

IT-Infrastruktur als Schlüsseltechnologie für die Forschung (Versorgungsdatenforschung, Single Source)

Perspektiven der Medizinischen Informatik, Biometrie und Epidemiologie



Hans-Ulrich Prokosch

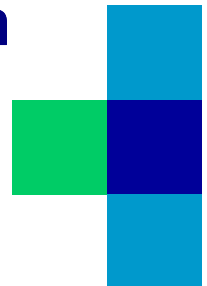
Lehrstuhl für Medizinische Informatik

Universität Erlangen-Nürnberg

CIO Universitätsklinikum Erlangen



FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
MEDIZINISCHE FAKULTÄT

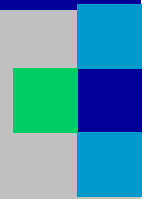


Meine Gießener Historie

- **Diplom Mathematik (1981)**
- **Zivildienst am Gießener Tumorzentrum (1984-1986)**
- **Promotion Medizinische Informatik (1988)**
- **Habilitation Medizinische Informatik (1994)**

- **Dozent an der Schule für Medizinische Dokumentare**
- **Wissenschaftlicher Mitarbeiter bei Prof. Dudeck**
- **Leiter der Abteilung Klinische Informationssysteme**

IT-Infrastruktur als Schlüsseltechnologie für die **Forschung in der Medizin**



Klinische Studien

Kohorten

IITs

Arzneimittelstudie

Good Clinical Practice

Identitäts-

Translationale Forschung

Management

Ethikkommission

Biobanking

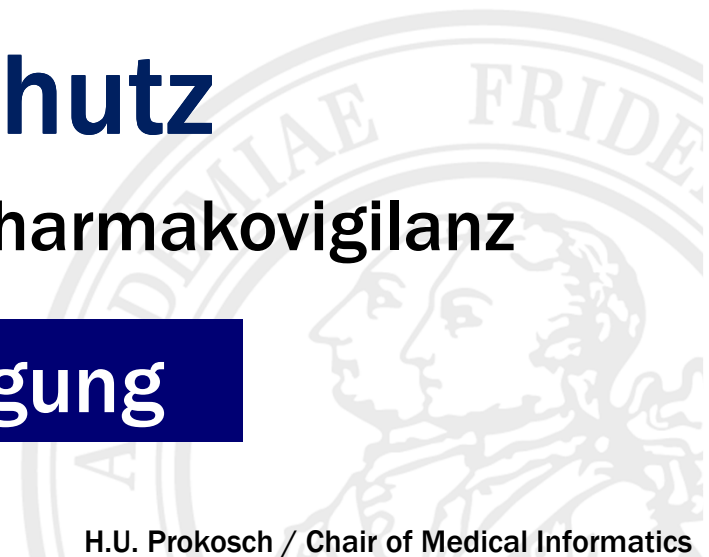
Patientenrekrutierung

Datenschutz

Feasibility Phase

Pharmakovigilanz

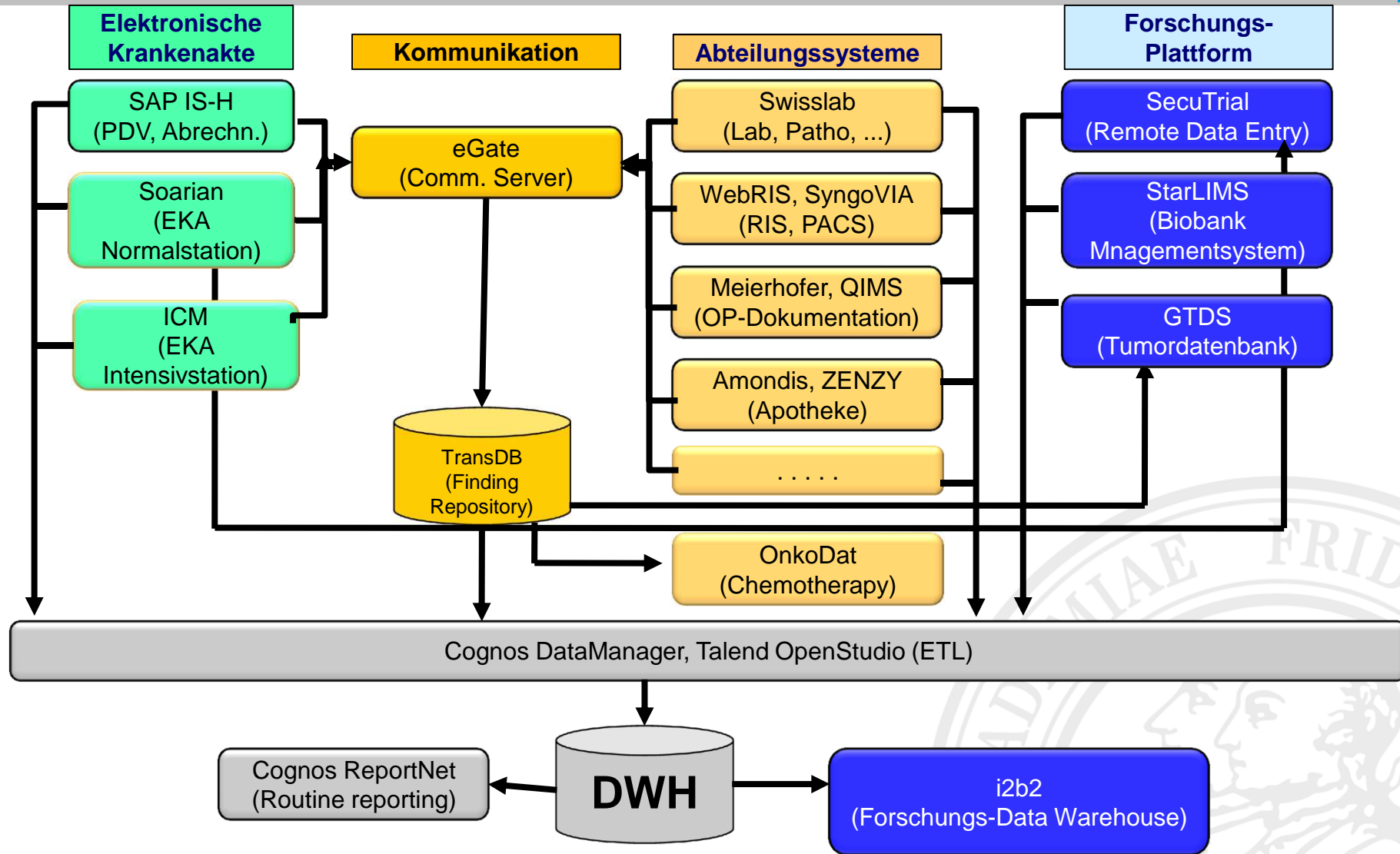
Patienteneinwilligung



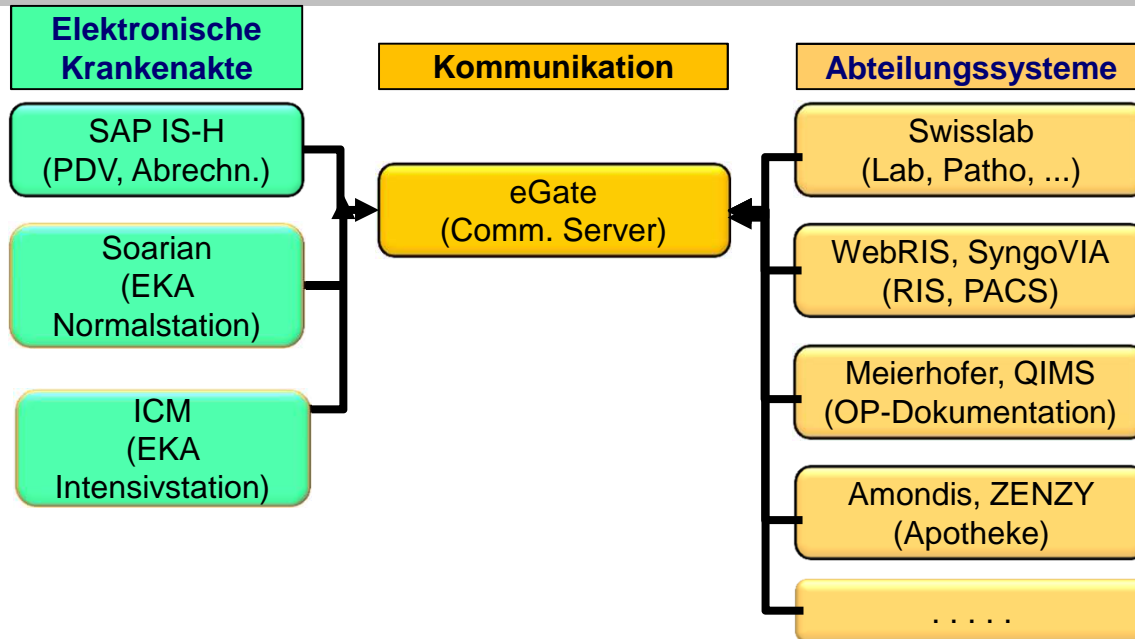
IT-Infrastruktur als Schlüsseltechnologie für die Forschung in der Medizin



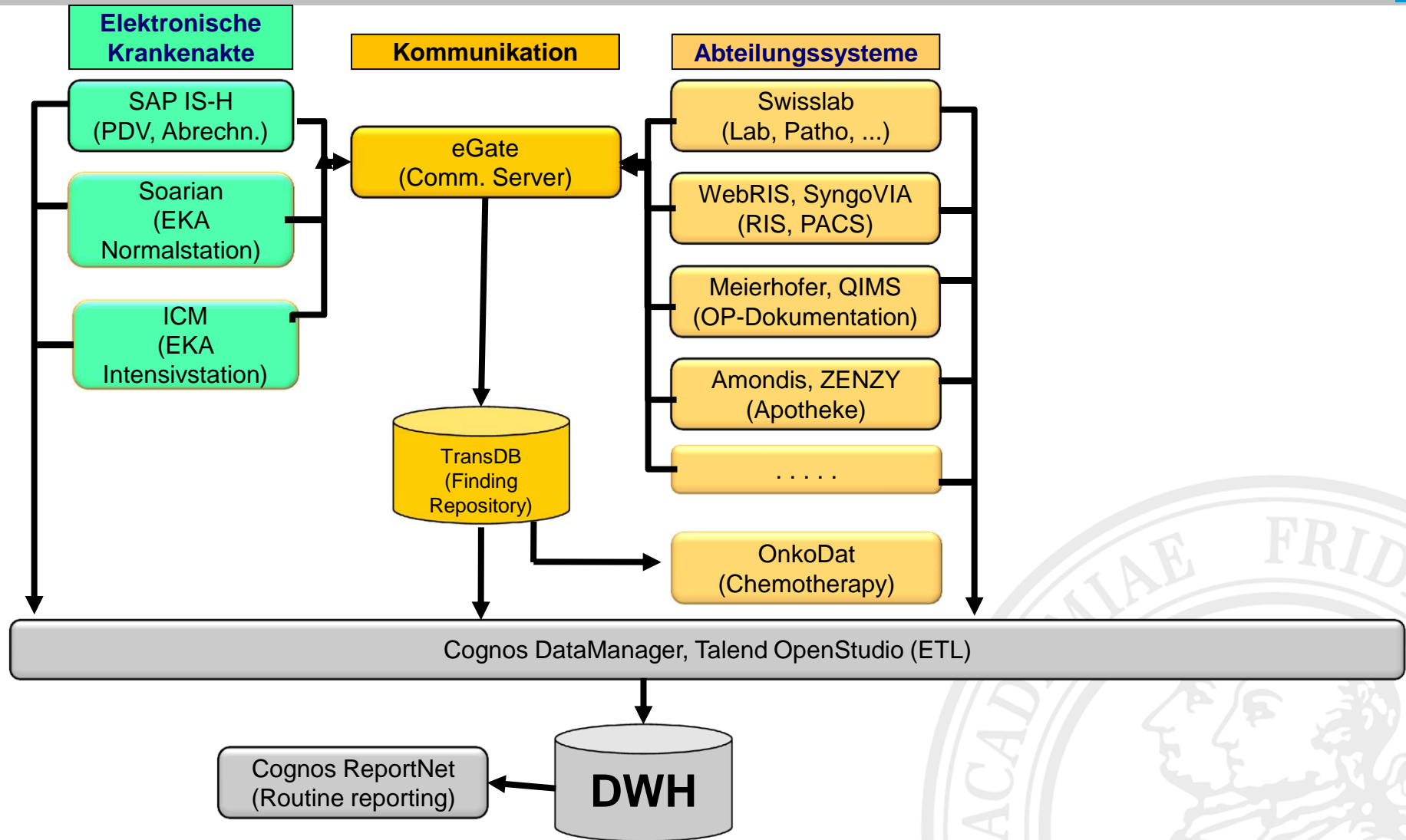
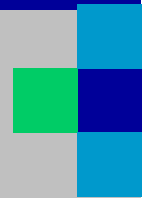
Das große Bild



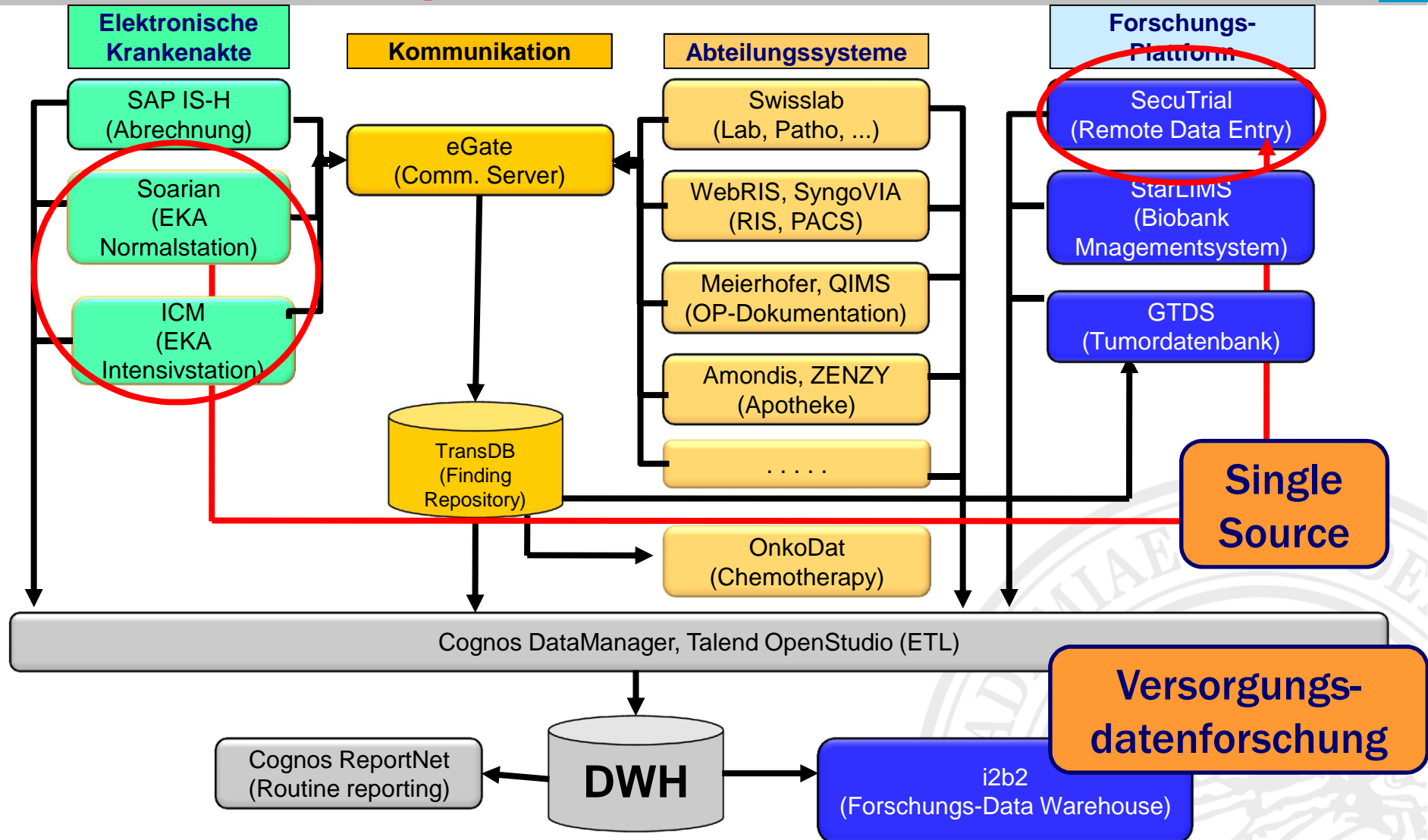
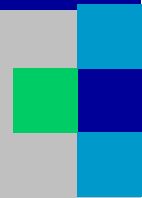
IT als Schlüsseltechnologie in der Krankenversorgung ⇒ KIS



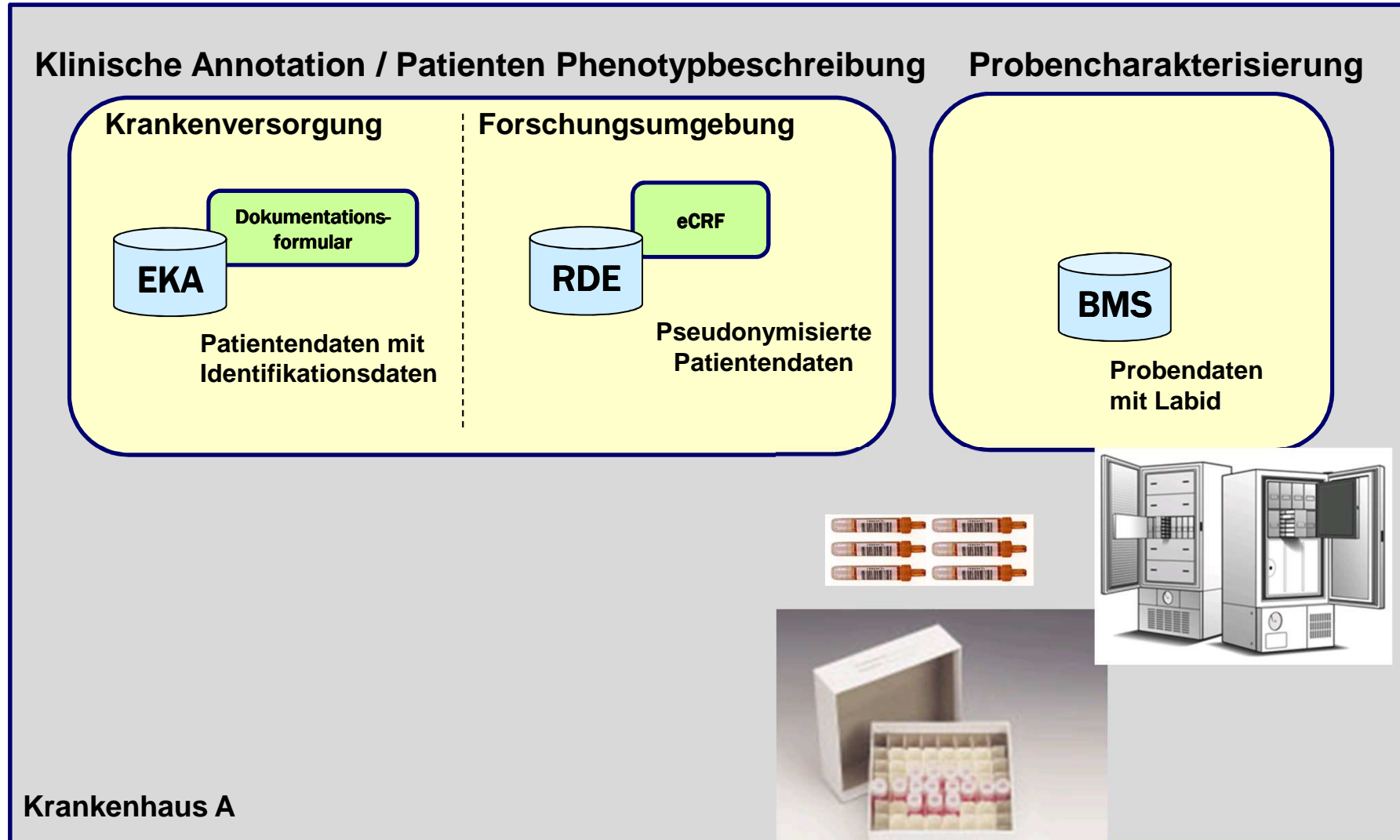
IT als Schlüsseltechnologie in der Krankenversorgung ⇒ KIS



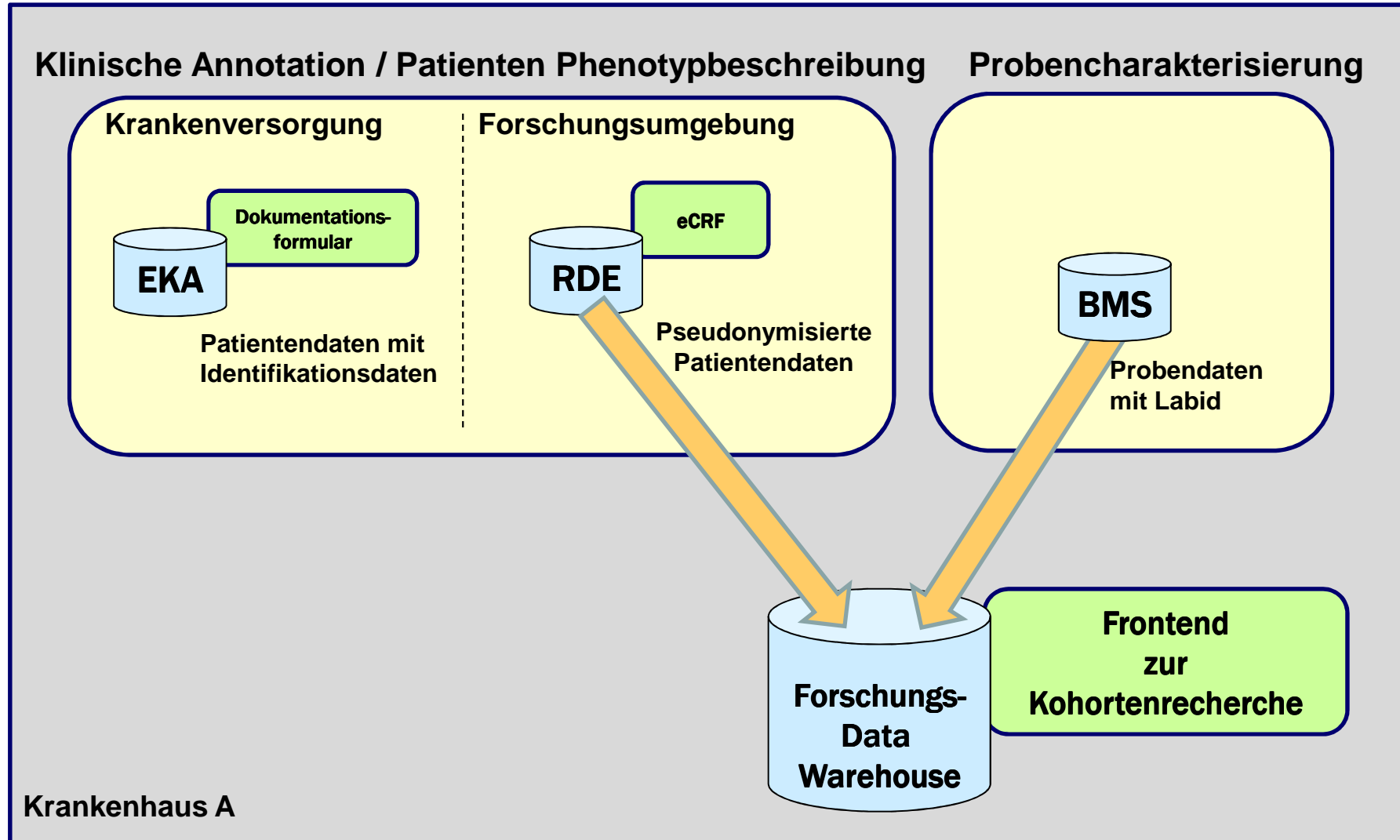
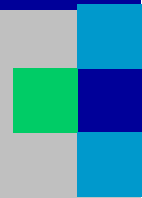
IT-Infrastruktur als Schlüsseltechnologie für die Forschung in der Medizin



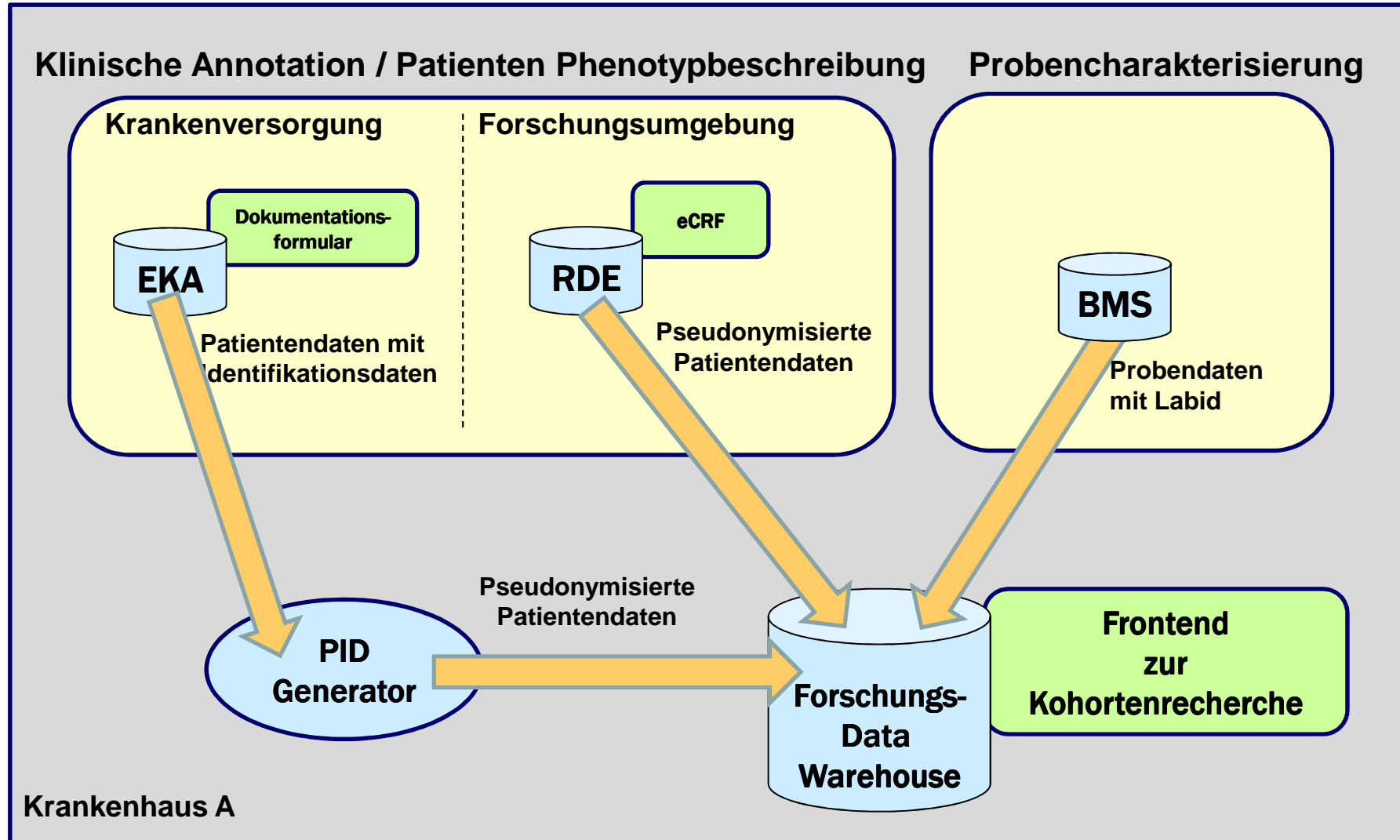
Kernkomponenten einer Standort-bezogenen Forschungsinfrastruktur



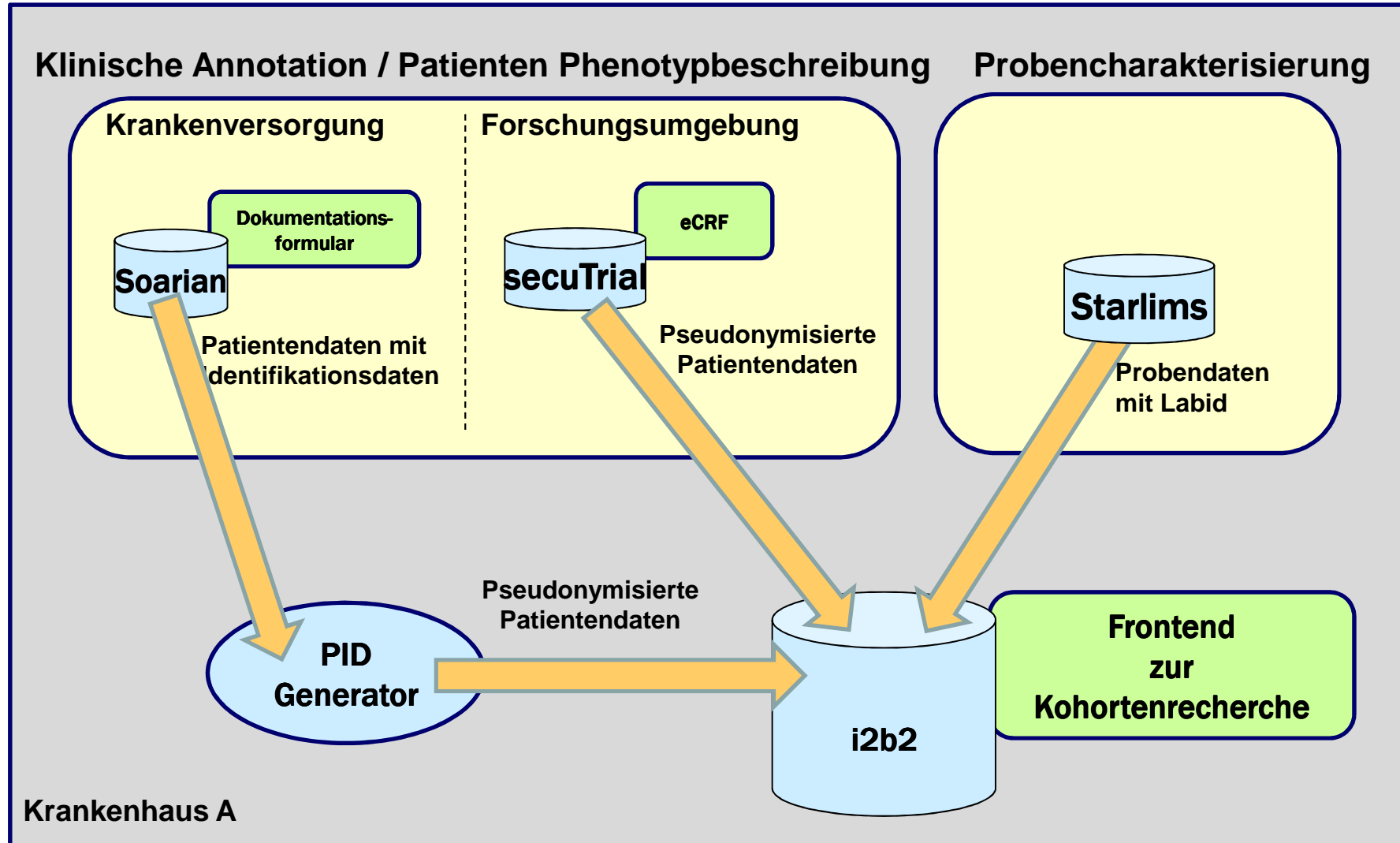
Kernkomponenten einer Standort-bezogenen Forschungsinfrastruktur



Kernkomponenten einer Standort-bezogenen Forschungsinfrastruktur



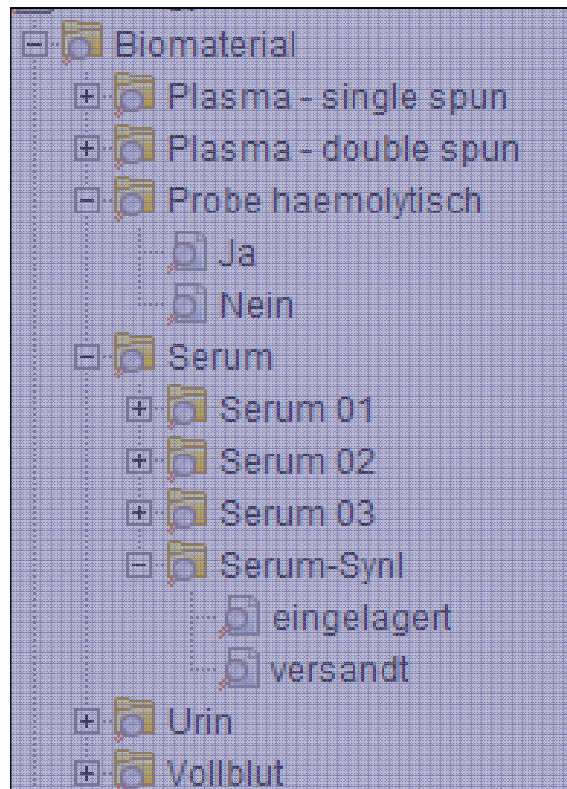
Kernkomponenten einer Standort-bezogenen Forschungsinfrastruktur



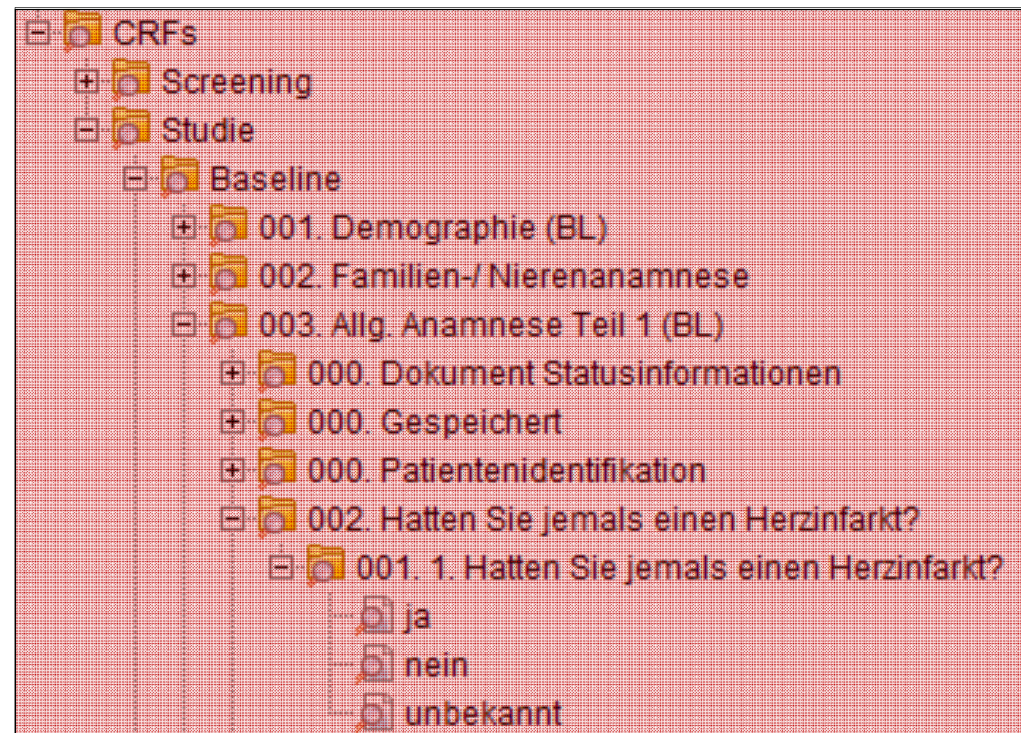
Beispiel 1: Zusammenführen von Daten aus unterschiedlichen Primärsystemen

GCKD Kohorte: 5.000 Patienten mit chronischer Nierenerkrankung

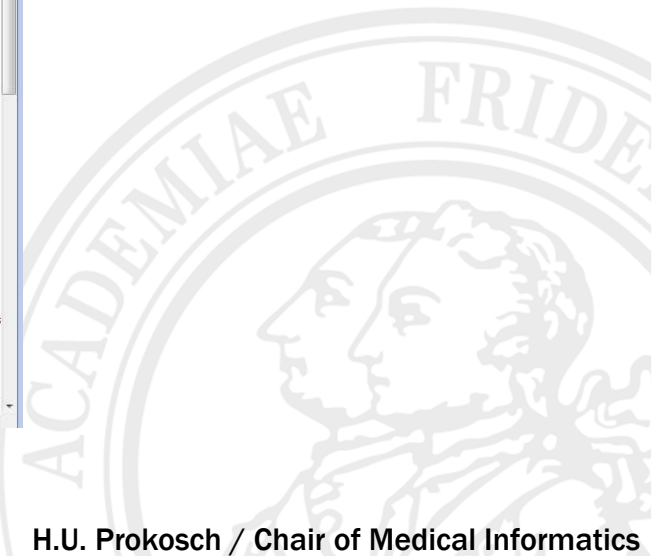
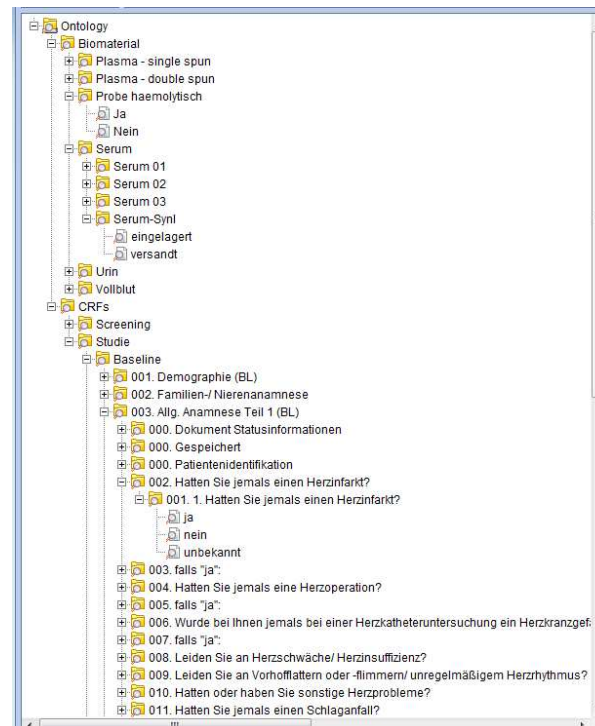
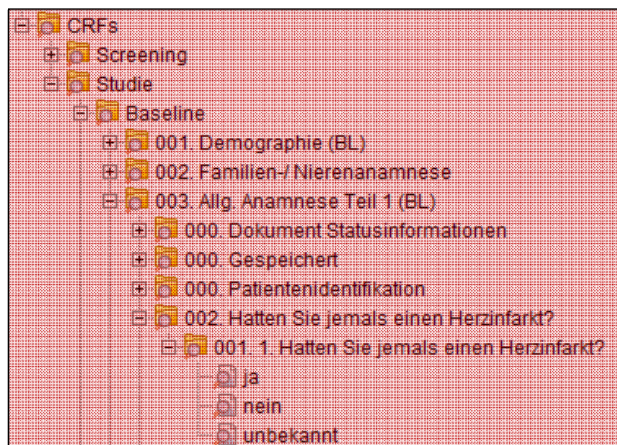
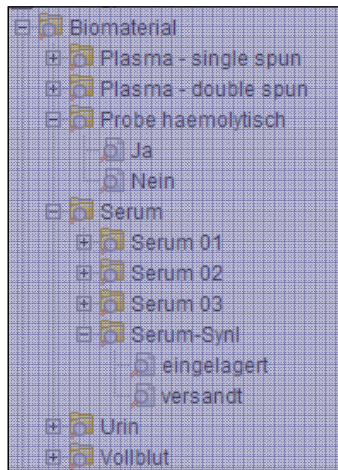
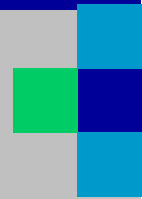
Biobank Managementsystem



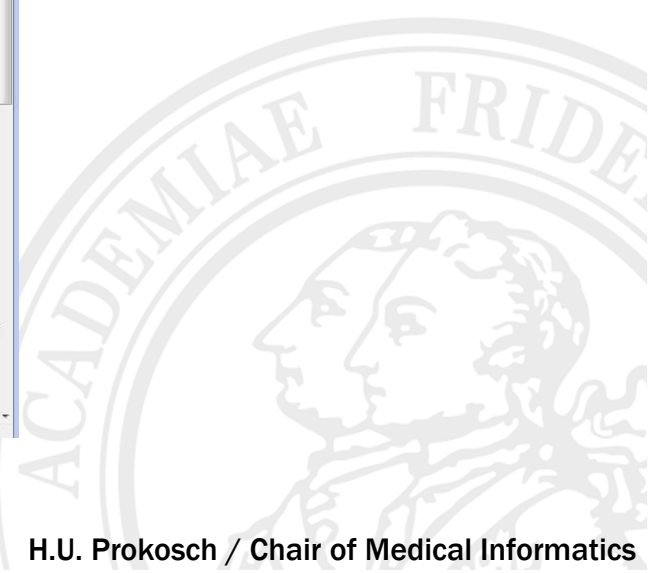
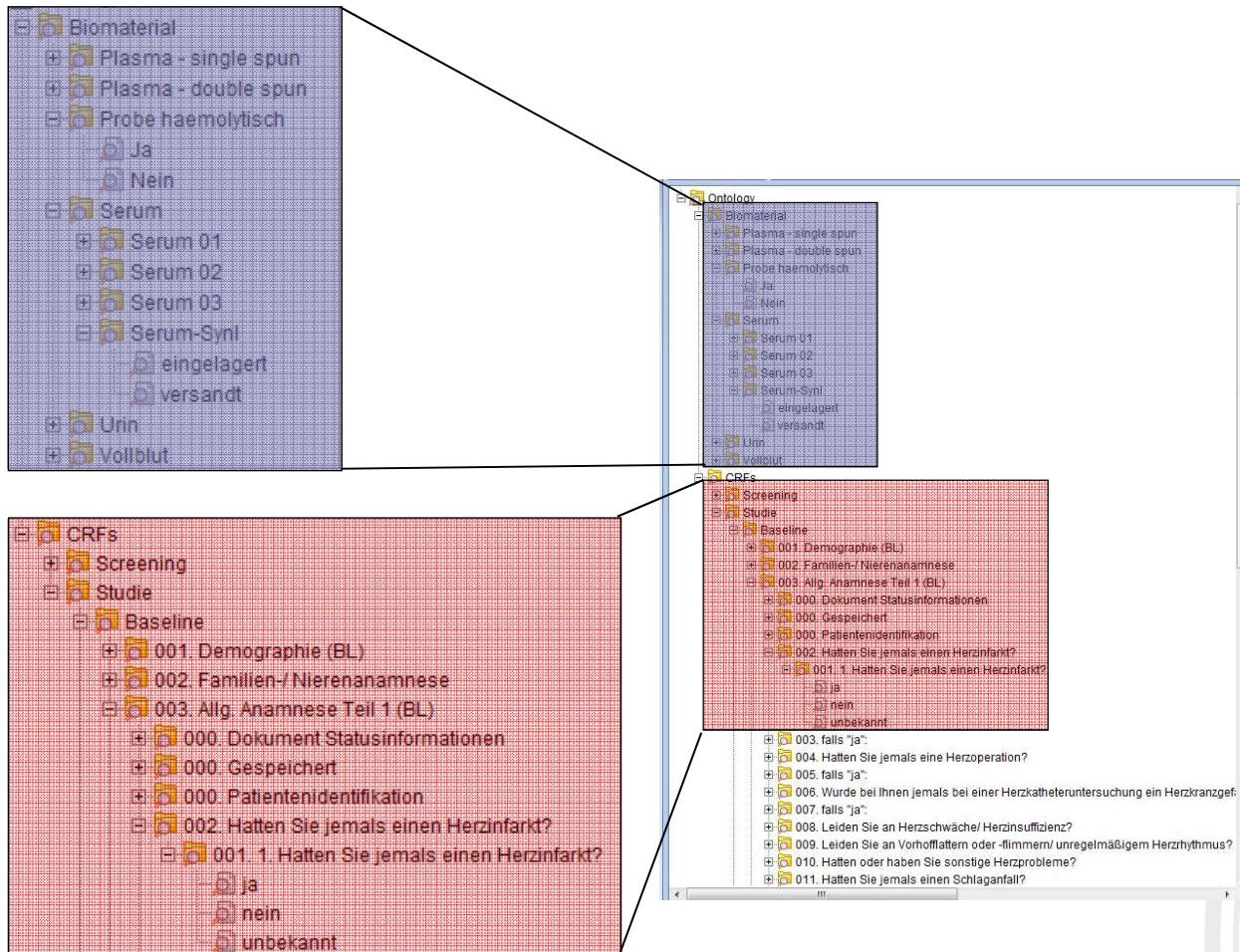
Remote Data Entry System



Beispiel 1: Zusammenführen von Daten aus unterschiedlichen Primärsystemen



Beispiel 1: Zusammenführen von Daten aus unterschiedlichen Primärsystemen



Beispiel 1: Zusammenführen von Daten aus unterschiedlichen Primärsystemen

The screenshot displays the i2b2 Query & Analysis Tool interface. On the left, a hierarchical ontology tree is shown. Three callout boxes provide detailed views of specific parts of this ontology:

- Top Callout (Blue):** Shows the 'Biomaterial' hierarchy, including 'Plasma - single spun', 'Plasma - double spun', 'Probe haemolytisch' (with 'Ja' and 'Nein' options), 'Serum' (with 'Serum 01', 'Serum 02', 'Serum 03', and 'Serum-Synl' sub-items), and 'Urin'.
- Bottom Callout (Red):** Shows the 'CRFs' hierarchy, including 'Screening' and 'Studie' (with 'Baseline' sub-items). The 'Baseline' section includes items like '001. Demographie (BL)', '002. Familien-/ Nierenanamnese', and '003. Allg. Anamnese Teil 1 (BL)'. The '003' item has sub-items '000. Dokument Statusinformationen', '000. Gespeichert', and '000. Patientenidentifikation'. The '002' item has sub-items '002. Hatten Sie jemals einen Herzinfarkt?' and '001. 1. Hatten Sie jemals einen Herzinfarkt?' (with 'ja', 'nein', and 'unbekannt' options).
- Middle Callout (Red):** Shows a detailed view of the 'CRFs' hierarchy, including 'Screening' and 'Studie' (with 'Baseline' sub-items). The 'Baseline' section includes items like '001. Demographie (BL)', '002. Familien-/ Nierenanamnese', and '003. Allg. Anamnese Teil 1 (BL)'. The '003' item has sub-items '000. Dokument Statusinformationen', '000. Gespeichert', and '000. Patientenidentifikation'. The '002' item has sub-items '002. Hatten Sie jemals einen Herzinfarkt?' and '001. 1. Hatten Sie jemals einen Herzinfarkt?' (with 'ja', 'nein', and 'unbekannt' options).

The main interface shows the 'i2b2 Query & Analysis Tool' with a 'Query Tool' section on the right. This section contains a query builder with three groups (Group 1, Group 2, Group 3) and a 'Query Status' section at the bottom. The 'Query Status' section shows the following information:

Query Name	Status	Time
Finished Query: "männl-ja-versa@14:25:22"	FINISHED	[7.6 secs]
Patient Set - 52 Patients		
Patient Count - 52 patients		

Beispiel 1: Zusammenführen von Daten aus unterschiedlichen Primärsystemen

Ontologiebaum zur Auswahl von Filterkriterien

The screenshot displays the i2b2 Web Client interface in Mozilla Firefox. The main content area is divided into two panes. The left pane, titled 'Ontologiebaum zur Auswahl von Filterkriterien', shows a hierarchical tree structure under 'Ontology'. The tree includes categories like 'Biomaterial', 'Serum', 'CRFs', 'Screening', and 'Studie'. The right pane shows a query builder interface with a list of criteria and logical connectors. The criteria include '001. Demographie (BL)', '002. Familien-/Nierenanamnese', '003. Allg. Anamnese Teil 1 (BL)', and various patient identification and medical history questions. The query builder shows three groups of criteria connected by 'AND' operators. The bottom right corner displays the 'Query Status' section, indicating that the query was finished at 14:25:22 with 52 patients in the patient set and 52 patients in the patient count.

Beispiel 1: Zusammenführen von Daten aus unterschiedlichen Primärsystemen

Recherche-
fenster

The screenshot shows the i2b2 Query & Analysis Tool interface. The main window is titled "Query Tool" and contains a "Query Name:" field and a table with three groups of columns: "Group 1", "Group 2", and "Group 3". Each group has columns for "Dates", "Occurs > 0x", and "Exclude". The table contains the following data:

Group 1	Group 2	Group 3						
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
männlich	ja	versandt						

Below the table, there is a list of search criteria:

- 006. Wurde bei Ihnen jemals bei einer Herzkatheteruntersuchung ein Herzkranzgefäß...
- 007. falls "ja":
- 008. Leiden Sie an Herzschwäche/ Herzinsuffizienz?
- 009. Leiden Sie an Vorhofflattern oder -flimmern/ unregelmäßigem Herzrhythmus?
- 010. Hatten oder haben Sie sonstige Herzprobleme?
- 011. Hatten Sie jemals einen Schlaganfall?

At the bottom right, a "Finished Query" window is displayed, showing the following information:

```

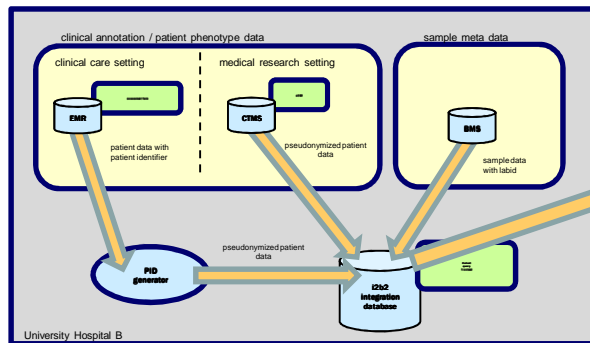
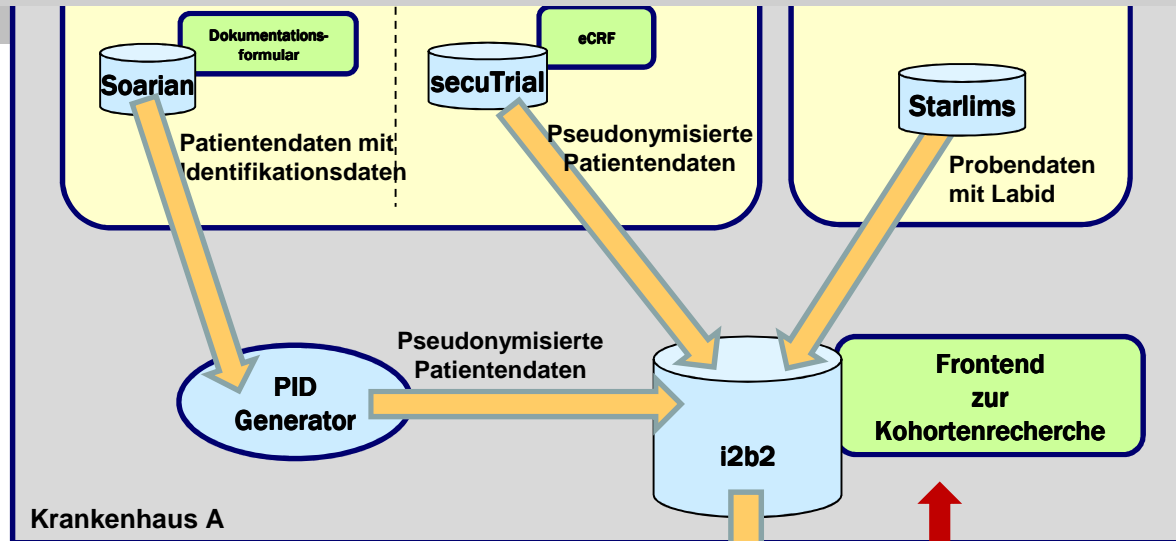
Finished Query: "männl-ja-versa@14:25:22"
Patient Set - 52 Patients - 52 patients
Patient Count - 52 patients
FINISHED [7.6 secs]
FINISHED [7.6 secs]
  
```

Von wie vielen

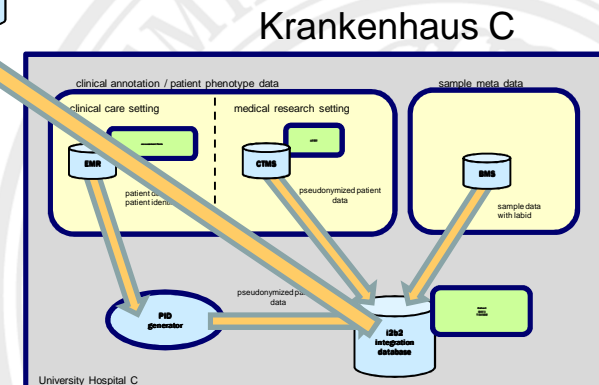
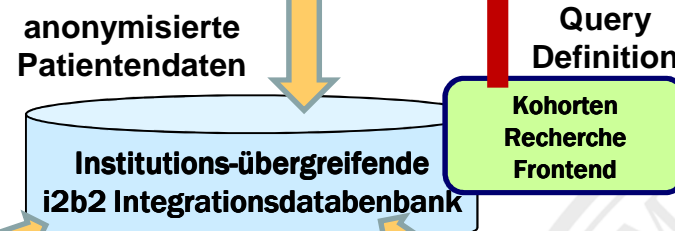
- männlichen Patienten unserer Nierenkohorte
- die einen Herzinfarkt hatten
- haben wir Biomaterial zur externen Analyse versandt?

Ergebnis-
fenster

Kernkomponenten einer Standort-übergreifenden Forschungsinfrastruktur



Krankenhaus B



Beispiele für Standort-übergreifende Integrationsdatenbanken zur Unterstützung der Versorgungsforschung

Downloaded from jamia.bmj.com on April 16, 2012 - Published by group.bmj.com

Brief communication

The Hub Population Health System: distributed ad hoc queries and alerts

Michael D Buck,^{1,2} Sheila Anane,¹ John Taver
Remle Stubbs-Dame,¹ Jesse Singer¹

ABSTRACT

The Hub Population Health System enables the creation of

offi
inv

Enabling collaborative research using the Biomedical Informatics Research Network (BIRN)

Research and applications

Identifying clinical/translational research cohorts: ascertainment via querying an integrated multi-source database

John F Hurdle,¹ Stephen C Haroldsen,² Andrew Hammer,² Cindy Spigle,²
Alison M Fraser,² Geraldine P Mineau,^{3,2} Samir J Courdy²

ABSTRACT

Background Ascertainment of potential subjects has been a longstanding problem in clinical research. Various methods have been proposed, including using data in electronic health records. However, these methods typically suffer from scaling effects—some methods work well for large cohorts; others work for small cohorts only.

Objective We propose a method that provides a simple identification of pre-research cohorts and relies on data available in most states in the USA: merged public health data sources.

Materials and methods The Utah Population Database Limited, a state-level database, was used to build

small, investigator-initiated studies to full-blown randomized controlled trials, and studies in health services or public health research.

Recruitment and participation failure

Research studies fail for many reasons. Estimates of the fraction of randomized controlled trials that fail, or that require extension because of problems with enrollment, range as high as 60%.⁵ To provide an initial assessment of recruitment failure in the broader context of all clinical investigations, we performed a study with our institutional review board (IRB) at the University of Utah. We reviewed the

Ambite,³ Joseph Ames,⁴ Rachana Ananthakrishnan,^{5,6}
Anak,³ Ian Foster,^{5,6} Lee Liming,^{5,6} David Keator,⁴
Dhuri,^{5,6} John-Paul Navarro,^{5,6} Steven Potkin,^{4,7}
Sis,^{8,9} Robert Schuler,³ Jessica A Turner,¹⁰ Arthur Toga,⁸
Sselman,³ for the Biomedical Informatics Research Network

becomes
s can probe ever
requirement that the
quires the
In biomedicine,
later shared
h technological and
r data to be passed
accessible to the
of the Biomedical
is to address the

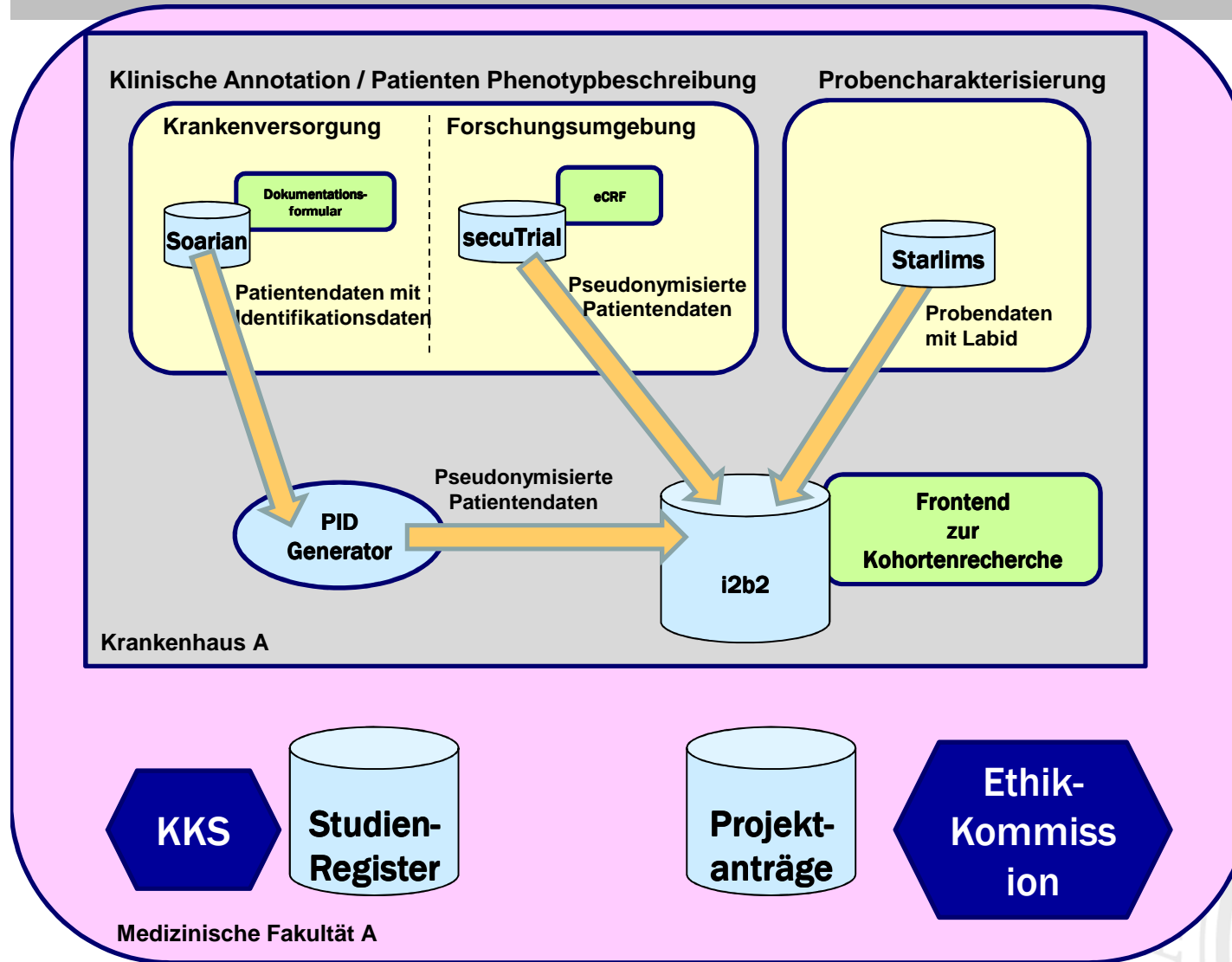
multiple sites, combined, and processed. In practice, however, there are myriad difficulties, both technological and sociological, that must be overcome for data to be accessible to the outside world and for the end result to be useful. For instance, data in clinical settings are generated behind firewalls, and moving those data to the outside world can be difficult, given that the institution must protect all data that contain private health information (PHI). Also, researchers who have generated the funding and acquired the data are often unwilling to give the data to a central storage location not under

Ausbau der Basis IT-Infrastruktur für die Forschung

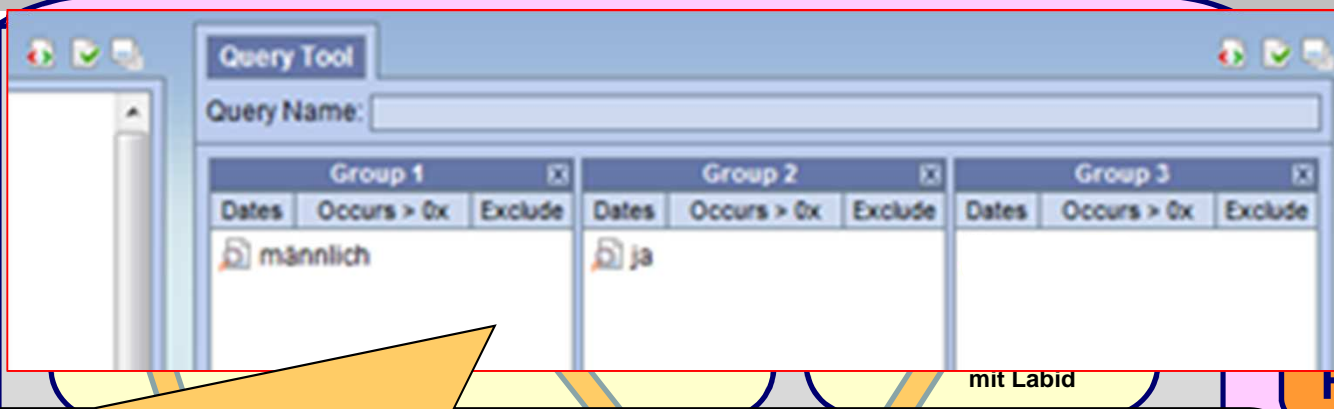
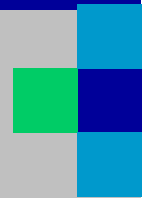
zur Unterstützung

- der Feasibility / Pre-Research Phase
- der Patientenrekrutierung

Kernkomponenten einer Standort-bezogenen Forschungsinfrastruktur

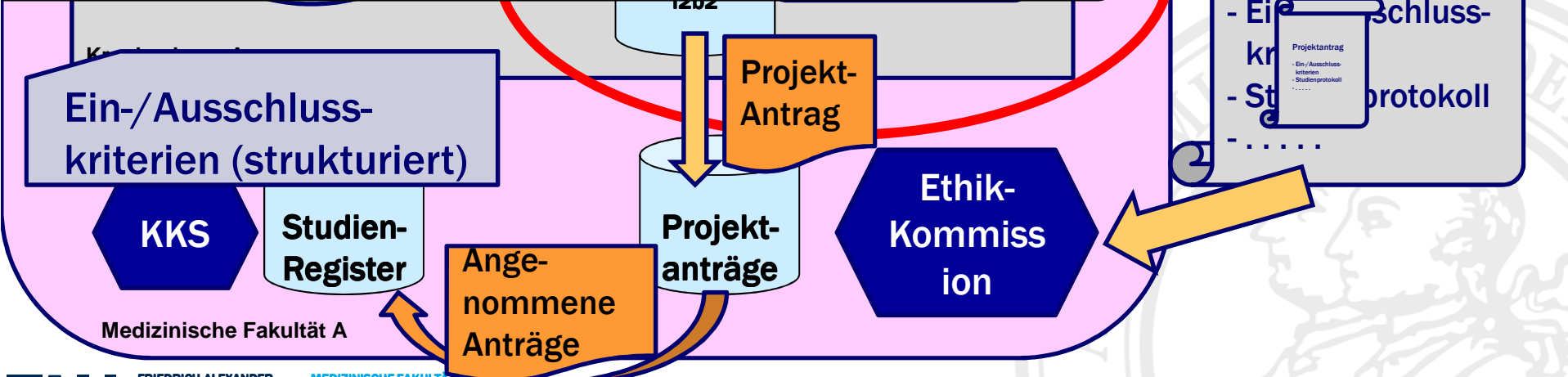


Feasibility Phase / Pre-Research Phase



Identifikation von Patientenkohorten

Alle männlichen Patienten unserer Nierenkohorte die einen Herzinfarkt hatten



KIS-basierte Patientenrekrutierung

■ BMBF Projekt KISREK (2011/2012)

■ Partner

- Universitätsklinikum Düsseldorf
- Universitätsklinikum Erlangen
- Universitätsklinikum Gießen
- Universitätsklinikum Heidelberg
- **Universitätsklinikum Münster**
- TMF e.V.

KIS-basierte Patientenrekrutierung

■ BMBF Projekt KISREK (2011/2012)

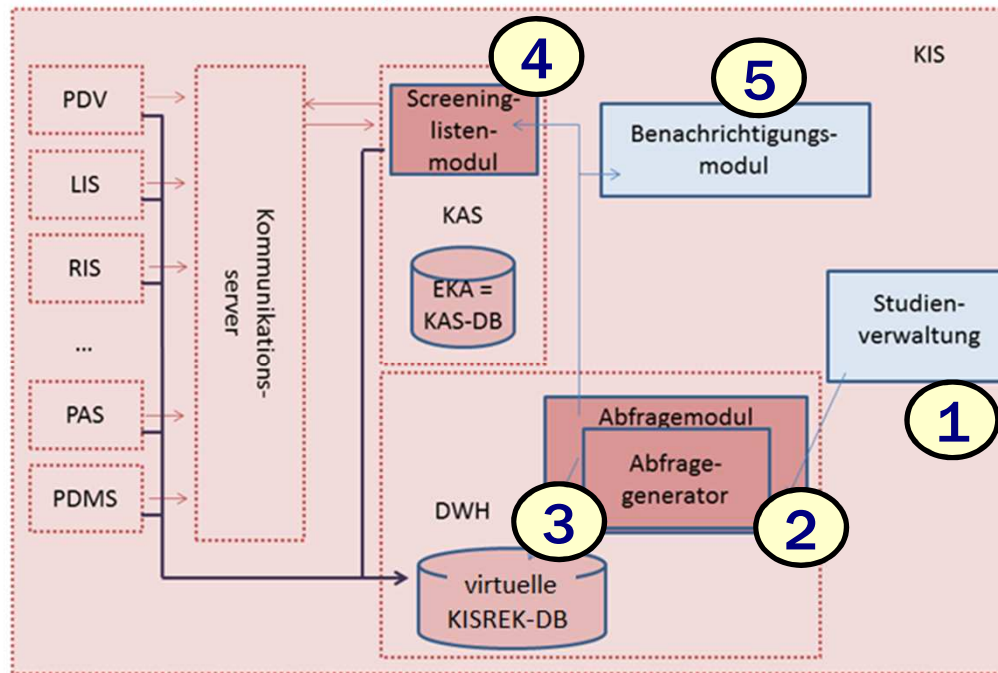
■ Partner

- Universitätsklinikum Düsseldorf
- Universitätsklinikum Erlangen
- **Universitätsklinikum Gießen**
- Universitätsklinikum Heidelberg
- **Universitätsklinikum Münster**
- TMF e.V.

KIS-basierte Patientenrekrutierung

generische KISREK-Architektur

Implementierungsvariante 2c der generischen KISREK-ARCH



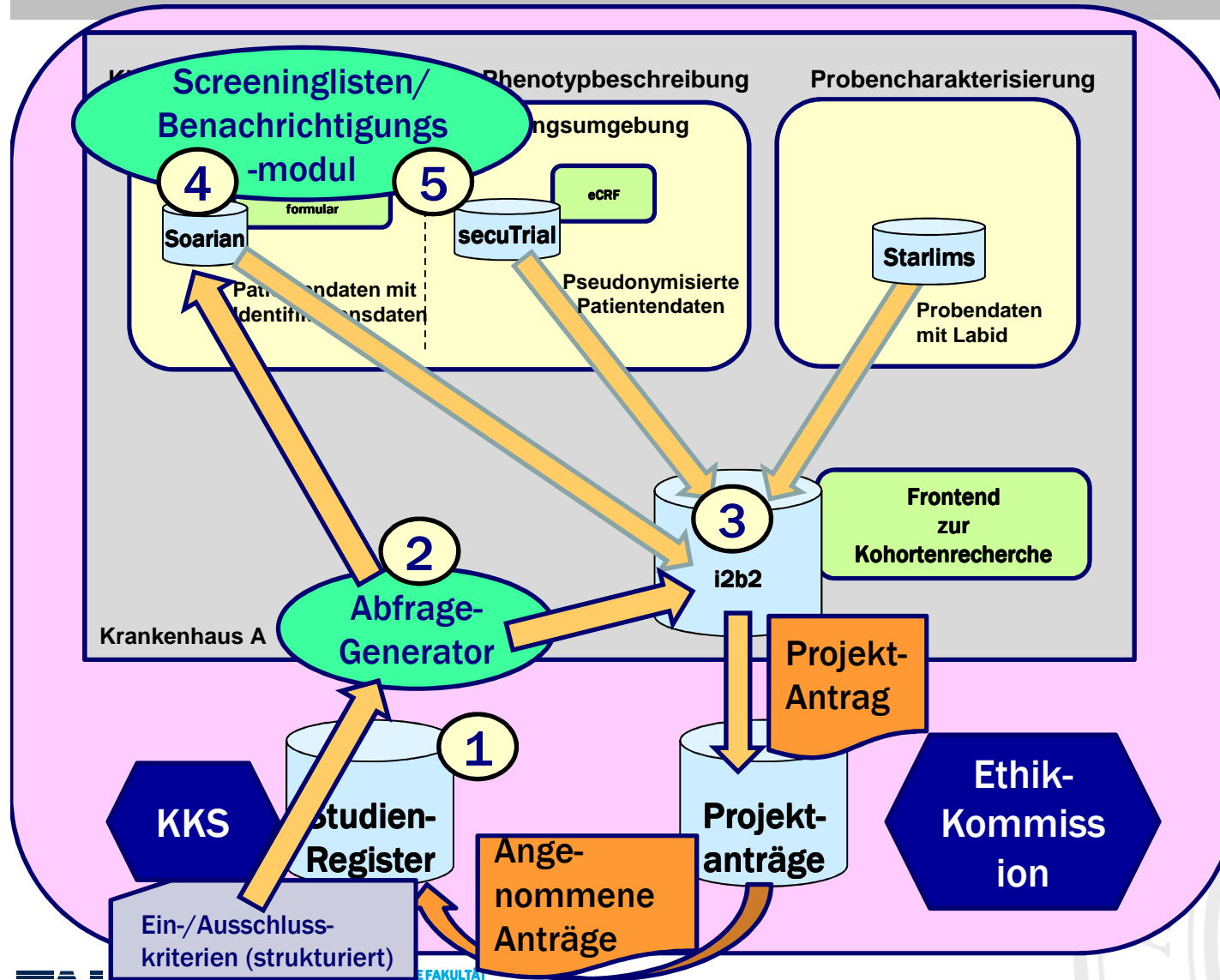
Felix Köpcke // "KIS-basierte Unterstützung der Patientenrekrutierung in klinischen Studien" // 14.12.2011



WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER



KIS-basierte Patientenrekrutierung



Nutzung von Daten der EKA für die Forschung

- Optimierung der Feasibility Phase / Pre-Research Phase
- KIS-basierte Patientenrekrutierung
 - Nutzung von Daten aus der Krankenversorgung
 - Für klinische Studien
- Zum Aufbau von Kohorten (auch Standort-Übergreifend)
 - Versorgungsforschung
- Optimierung der Erkennung unerwünschter Nebenwirkungen
 - Wissensverarbeitende Funktionen (Pharmakovigilanz)

Die Visionen der Pharmaindustrie

EHR4CR Project 2011-2014



- **Electronic Health Records for Clinical Research**

- enhance clinical trials and pharmacovigilance
- through re-use of routine clinical data

- **Approach**

- develop technical platform
- demonstrate functionality at pilot hospitals
- establish business model for long-term sustainability

- **Collaboration of Pharma & Academics**



EHR4CR Consortium

The image displays a collection of logos for the EHR4CR Consortium. The logos are arranged in a grid-like fashion. Several logos are circled in red: TMF, Friedrich-Alexander-Universität Erlangen-Nürnberg, HEINRICH HEINE UNIVERSITÄT DÜSSELDORF, and WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER. The logo for the Friedrich-Alexander-Universität Erlangen-Nürnberg is the largest and most prominent among the circled ones.

Die Visionen der Pharmaindustrie

EHR4CR Project 2011-2014



- **Electronic Health Records for Clinical Research**

- enhance clinical trials and pharmacovigilance
- through re-use of routine clinical data

- **Approach**

- develop technical platform
- demonstrate functionality at pilot hospitals
- establish business model for long-term sustainability

- **Key challenges**

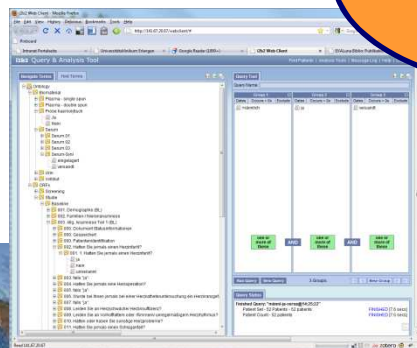
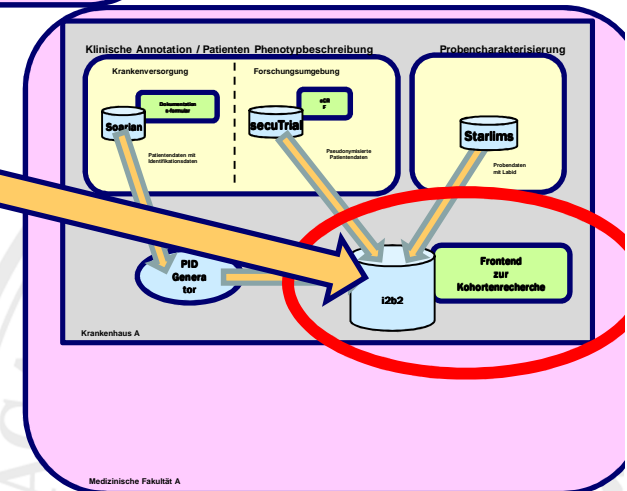
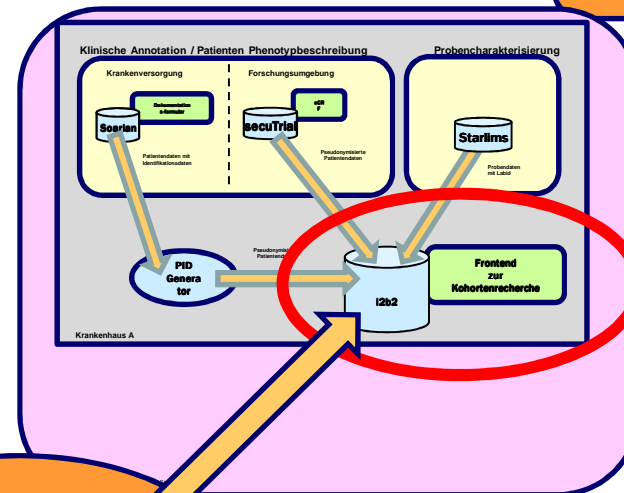
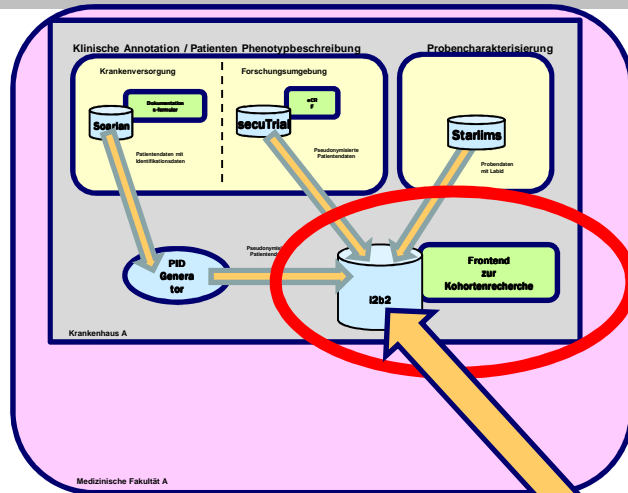
- distributed datamart & query infrastructure
- semantic interoperability
- data protection & ownership

- **Collaboration of Pharma & Academics**



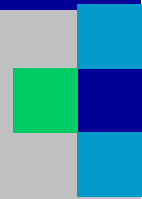
Feasibility Phase / Pre-Research Phase der Pharma Industrie

Identifikation von
Patientenkohorten



MEDIZINISCHE FAKULTÄT

IT-Infrastruktur als Schlüsseltechnologie für die Forschung in der Medizin



IT-Infrastruktur als Schlüsseltechnologie für die Forschung in der Medizin

- Verknüpfung molekularer/genetischer Daten mit Phenotyp Daten
 - DNA-Sequenzierung,
 - Whole Genom Mapping,
 - Next Generation Sequencing

Ohne eine leistungsfähige zentrale IT-Infrastruktur
lässt sich keine innovative klinische / translationale
Forschung in der Medizin betreiben

**Ohne eine leistungsfähige zentrale IT-Infrastruktur
lässt sich keine innovative klinische / translationale
Forschung in der Medizin betreiben**

■ **Kernkomponenten heute**

- GCP-zertifizierte (RDE) Studiendatenbank
- Biobankmanagementsystem
- Standortbezogenes Studienregister

■ **Notwendige Ausbaustufen**

- Integrationsdatenbank / Forschungs Data Warehouse
- Single Source Konzepte
 - IT-Unterstützung in Feasibility Phase und
 - Patienten-Rekrutierung und
 - Studiendurchführung

**Ohne eine leistungsfähige zentrale IT-Infrastruktur
lässt sich keine innovative klinische / translationale
Forschung in der Medizin betreiben**

- **Kernkomponenten heute**

- **GCP-zertifizierte (RDE) Studiendatenbank**

Ziel:

- **Eine in sich geschlossene
Workflow-unterstützte IT-Infrastruktur
die alle Phasen der medizinischen Forschung
durchgängig unterstützt**

- **Patienten-Rekrutierung und**
- **Studiendurchführung**

Ohne eine leistungsfähige zentrale IT-Infrastruktur lässt sich keine innovative klinische / translationale

Forschung in der Medizin betreiben

- Kernkomponenten

- GCP-zertifizierte

Symposium

Perspektiven der
Medizinischen Informatik,
Biometrie und Epidemiologie

W...tur
die alle Phasen der medizinischen Forschung
durchgängig unterstützt

- Patienten-Rekrutierung und
- Studiendurchführung

ulli.prokosch@uk-erlangen.de

Vielen Dank an

Prof. Dudeck,

Dr. Bödeker,

und

Dr. Udo Altmann

Dr. Rainer Röhrig

**Vielen Dank
für Ihre Aufmerksamkeit**