

Eine Referenzarchitektur für semantische Interoperabilität und ihre praktische Anwendung



Christian Zunner

**Thomas Ganslandt, Hans-Ulrich Prokosch,
Thomas Bürkle**

Lehrstuhl für Medizinische Informatik
Am Wetterkreuz 13, 91058 Erlangen

**Friedrich-Alexander-Universität
Erlangen-Nürnberg**



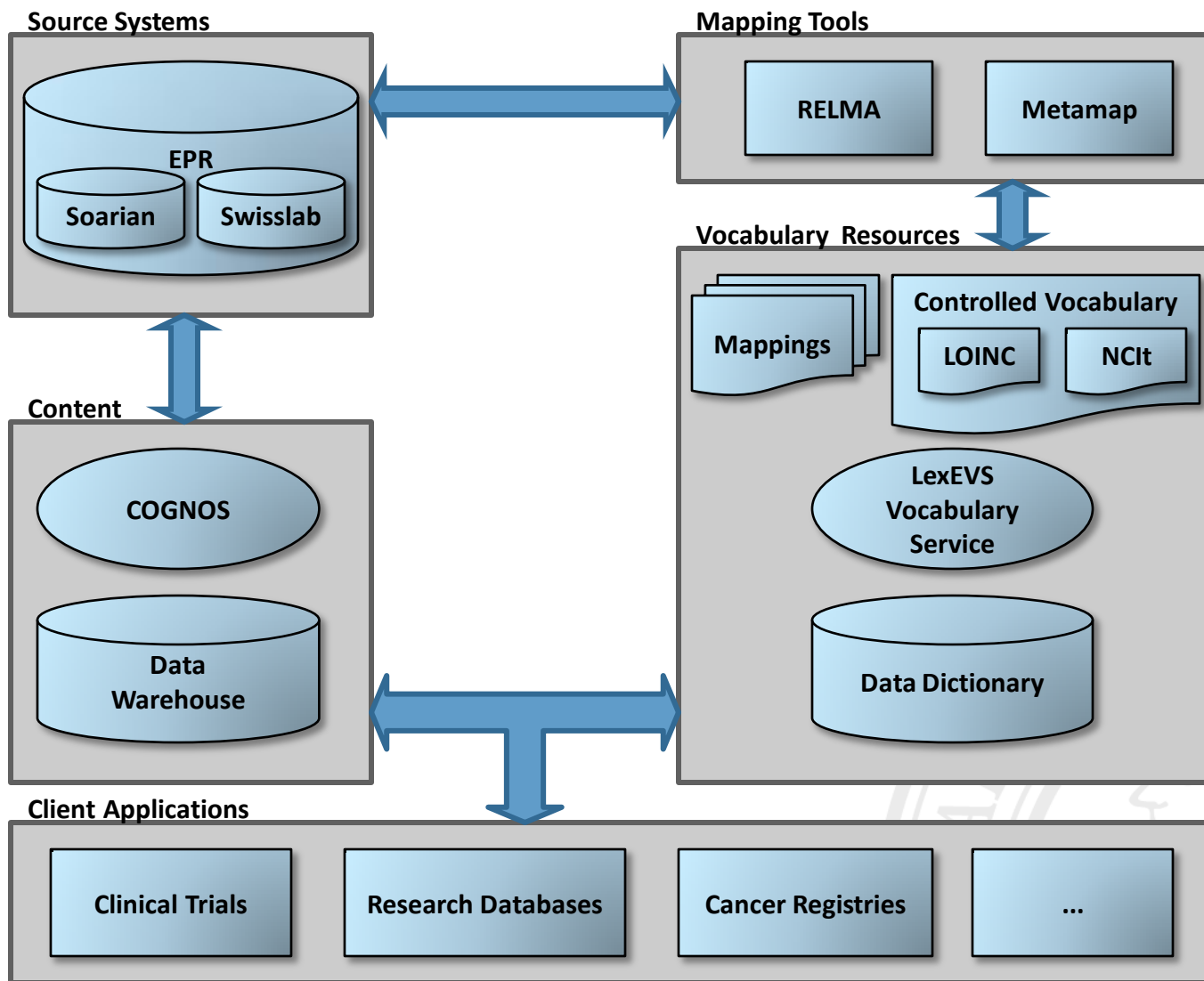
Einleitung: Semantische Interoperabilität der Ansatz

- Sekundärnutzung klinischer Daten erfordert semantische Interoperabilität
- Standards existieren, müssen in Single-Source-Umgebungen aber auch implementiert werden
- **Ziel:**
generische Single-Source-Architektur
aus frei verfügbaren Tools und Komponenten



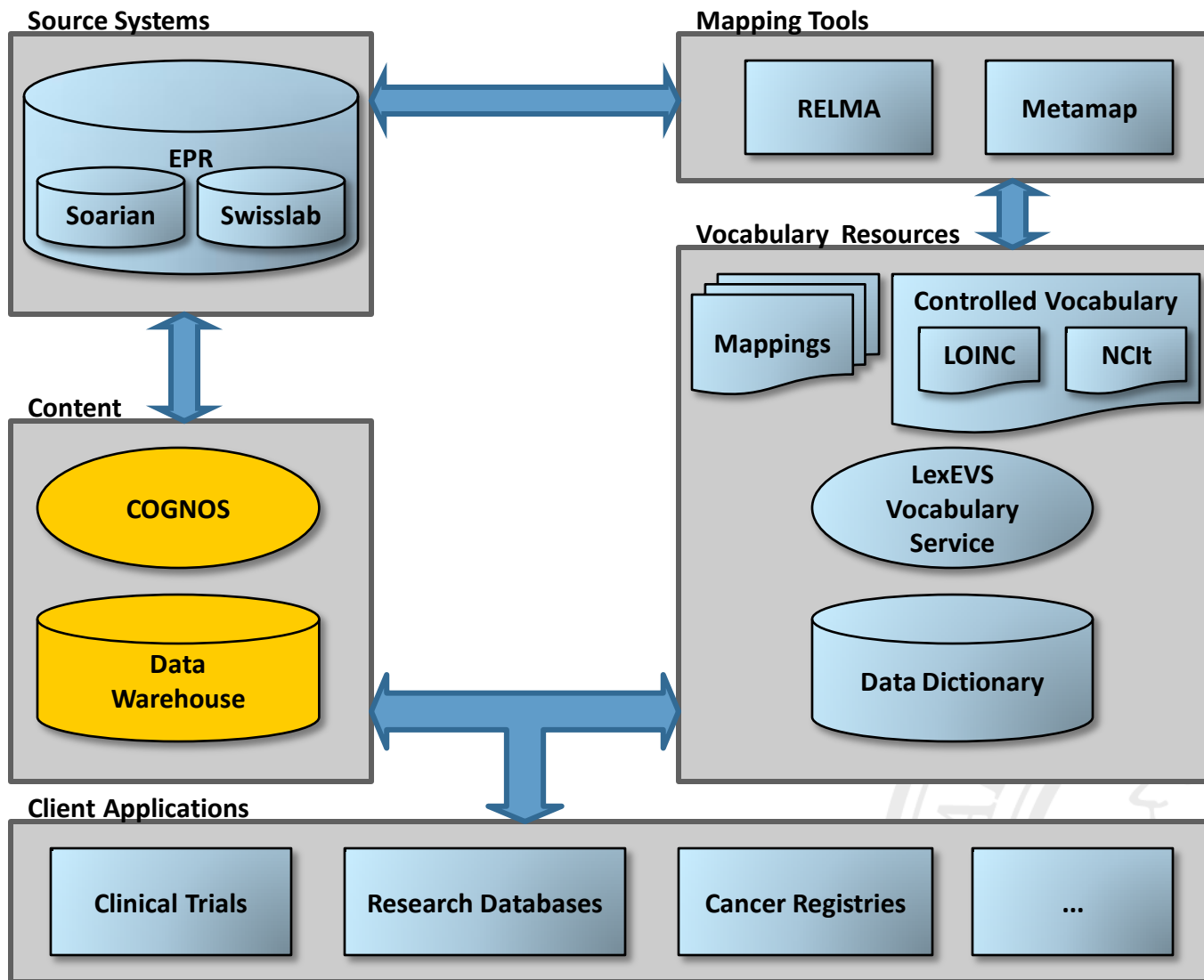
Methoden: Komponenten der Architektur

logische Sicht



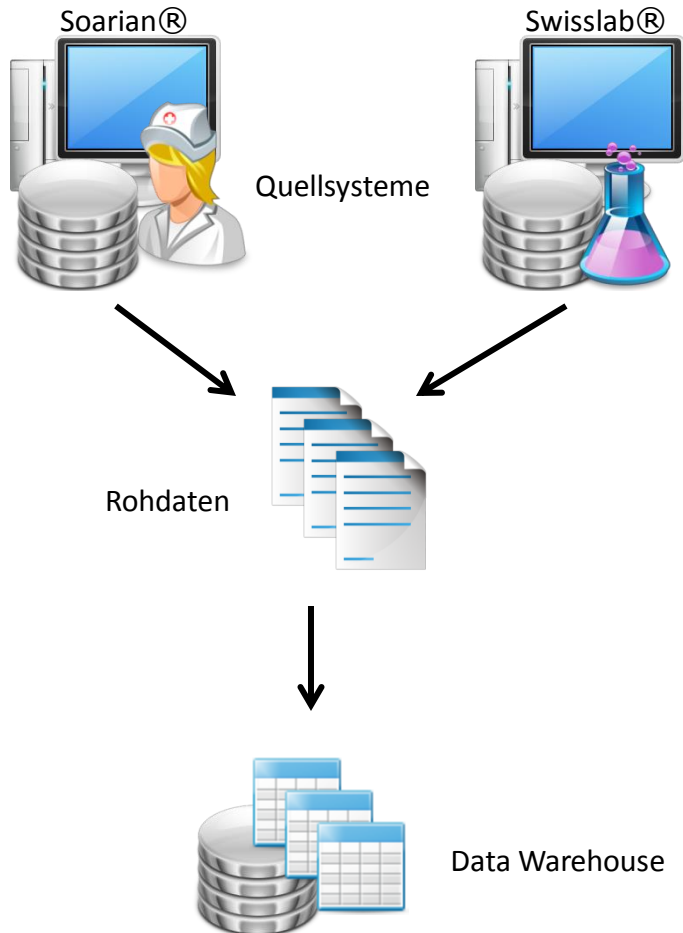
Methoden: Komponenten der Architektur

logische Sicht – Modul Content/DWH



Methoden: Komponenten der Architektur

logische Sicht – Modul Content/DWH

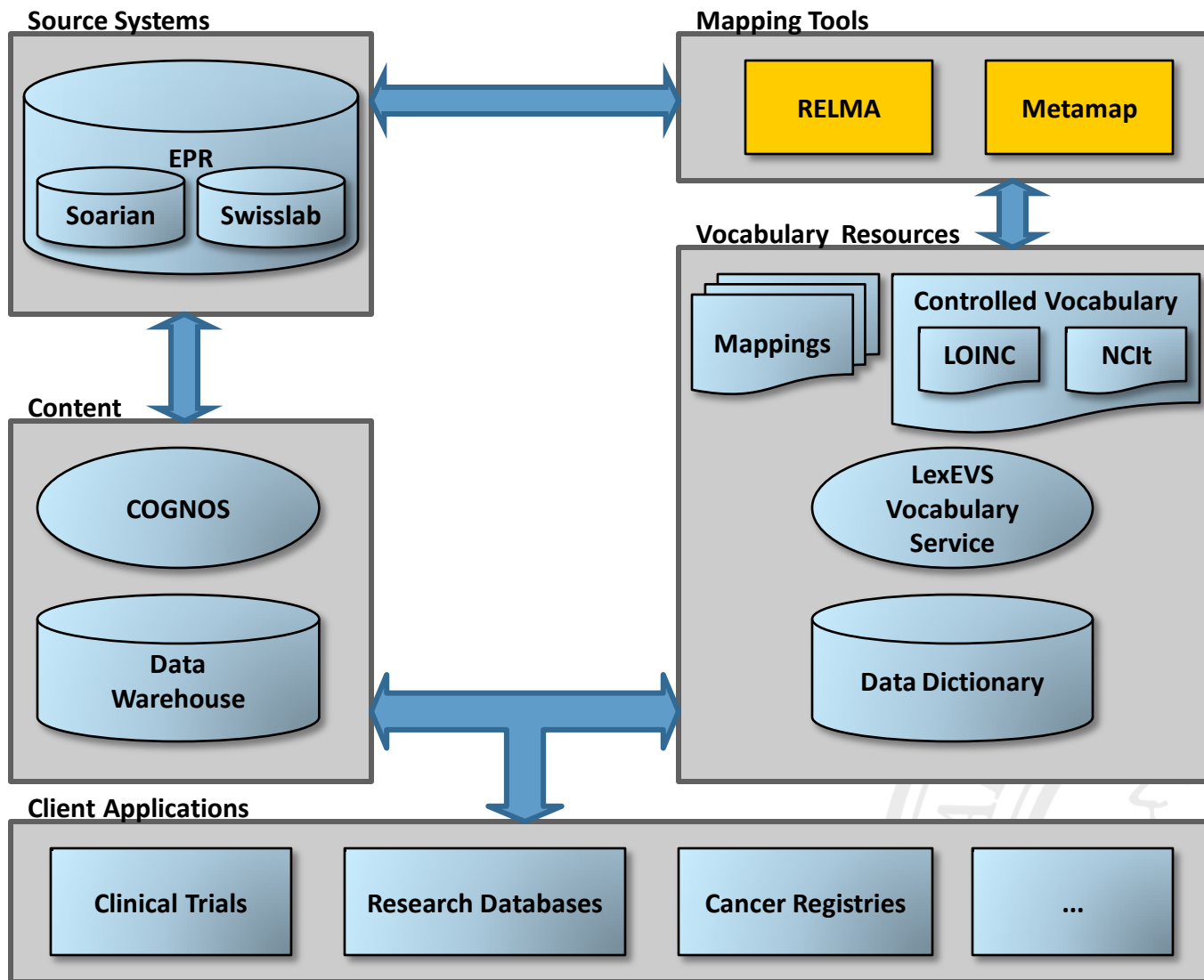


- DWH basierend auf Oracle®-Datenbank
- Aufgebaut mit COGNOS®-Tools (von IBM Inc.)
- Import von Rohdaten der einzelnen Quellen:

1x täglich ETL

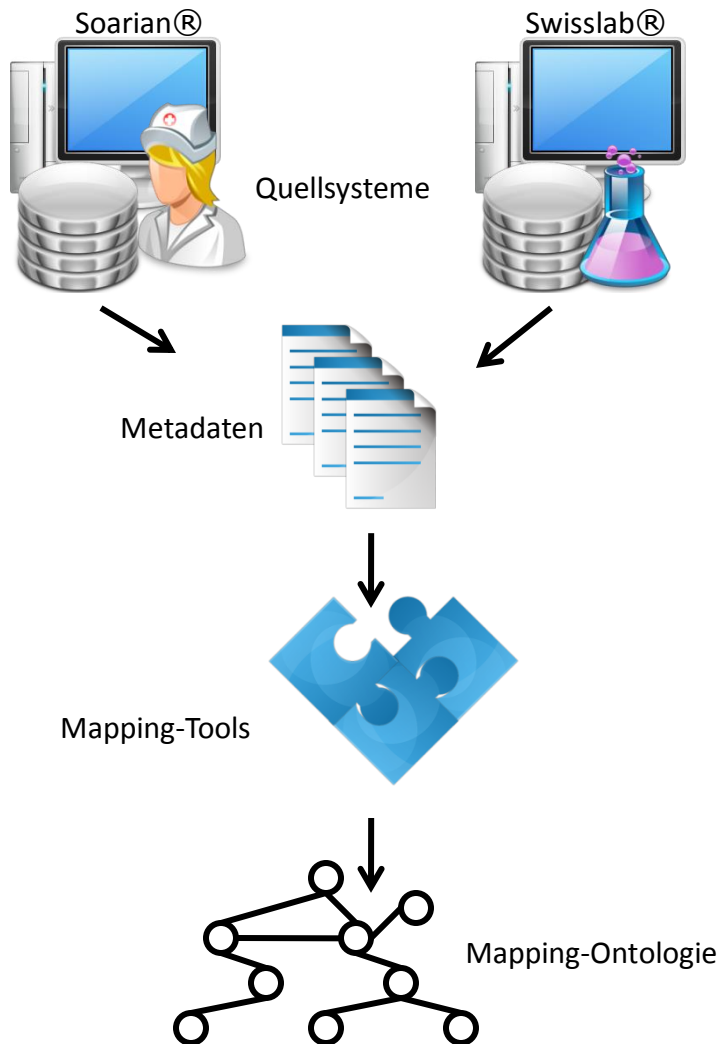
Methoden: Komponenten der Architektur

logische Sicht – Modul Mapping Tools



Methoden: Komponenten der Architektur

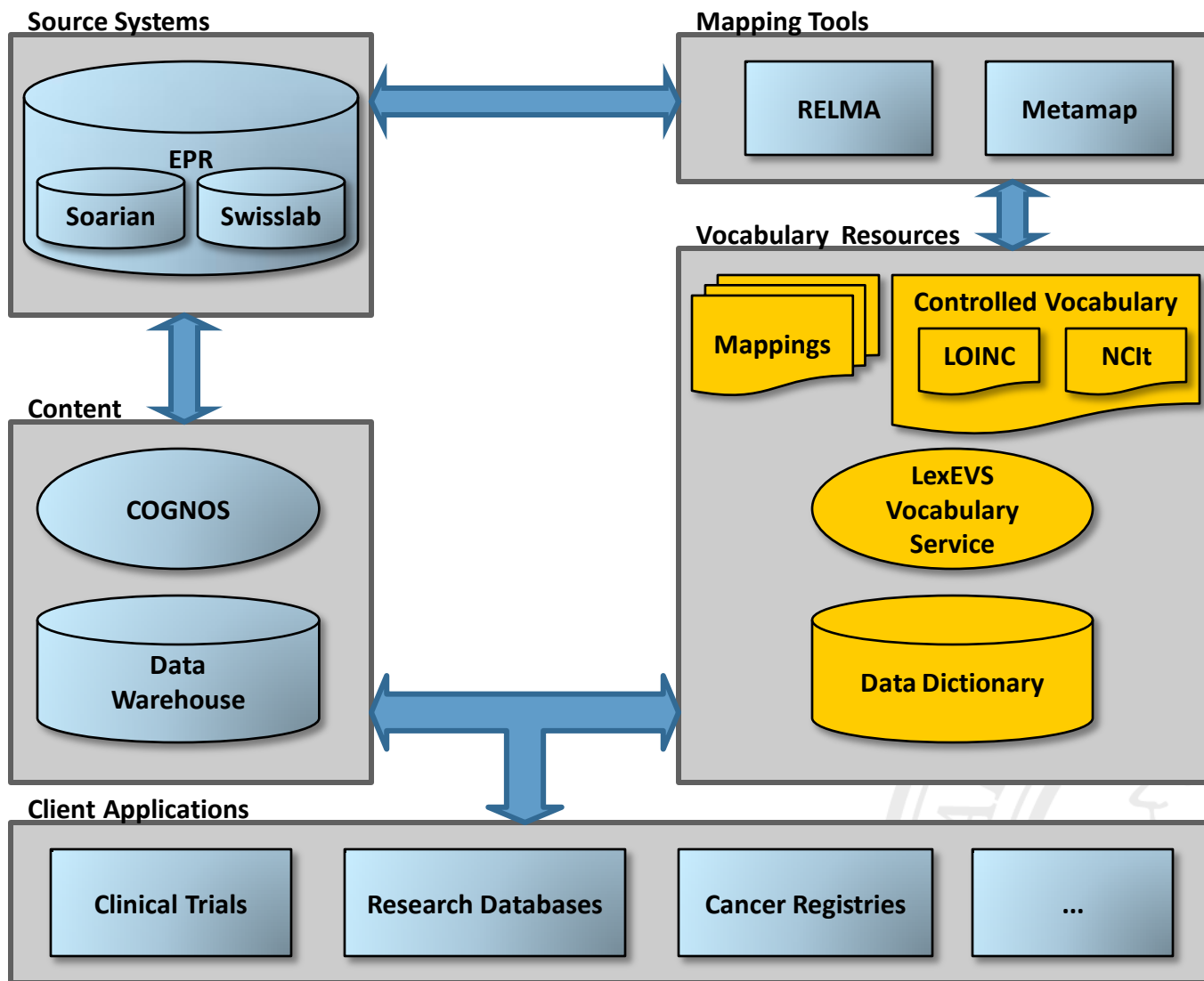
logische Sicht – Modul Mapping Tools



- Export der Metadaten aus den Quellsystemen
- Verwendung der Mapping-Tools
 - **RELMA** (www.loinc.org) für LOINC und
 - **Metamap** (www.nlm.nih.gov) für NCI-Thesaurus (www.ncit.nci.nih.gov)
- Abbildung der Items auf entsprechende Konzepte

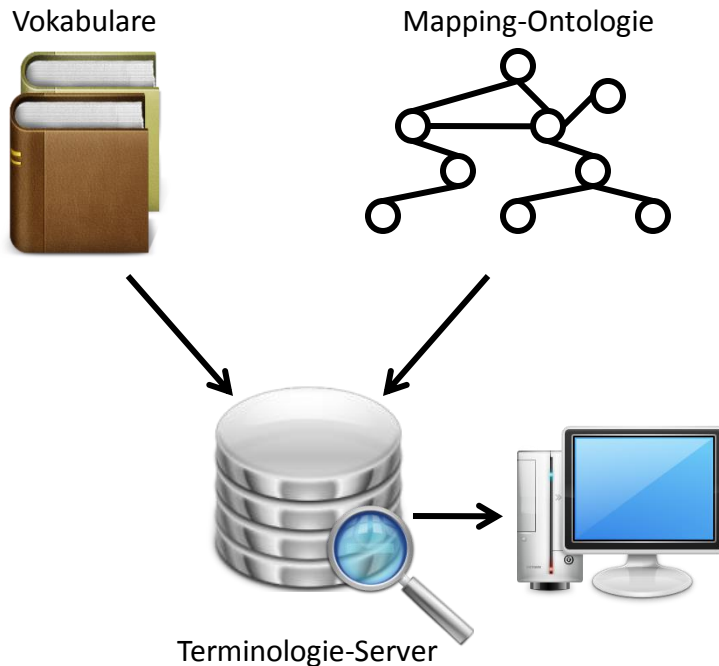
Methoden: Komponenten der Architektur

logische Sicht – Modul Vocabulary Resources



Methoden: Komponenten der Architektur

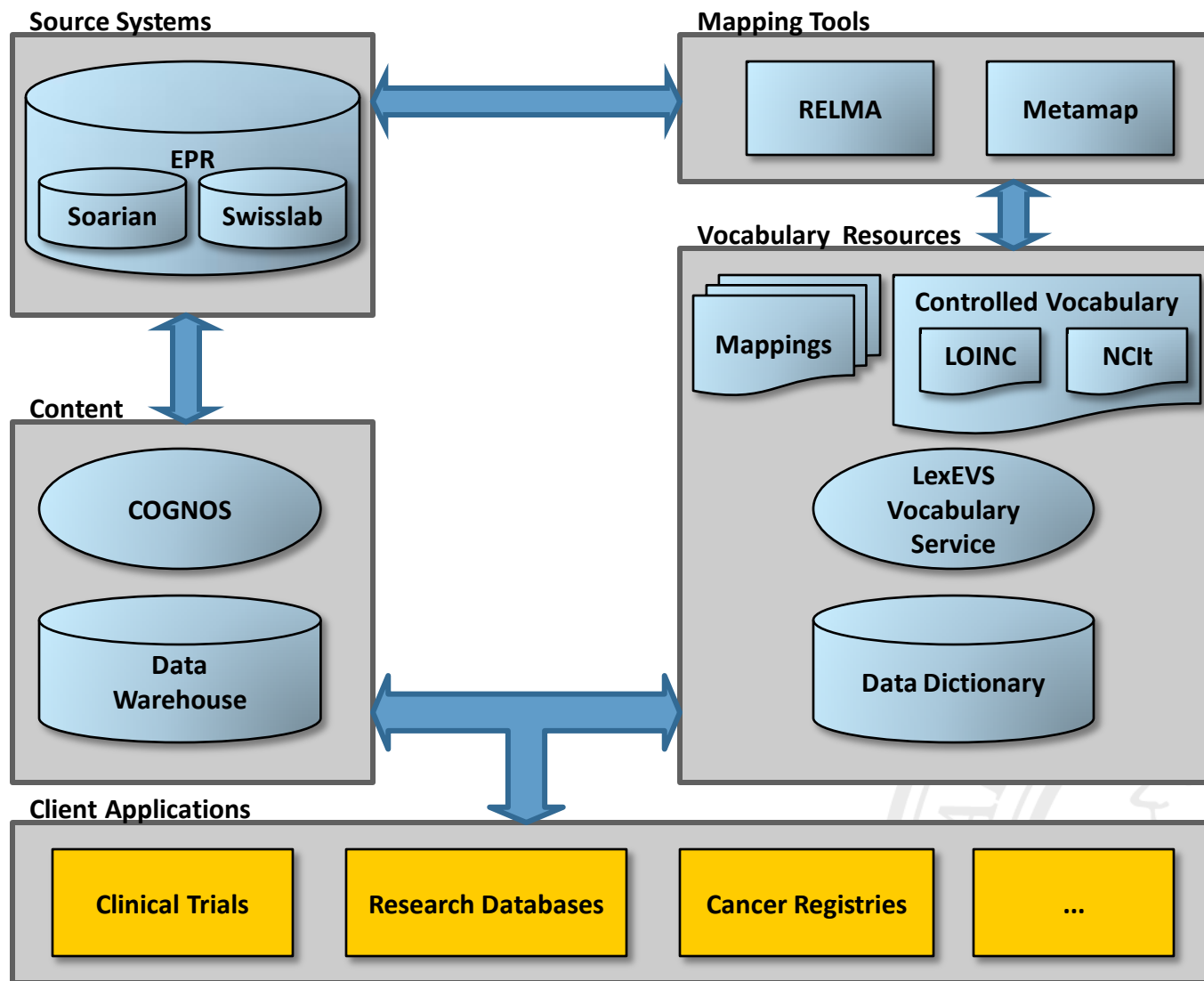
logische Sicht – Modul Vocabulary Resources



- **LexEVS** als Terminologie-Server (www.ncicb.nci.nih.gov)
- Import von Mappings und LOINC (www.loinc.org) sowie NCI-Thesaurus (www.ncit.nci.nih.gov) als Pivot-Terminologie
- GUI und Schnittstellen für gezielte Abfragen

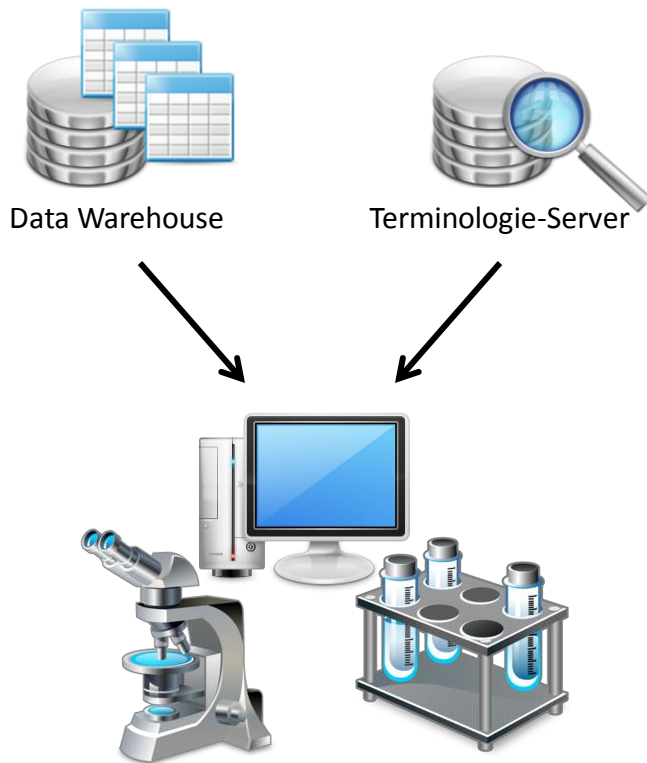
Methoden: Komponenten der Architektur

logische Sicht – Modul Client Applications



Methoden: Komponenten der Architektur

logische Sicht – Modul Client Applications



- ETL-Jobs zum Export in Client-Anwendungen
- Semantische Aufbereitung der Rohdaten via Abfrage der ontologischen Mappings
- Mögliche Ziele am UKE: GTDS[®]-Krebsregister (www.gtds.de) oder i2b2 (www.i2b2.org)

Ergebnisse: Anwendung der Architektur in zwei Pilot-Szenarien am UKE

Szenario 1: Labordaten

Mapping von
Swisslab[®]-Analyten
auf LOINC

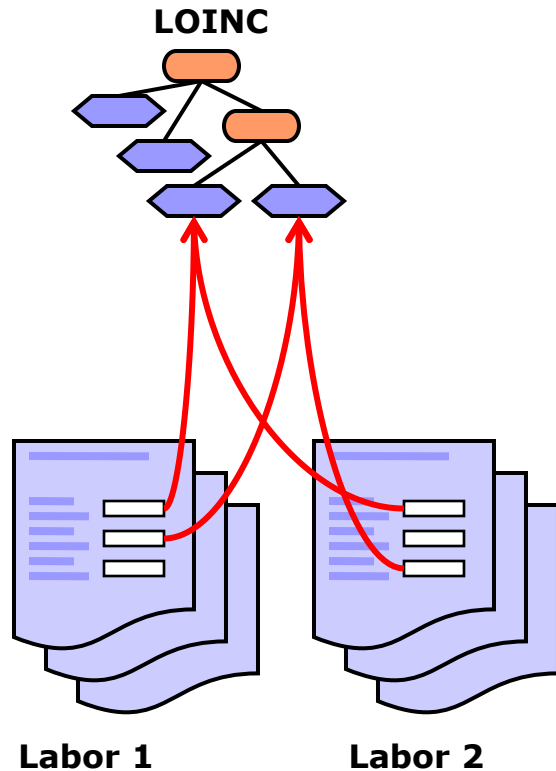
Szenario 2: Tumordaten

Mapping von
Soarian[®]- und GTDS[®]-Items
auf den NCI-Thesaurus



Ergebnisse: Anwendung der Architektur

Szenario 1: Labordaten



- Semiautomatisches Mapping der Swisslab[®]-Analyte von 13 Laboren auf LOINC mit RELMA
- LOINC als Pivot-Terminologie für Abbildungen zwischen den Laboren
- Abbildung von 10.206 Analyten auf 2.564 LOINCs
- Korrektheit: 98%

Ergebnisse: Anwendung der Architektur

Szenario 1: Labordaten

LOINC

■ Semiautomatisches

Downloaded from jamia.bmj.com on May 17, 2014 - Published by group.bmj.com

Research and applications

Mapping local laboratory interface terms to LOINC at a German university hospital using RELMA V.5: a semi-automated approach

Christian Zunner,¹ Thomas Bürkle,¹ Hans-Ulrich Prokosch,^{1,2} Thomas Ganslandt²

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

²Medical Center for Information and Communication, Erlangen University Hospital, Erlangen, Germany

Correspondence to

Christian Zunner, Chair of Medical Informatics, University of Erlangen-Nuremberg, Krankenhausstraße 12, Erlangen 91054, Germany; christian.zunner@uk-erlangen.de

Received 24 April 2012

Accepted 23 June 2012

Published Online First

16 July 2012

ABSTRACT

Objective Logical Observation Identifiers Names and Codes (LOINC) mapping of laboratory data is often a question of the effort of mapping compared with the benefits of the structure achieved. The new LOINC mapping assistant RELMA (version 2011) has the potential to reduce the effort required for semi-automated mapping. We examined quality, time effort, and sustainability of such mapping.

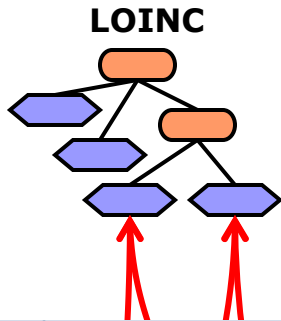
Methods To verify the mapping quality, two samples of 100 laboratory terms were extracted from the laboratory system of a German university hospital and processed in a semi-automated fashion with RELMA V.5 and LOINC V.2.34 German translation DIMDI to obtain LOINC codes. These codes were reviewed by two experts from each of two laboratories. Then all 2148 terms used in these two laboratories were processed in the same way.

However, other content of electronic medical records (EMRs), even if stored in a structured fashion, is often coded using disparate standards or locally developed schemes. Unless these individual interface terminologies are carefully inspected and mapped to a common set of concepts, querying these datasets remains cumbersome and error-prone. Thus the lack of 'semantic interoperability' remains a fundamental barrier to the full realization of the single-source paradigm.^{9–13}

Laboratory analysis results offer an exceptional potential for reuse in a single-source scenario, as they are typically well-structured and well-defined data elements, often generated in an automated, quality-controlled fashion.¹⁴ As EMRs become more widespread, clinical laboratories are increasingly delivering reports electronically in a form that

Ergebnisse: Anwendung der Architektur

Szenario 1: Labordaten



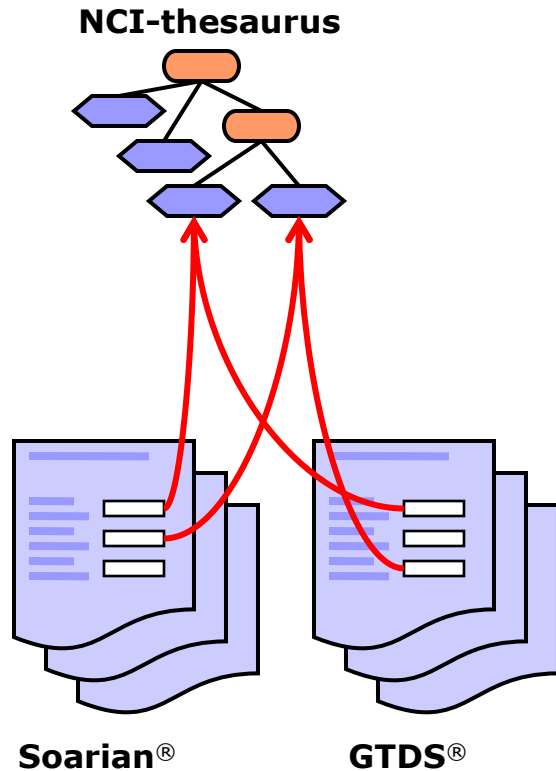
- Semiautomatisches Mapping der Swisslab[®]-Analyte von 13 Laboren auf LOINC mit RELMA

TESTCODE	TESTDESCRIPTION	Einheit	LOINC	LOINCLONGCOMMONNAME
6GRAVS	Grav.-Schnelltest (Serum)		2118-8	Choriogonadotropin (pregnancy test) [Presence] in Serum or Plasma
6K	Kalium	mmol/l	2823-3	Potassium [Moles/volume] in Serum or Plasma
6NA	Natrium	mmol/l	2951-2	Sodium [Moles/volume] in Serum or Plasma
6CL	Chlorid	mmol/l	2075-0	Chloride [Moles/volume] in Serum or Plasma
6GLUS	Glucose	mg/dl	2345-7	Glucose [Mass/volume] in Serum or Plasma
6HB1C	HbA1c	%	4548-4	Hemoglobin A1c/Hemoglobin.total in Blood
6CA	Calcium	mmol/l	2000-8	Calcium [Moles/volume] in Serum or Plasma
6HARS	Harnsäure	mg/dl	3084-1	Urate [Mass/volume] in Serum or Plasma
6HASt	Harnstoff	mg/dl	3091-6	Urea [Mass/volume] in Serum or Plasma
6KREA	Kreatinin	mg/dl	2160-0	Creatinine [Mass/volume] in Serum or Plasma
6GFR	GFR (geschätzt MDRD-Formel)	ml/min	33914-3	Glomerular filtration rate/1.73 sq M.predicted by Creatinine-based formula (MDRD)
6BILI	Bilirubin gesamt	mg/dl	1975-2	Bilirubin.total [Mass/volume] in Serum or Plasma
6BILD	Bilirubin direkt	mg/dl	1968-7	Bilirubin.direct [Mass/volume] in Serum or Plasma
6BIND	Bilirubin indirekt	mg/dl	1971-1	Bilirubin.indirect [Mass/volume] in Serum or Plasma

- Korrektheit: 98%

Ergebnisse: Anwendung der Architektur

Szenario 2: Tumordaten

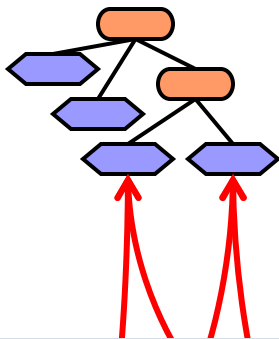


- Vollautomatisches Mapping der Soarian®- und GTDS®-Items auf den NCI-thesaurus mit Metamap
- NCI-thesaurus als Pivot-Terminologie zwischen Soarian® und GTDS®
- Abbildung von 1.142 items auf 249 GTDS®-items
- Korrektheit: 79%

Ergebnisse: Anwendung der Architektur

Szenario 2: Tumordaten

NCI-thesaurus



- Vollautomatisches Mapping der Soarian[®]- und GTDS[®]-Items auf den NCI-thesaurus mit Metamap

Displayname_Krebsregister	NCIt_Name	NCIt_CUI	Displayname_Gyn
Histologie Code	ICD-O-3 Code	C66766	ICD-O:
M	Pathologic Distant Metastasis TNM Finding	C48886	pM:
T	Pathologic Primary Tumor TNM Finding	C48887	pT:
N	Pathologic Regional Lymph Nodes TNM Finding	C48888	pN:
Operation	Operation	C15329	OP als Therapie
Bestrahlung	Radiotherapy	C15313	Bestrahlung als Therapie
Chemotherapie	Chemotherapy	C15632	Chemotherapie als Therapie

- Korrektheit: 79%

Diskussion

- Generische Single-Source-Architektur aus weitgehend frei verfügbaren Komponenten
- Standardisierte erweiterbare Pivot-Terminologie
- Einzelne Komponenten ohne größeren Aufwand austauschbar
- Mapping-Aufwand nicht vermeidbar
- Ontologischer Ansatz, erfordert strukturierte Daten
- Je nach Verwendung Validierung der Mappings erforderlich



VIELEN DANK

