur Geschlossenen Hüft-TEP

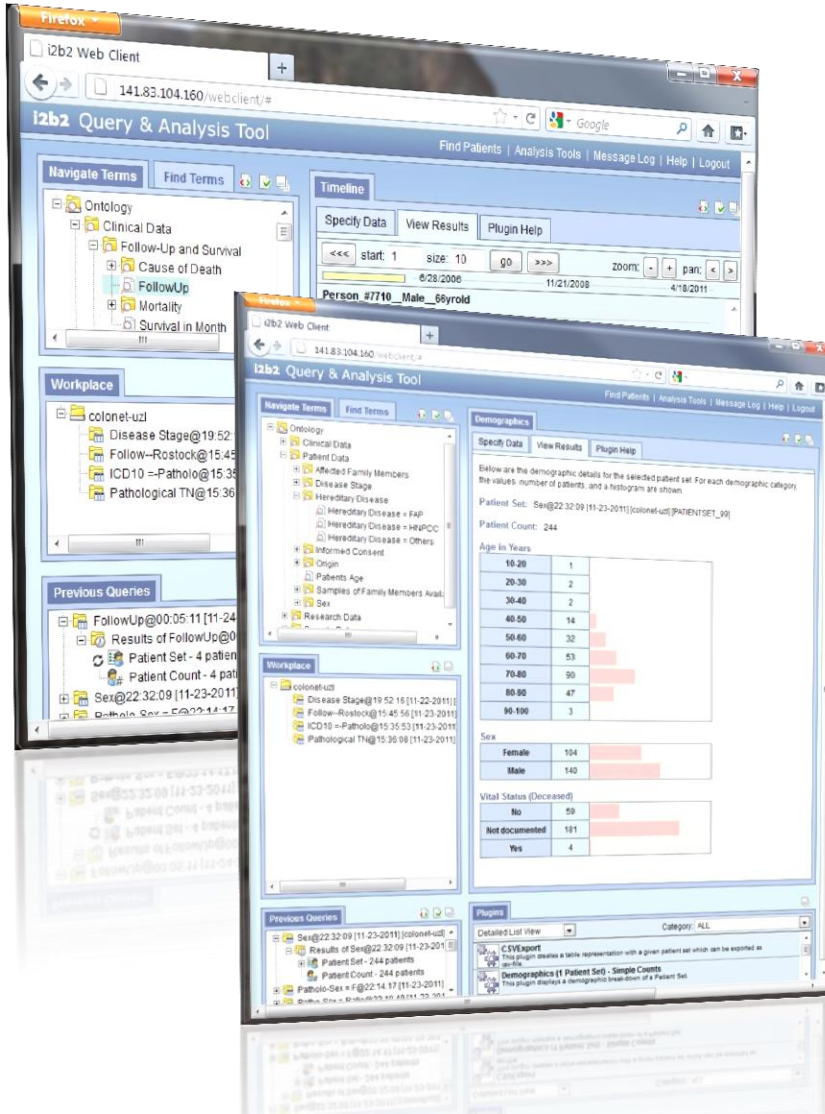
1. Bericht: Larynxmaskenintubation  
Unter Durchleuchtung erfolgte die Einlagerung des Hüftkopfes auf die Pfanne sowie das korrekte Eintreten des Hüftkopfes in die Pfanne.

```
- <ItemGroupData ItemGroupOID="de.cloud4health.odm.eprd.itemgroup.finding" ItemGroupRepeatKey="11">
  <ItemData ItemOID="de.cloud4health.odm.confidence" Value="1.0"/>
  <ItemData ItemOID="de.cloud4health.odm.evidence" Value="Hüft-TEP Luxation"/>
  <ItemData ItemOID="de.cloud4health.odm.eprd.finding.conceptid" Value="c4n.eprd.finding.luxation"/>
</ItemGroupData>
- <ItemGroupData ItemGroupOID="de.cloud4health.odm.eprd.itemgroup.operation" ItemGroupRepeatKey="12">
  <ItemData ItemOID="de.cloud4health.odm.confidence" Value="1.0"/>
  <ItemData ItemOID="de.cloud4health.odm.evidence" Value="Hüft-TEP"/>
  <ItemData ItemOID="de.cloud4health.odm.eprd.operation.conceptid" Value="5-820.0"/>
</ItemGroupData>
- <ItemGroupData ItemGroupOID="de.cloud4health.odm.eprd.itemgroup.operation" ItemGroupRepeatKey="10">
  <ItemData ItemOID="de.cloud4health.odm.confidence" Value="1.0"/>
  <ItemData ItemOID="de.cloud4health.odm.evidence" Value="Osteosynthese"/>
  <ItemData ItemOID="de.cloud4health.odm.eprd.operation.conceptid" Value="NA"/>
</ItemGroupData>
```

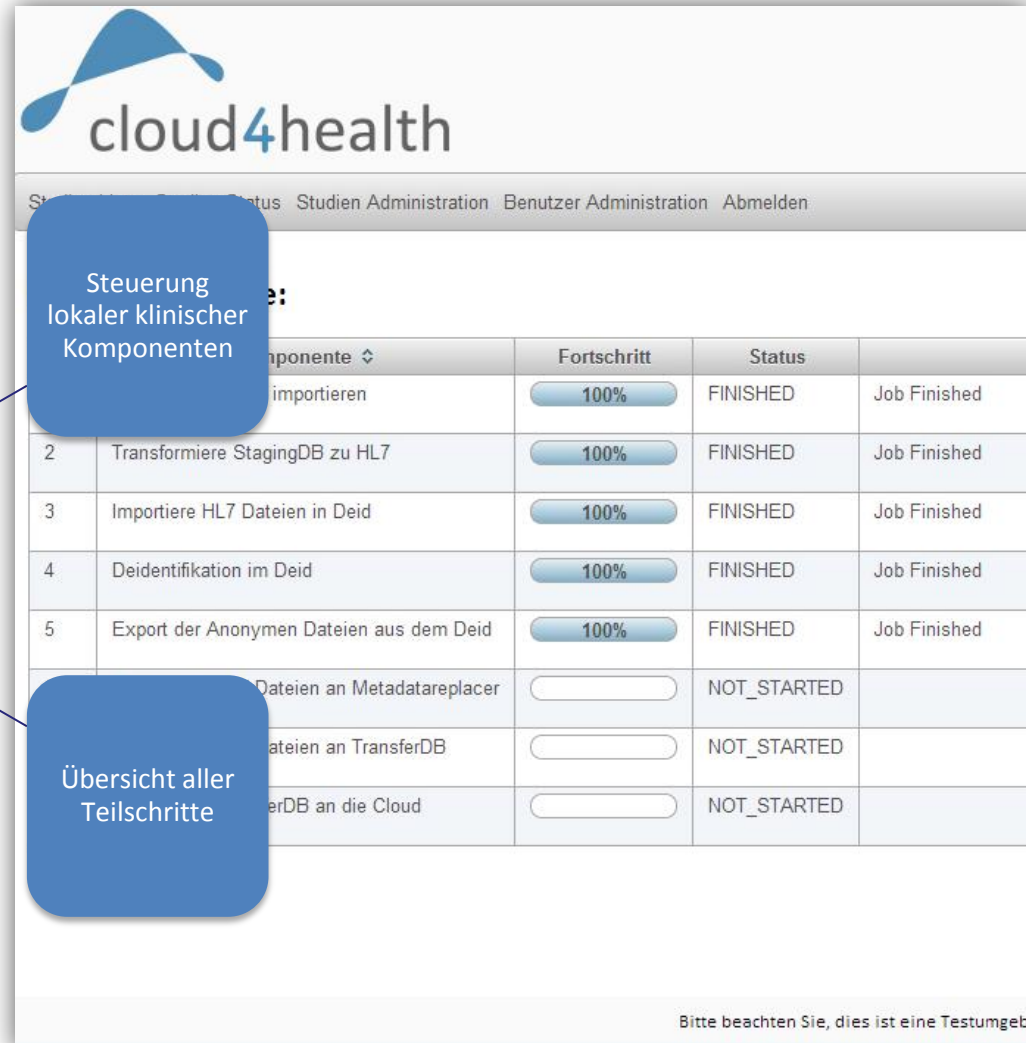
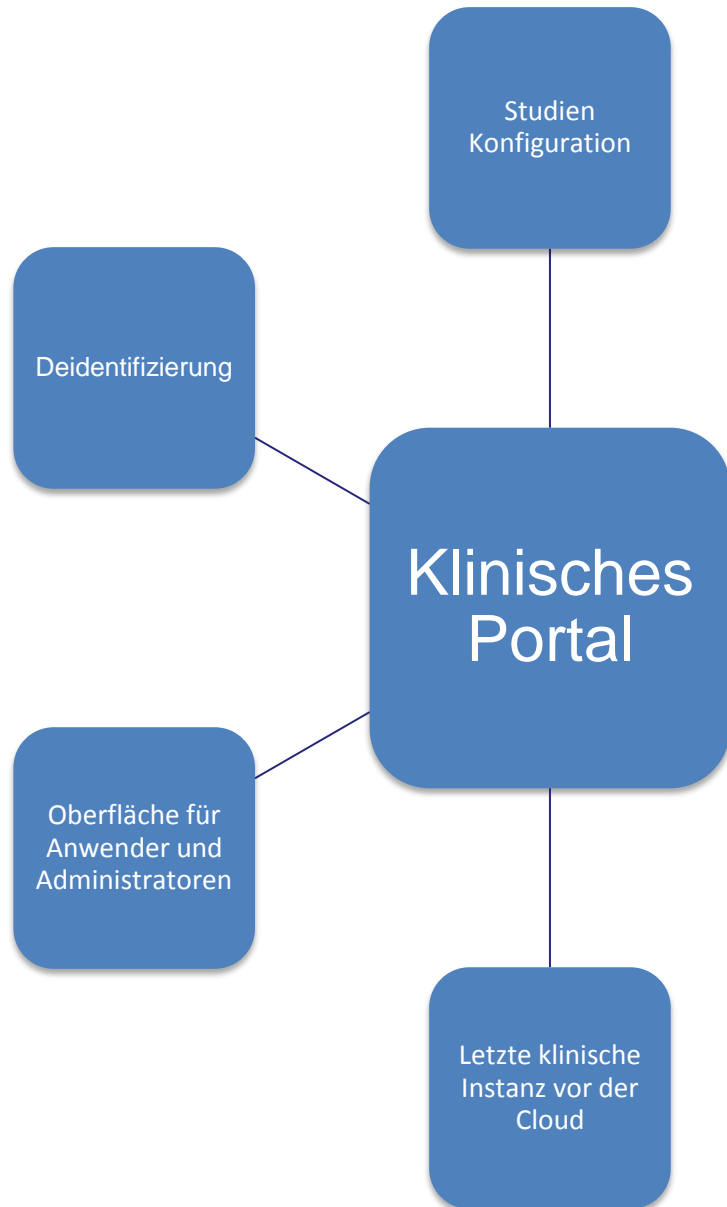








- Zugriff auf Daten
- Mehrwertdienste
  - Statistische Analyse
  - Datamining
- i2b2, R, tranSMART,  
...



The screenshot shows the cloud4health interface with a navigation bar and a table of clinical components.

Navigation: Studien Administration Benutzer Administration Abmelden

komponente	Fortschritt	Status	
Importieren	100%	FINISHED	Job Finished
2 Transformiere StagingDB zu HL7	100%	FINISHED	Job Finished
3 Importiere HL7 Dateien in Deid	100%	FINISHED	Job Finished
4 Deidentifikation im Deid	100%	FINISHED	Job Finished
5 Export der Anonymen Dateien aus dem Deid	100%	FINISHED	Job Finished
Exportieren der Dateien an Metadatatereplac	<input type="text"/>	NOT_STARTED	
Exportieren der Dateien an TransferDB	<input type="text"/>	NOT_STARTED	
Exportieren der Dateien an die Cloud	<input type="text"/>	NOT_STARTED	

Bitte beachten Sie, dies ist eine Testumgebung

Steuerung lokaler klinischer Komponenten

Übersicht aller Teilschritte

## Angelehnt an den Vorgaben des Deutschen Endoprothesenregisters

Bundesweites Register für Endoprothesen zur Ermöglichung von Qualitätskontrollen von künstlichen Hüft und Knieprothesen



## Der Cloud4health Workflow kann folgende Daten extrahieren:

- Patientenbezogene Daten
- Gelenk (Knie/Hüfte), Seite (rechts/links)
- Art: (Erst-OP, Wechsel),
- Voroperationen am endoprothetisch versorgten Gelenk
  - Können großen Einfluss auf Standzeit haben
- Standzeiten bei Versorgung im gleichen Haus
- Gründe für Wechsel von Prothesenkomponenten
- Nebendiagnosen
- Komplikationen

## Datenmigration von unstrukturierten textuellen medizinischen Berichten zu hochstrukturierter Information in klinischen Datenbanksystemen

### Der Cloud4health Workflow kann folgende Daten bereitstellen:

- Patientenbezogene Daten
- Stadieneinteilung nach TNM Klassifikation  
Erkennung der Terme und Interpretation der Information  
Bsp: ypT2 ypN1a (2/4) pM1(HEP) L1 V0 Pn0
- Anzahl der Nachbefunde
- Tumorgraduierung



Große Datenmengen mit mehr als 500.000 Pathologieberichten aus der 1. Klinik konnten automatisch in hoher Qualität strukturiert werden

Finale Test für die Strukturierung weiterer Bestände laufen an einer 2. Klinik

## Überprüfung von Plausibilität von Verordnungen am Beispiel des Projektes P3 – Personalisierte Pharmakotherapie

### Der Cloud4health Workflow kann folgende Daten bereitstellen:

- Diagnosen
- Verordnungen: Arzneimitteldaten
- Laborwerte aus strukturierten sowie unstrukturierten Daten

Anhand der extrahierten Parameter kann die Richtigkeit der Therapieempfehlung entsprechend der eingesetzten Guidelines automatisch überprüft werden

Annotation Results for DOC001.docx.txt.xml in C:\Users\vhle\projekte\cloud4health\phoch3\brieffe\_xmi

Medikation bei Entlassung (unverändert)

Duloxetine (z.B. Cymbalta) 60-30-0-0 mg  
 Zolpidem (z.B. Stilnox) 0-0-0-15 mg  
 Prednisolon (z.B. Decortin) 2,5-0-2,5-0 mg  
 Hydrochlorthiazid + Valsartan (z.B. Co Diovan 80/12,5) 1-0-0-0-Tbl.  
 Risperidon (z.B. Risperdal®) 0-0-1-0 mg  
 Pantoprazol (z.B. Pantozol) 40-0-0-0 mg  
 Indacaterol (z.B. Onbrez®) 300µg/d  
 Tiotropiumbromid (z.B. Spiriva®) 18µg/d  
 Brimonidin (Alphagan® Augentropfen) 2gtt/d  
 Natriumcarboxymethylcellulose (Optive® Augentropfen) 2gtt/d  
 Latanoprost (z.B. Xalatan® Augentropfen) 2gtt/d

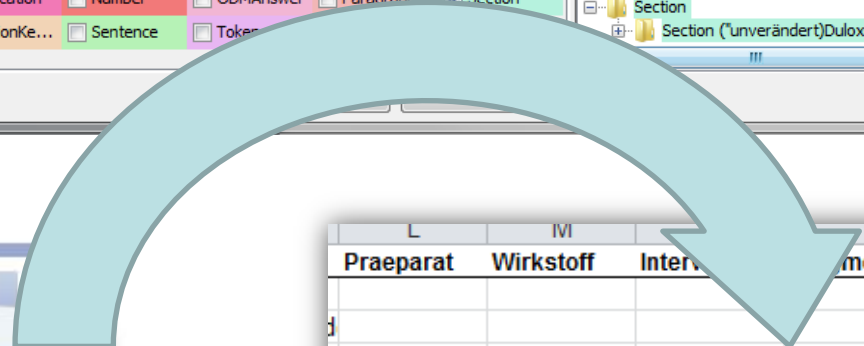
Legend

- Concept
- Diagnose
- Document...
- Drug
- Entity
- Header
- Labor
- LaborUnit
- LaborWert
- Line
- Medication
- Number
- ODMAnswer
- Paragraph
- Section
- SectionKe...
- Sentence
- Token

Click In Text to See Annotation Detail

Annotations

- Medication
  - Medication ("Tiotropiumbromid (z.B. Spiriva®) 18µg/d")
    - begin = 1122
    - end = 1164
    - componentId = null
    - confidence = 0.0
    - drug = null
    - praeparat = Praeparat ("(z.B. Spiriva®)")
    - wirkstoff = Wirkstoff ("Tiotropiumbromid")
      - dose = 18
      - unit = µg
      - regimen = 18 /d
      - form = null
      - route = null
      - trade\_name = (z.B. Spiriva®)
      - active\_ingredient = Drug ("Tiotropiumbromid")
- Section
  - Section ("unverändert")Duloxetine (z.B. Cymbalta) 60-3



L	MI	U	P	Q	
Praeparat	Wirkstoff	Inter...	MonitoringID	Dosisschema	Dosiseinheit
Cymbalta	Duloxetine			60-30-0-0	mg
Stilnox	Zolpidem			0-0-0-15	mg
Decortin	Prednisolon			2,5-0-2,5-0	mg
Co Diovan 80	Hydrochlorthiazid + Valsartan			1-0-0-0	Tbl
Risperdal®	Risperidon			0-0-1-0	mg
Pantozol	Pantoprazol			40-0-0-0	mg
Onbrez®	Indacaterol			300 /d	µg
Spiriva®	Tiotropiumbromid			18 /d	µg
Alphagan® A	Brimonidin			2 /d	gtt
Optive® Aug	Natriumcarboxymethylcellulose			2 /d	gtt
z.B. Xalatan®	Latanoprost			2 /d	gtt

## Use Case Description

Fragestellung  
Einschlusskriterien  
Notwendige Daten zur Beantwortung

## Identification of Data Sources

Klinische  
Quellsysteme  
Schnittstellen,  
Formate, Qualität  
...  
Eigentümer und  
Schutzbedarf

## Allowance

Szenario  
Eigentümer  
Datenschützer  
Ggfs.  
Einverständnis  
des Patienten

## Data Extraction

Technische  
Realisierung  
Syntaktisch &  
semantisch



- Sekundärnutzung
- Strukturierte und **unstrukturierte** Daten
  - Deidentifizierung
  - Text Mining
- Cloud computing
  - Dynamische Infrastruktur
  - Infrastructure / Service on Demand
  - Externe und Inhouse-Nutzung
- Anwendungsbeispiele
  - Füllen von Registern, Medizinische Forschung, technischer Betrieb und Qualitätsmanagement



natural language  
processing

