

Daten, Standards und IT-Architektur der Hannover Unified Biobank (HUB)

11. Sitzung der GMDS Arbeitsgruppe
21.11.2014, Erlangen

Markus Kersting

Kersting.Markus@mh-hannover.de



Medizinische Hochschule
Hannover

Themen

- Die Hannover Unified Biobank (HUB)
- Die Biobank IT-Systeme der HUB
- SPREC
- ID-Management
- Klinischen Annotation
- Big (Data) Picture

Medizinische Hochschule Hannover (MHH)

- **Über 9000 Mitarbeiter**
- **110 Kliniken und Institute**
- Medizinische Fakultät, über **3000 Studenten**
- Enge Kooperation mit den beiden Universitäten in Hannover, sowie Fraunhofer ITEM und HZI
- **Biobank im Aufbau seit 2012**



Forschung

- 2 Deutsche Gesundheitszentren (DZL, DZIF)
- 2 Exzellenz Cluster
- 7 SFBs
- 10 Forschergruppen
- 5 „Kompetenznetzwerke“
- 1 „IFB“ für Transplantation

Die Biobank im Clinical Research Center (CRC)



Umzug 2014 erfolgt

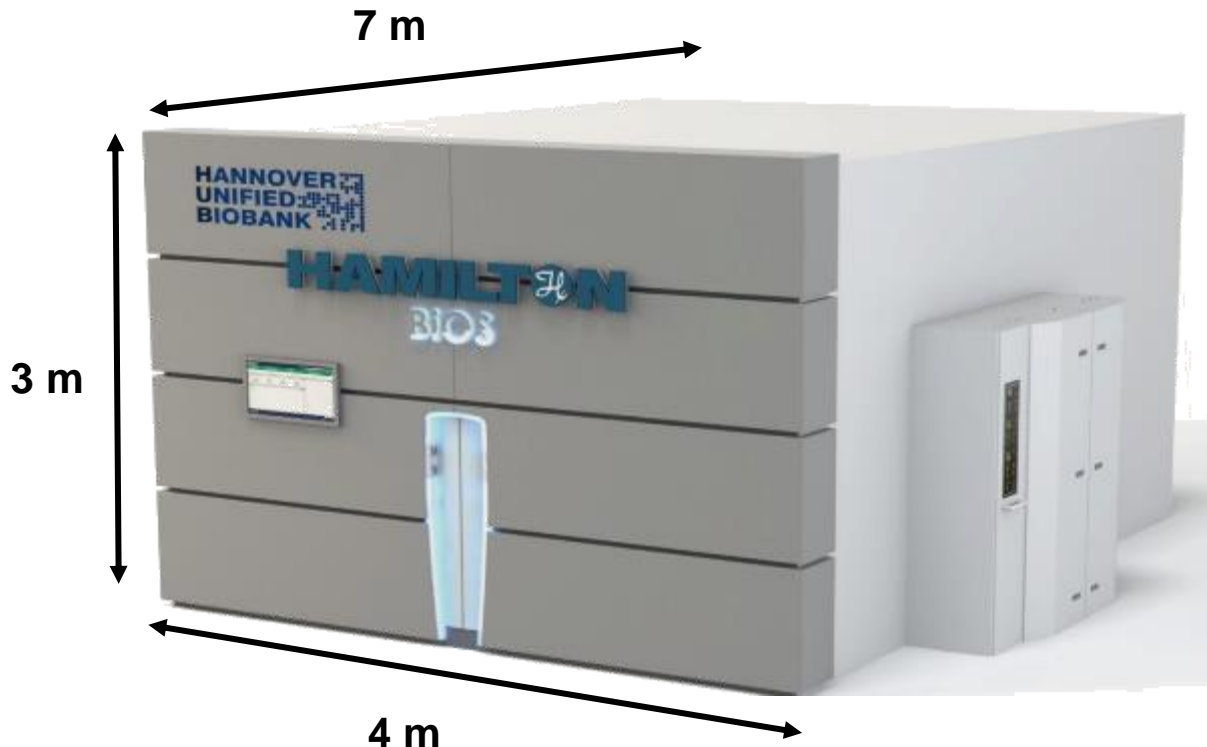
ca. 6800 qm NF

Fraunhofer ITEM
HZI Braunschweig
MHH



Central Storage Room

- 400m²
- Now: Manual Storage: 7 freezer, 4 tanks
- **N₂-Tanks**, Maximum: n=21
- 120.000 1.0ml tubes per tank
- Mostly used for storage of liquid samples



Hamilton BIOS

- Stores multiple labware
- Sample handling and storage at -80°C
- Capacities 600.000 1,0ml tubes
- Start 2015

Organisation der HUB

- **Zentrale Leistungseinheit der MHH**
 - Stabsstelle des Präsidiums (analog zum Rechenzentrum)
 - **Integration aller** Biobank-relevanten IT-Strukturen an der MHH

- **Aktivitäten**
 - Ziel: Eine Biobank für die gesamte MHH, zentrale Lagerung
 - Broad Consent von Ethikrat genehmigt und eingeführt
 - Interne Verrechnung von Biobankleistungen
 - Enge **Kooperation mit Rechenzentrum**
(Zentrales Datawarehouse und ID-Management/Datenschutz)
 - Enge **Kooperation** mit den Partnern im CRC (HZI, NK, ITEM)



Projekte und Sammlungen der HUB 1/2

- **DZL** (Deutsches Zentrum für Lungenforschung) 75.000 Probe
- KIRA (Asthma bei Kindern Studie) 2500 Proben
- CF-DNA Biobank (Lungenfibrose bei Kindern Studie) i.p.
- Zystische Fibrose (Familien- und Zwillingsstudien zu CF) und 11.000 Proben
- PH-BB (Pulmonale Hypertonie) 500 Proben
- **Capnetz** (Kompetenznetz für Ambulant erworbene Pneumonien; CAP) 48.000 Proben
- ABACOPD (Multizentrische Studie zu chronischen obstruktiven Lungenkrankheiten) 600 Proben
- Lung-Tx Lungentransplantation Biobank (Lung Transplant Clinical Specimen and Organ Bank) 14.000 Proben
- **DZIF** (Deutsches Zentrum für Infektionsforschung) i.p.
- DZIF Tx-Kohorte (multizentrische Transplantationskohorte im Bereich KM) i.p.

Projekte und Sammlungen der HUB 2/2

- **BMBF Probase** (Deutsche Studie zum Prostatakrebs Screening) 12.400 Proben
- Hepnet (nationales Hepatitis Netzwerk) 2.000 Proben
- **EU IMI Combacte** (Europäische multizentrische Studie zu antimikrobiellen Resistenzen und der Verbesserung von Antibiotika Entwicklungen) i.p.; 1000 Proben
- **IFB-Tx** (Integriertes Forschungs-und Behandlungszentrum Transplantationen) 80.000 Proben
- Rebirth - active women (Standort Studie zur Wirkung von regelmäßigen Sport auf Gesundheit) 1.190 Proben
- VIANNA (Implantat-Forschung, Gehörlosigkeit)
- **KMT-Biobank** (Knochenmarktransplantationen) 50.000 Proben
- Pädiatrische Tumorbank, FANCONI (Studie Fanconi-Anämie bei Kindern) 100 Proben

IT-Systeme - Überblick

• IT-Personal

- Stabsstelle Biobank: 3 (2 ½)
- Stabsstelle Rechenzentrum: 130

• Biobank-Management-Software

- MySamples, Fa. MyData (www.mysamples.de)
- Java, MySQL, Lokale Installation, Enterprise-Lizenz
- Alle Server im Hochverfügbarkeits-Rechenzentrum (VM)
- Ab 2015: CentraXX von Kairos

• Weitere Biobank-Systeme

- Sample Registration Tool (Web)
 - <https://www.mh-hannover.de/sample-registration/>
 - Eigenentwicklung, LAMP
- SharePoint (QM, Kunden- und Auftragsmanagement)
- ELPRO (Temperatur / Monitoring)
- Zebra (Etikettendruck)
- Hamilton + weitere (Roboter, Logger, ...)



IT-Systeme - Historie

- Vor 2012: Lokaler, vereinzelter Einsatz von MySamples an der MHH
- 2012/08: Beschaffung Enterprise-Lizenz für MySamples
 - Freie Nutzung für alle > Unabhängig von zentraler Lagerung
 - Installation / Serverbetrieb durch Rechenzentrum
 - Inhaltlicher Support durch HUB
 - User-Support
 - Verwaltung der Stammdaten und Probeneigenschaften
 - Schulung von Usern (TA, Dokumentare, Ärzte, ...) n > 200
 - Weiterentwicklung/Customizing von MySamples durch MyData
 - Scanner-Anbindung, Schnellumlagerung, Globale Suche, Protokolle/Berichte, Verwaltung von Merkmalen, Übergabe von Proben an anderen Abteilungen / Organisationen, ...
- 2013: Neue Eigenentwicklung: Sample Registration Tool (COMBACTE)
- 2014/07: Ausschreibung/Beschaffung neues Biobank-System
- 2014/11: Zuschlag CentraXX

SPREC

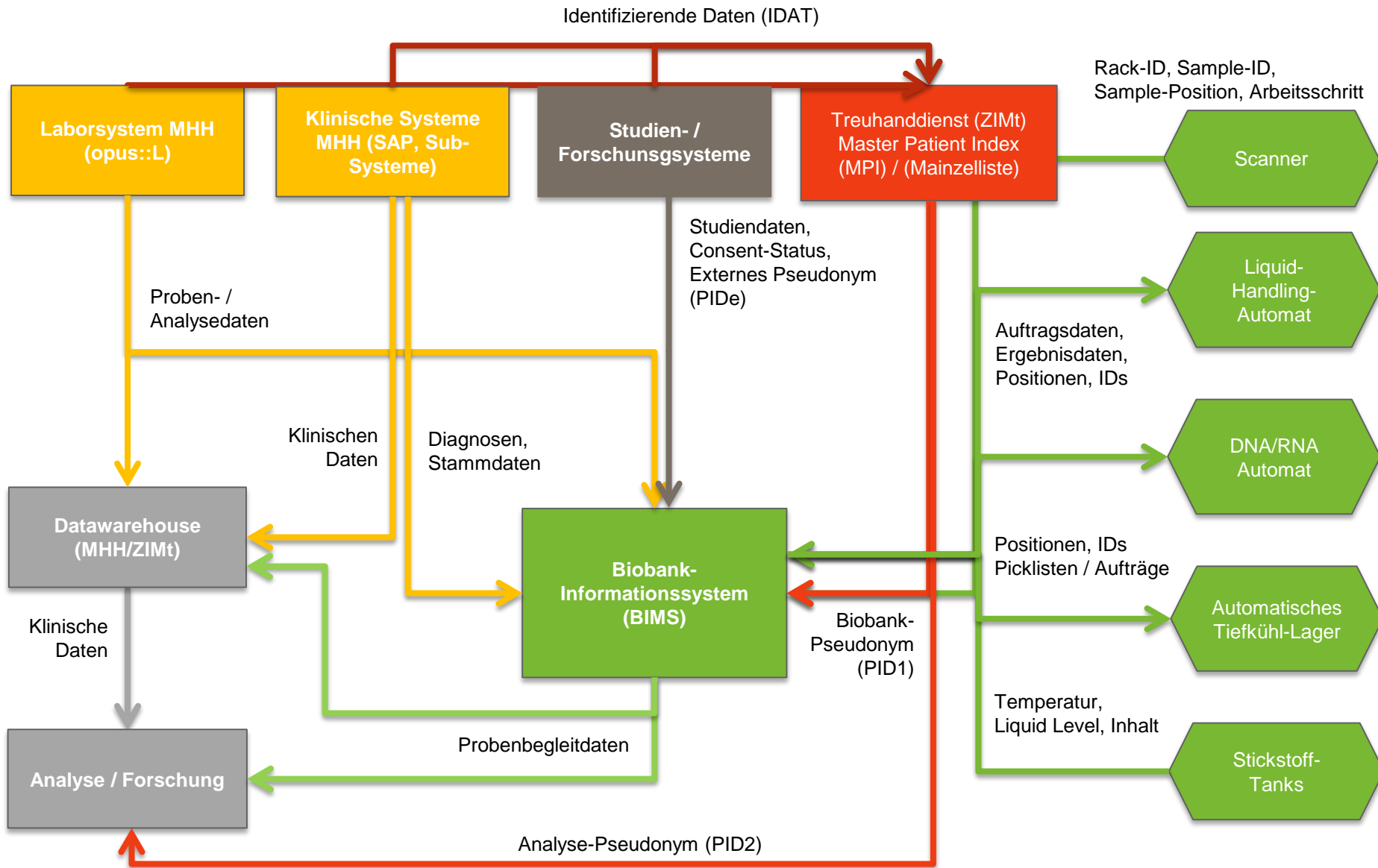
- Sample PREanalytical Code (SPREC)
- Nutzung nur als Probeneigenschaft, ohne Logik
- Implementierung mit Anpassungen der Felder **SampleType** & PrimaryContainer
- Zusammenlegung der **SampleType**-Kataloge von SPREC-Fluid & SPREC-Solid in einem Feld
- **Anpassung des SampleType - Katalogs**
 - z. B. „Microorganism“ mit SPREC-Code “ZZZ(mo)”
 - ZZZ = Other > abwärtskompatibel ohne Anhang/Klammern
 - Ziel: **Zentraler Materialart-Katalog für alle Bioproben**
>>> MIABIS 2.0 !?
- Anpassung der Primärschlüssel/Codes für Eindeutigkeit



Datenschutz / ID-Management

- Aktuelle Lösung
 - Keine identifizierenden Daten in der Biobank
 - Keine direkte Anbindung an klinische Systeme
 - Verknüpfung zur Behandlung über I-Zahl und Fallnummer
 - Pseudonyme werden durch Studien verwaltet
 - Daten- / Probenherausgabe nur an Eigentümer
- Ziele
 - MHH-einheitliche Lösung für Biobank und Datawarehouse
 - Integration der klinischen Systeme (KIS/SAP, Labor, ...)
 - Integration von Studiensystemen
 - Berücksichtigung/Migration lokaler Forschungsdatenbanken
 - Berücksichtigung lokale vs. zentrale Lagerung
 - Zentrale Datenhaltung
 - Implementierung des TMF-Datenschutzkonzeptes
 - Benutzerfreundlichkeit, Automatisierung





Klinische Annotation - Systeme

- **Keine einheitliche Lösung** derzeit > auch nicht in Sicht !?!
- **„Große“ Projekte**
 - Eigenentwicklungen (PHP/MySQL, FileMaker, Access)
 - SecuTrial, Marvin, OpenClinica, (ITEM: ClinBase)
 - Proben/QM-Daten > Biobank-System
 - Alle Daten > Externes, Studienbezogenes DataWareHouse (z.B. DZL)
- **„Kleine“ Projekte**
 - Access, FileMaker, Papier, „nix“ ...
 - Kerndaten > Biobank-System (MySamples)
- **Routinedaten**
 - Primärsysteme SAP (Administration) und opus::L (Labor) und weitere (OP, Intensiv, Zentrales Patientenaktenarchiv (PDF) ...)
 - Alle Daten > Datawarehouse
 - Kerndaten > Biobank-System

Klinische Annotation - Status

- Import über Schnittstellen?
 - Derzeit nur manuell per CSV
 - Bei größeren Einlagerungen oder Altdaten-Migration
- Redundante Eingabe
 - Teilweise ja
 - Vor allem Kerndaten (Diagnose, PIDs)
- Kerndatensatz vorhanden?
 - Ja, per Richtlinie > Teil der Schulung
 - Kerndaten > Biobank-System (MySamples)
 - Neudefinition geplant (XSD/Message)
- Biobank-Query-Tool?
 - (Globale) Suchfunktionen vorhanden für Proben- und Kerndaten in MySamples
 - Query-Tool für klinische Daten geplant als Modul des DataWarehouse
 - Policies & Zugriffsmöglichkeiten für Bioproben und DataWarehouse in Arbeit
 - Geplante Technologien: Query mit SharePoint / MS SQL-Server als Backend
 - Feasibility Queries über Kerndaten und ggf. Metadaten : CentraXX

2 Referenz globaler Merkmale

2.1 Basiseigenschaften in MySamples

Name	Kategorie	Beschreibung/Description deutsch [englisch]	Beispiel	Datentyp
Kennzeichnung	Sample/Barcode	Im Feld „Kennzeichnung“ einer Probe muss eine eindeutige Kennung hinterlegt werden. Hierbei handelt es sich um die eindeutig Nummer eines Röhrchens. Früher waren hier die sogenannten MySamples Nummern eingetragen. Zukünftig soll in dieses Feld ausschließlich der auf dem Röhrchchen aufgedruckte 2D-Barcode per Scan eingegeben werden.	NU01492253	Alphanum
Standort	Position	Beschreibt die oberste Ebene einer Positionsangabe. Hilfestellung zum Anlegen sind voller Lagerorte, Lager, Gestecke etc. für Ihre Organisation erhalten Sie von der Biobank	Hauptlager, München, MHC, 808 50-4711	Auswahl/Vorgabe

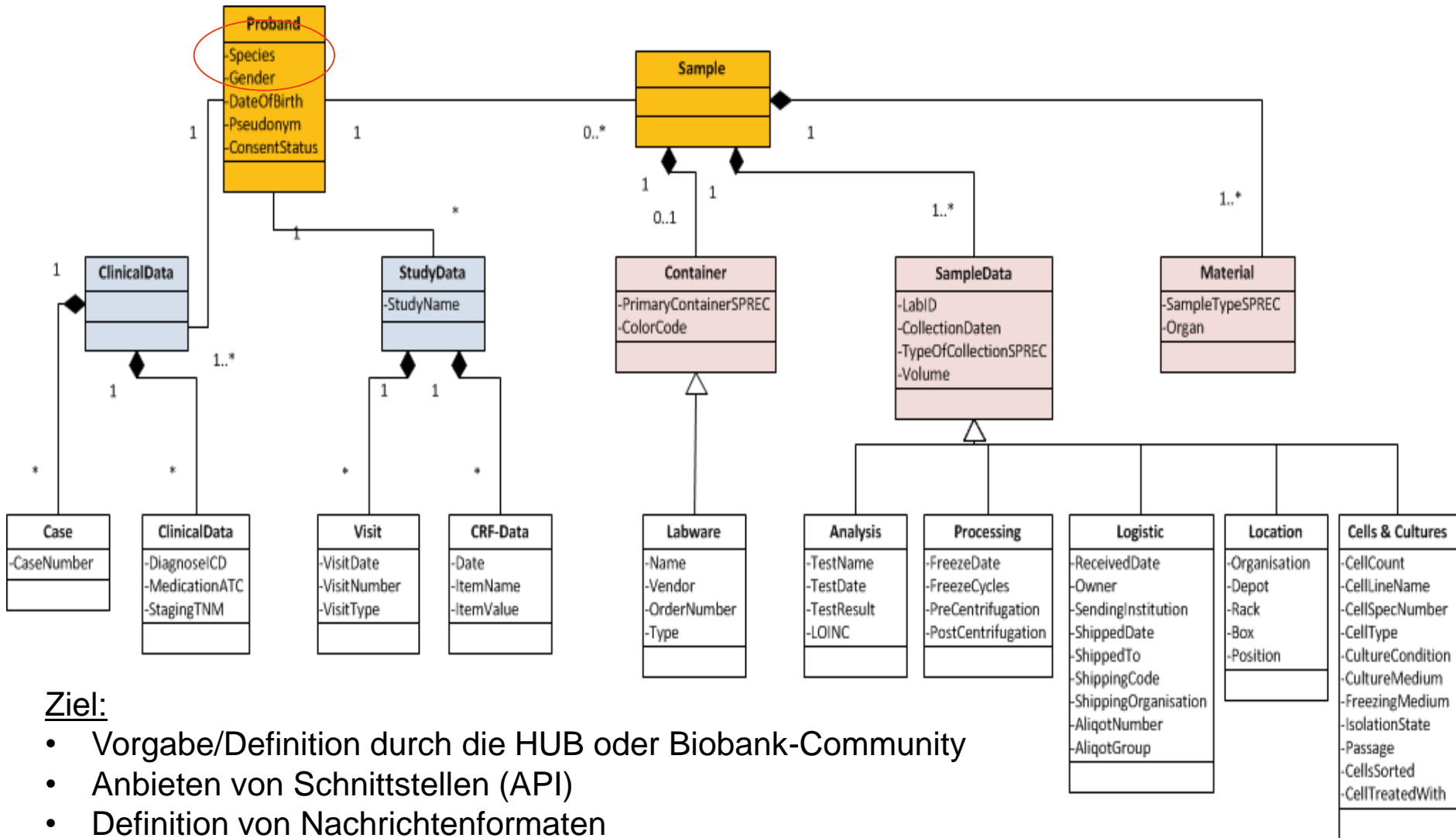
2.2 Pflichtfelder für die Einlagerung in die Biobank

Name	Kategorie	Beschreibung/Description deutsch [englisch]	Beispiel	Datentyp
SampleType	Sample/SPREC	Material / Probenart: Die Angabe erfolgt gemäß der Probenarten nach SPREC. [Type of Sample according to SPREC]	Blood, Solid Tissue	Auswahl (Alphanum)
PrimaryContainer	Sample/SPREC	Art des primären Containers nach SPREC. [Type of Primary Container according to SPREC]. (Pflichtangabe für "fluids")	Additives, Lithium heparin, Sodium EDTA	Auswahl (Alphanum)
TypeOfCollection	Sample/SPREC	Art der Probengewinnung nach SPREC (Pflichtangabe für „solids“)	Biopsy	Auswahl (Alphanum)
CollectionDate	Sample	Datum (Uhrzeit) der Abnahme / Erhebung der Probe [date/time of sample collection]	19.03.2010 22.04.2012 13:32	Date
PatientID_Department	Patient	Spezifische Patienten-ID oder Labornummer aus einem Abteilungs- oder Studiensystem. Diese Nummer sollte nicht sprechend, immer angefüllt und für jeden Patienten innerhalb der jeweiligen Abteilung eindeutig sein. [Patient ID / Lab ID that is used in the department]	ZK34782957	Alphanum
Gender	Patient	Geschlecht des Patienten [male /m/, female /f/] [gender of the patient/subject]	m f	Auswahl (1 Zeichen)
DepartmentID	Department	Eindeutige Kennung einer Abteilung bzw. der Ursprungsorganisation einer Probe, z. B. „KMTA“. Die Kennung wird von der Biobank vergeben, wird für alle Proben abteilungsspezifisch voreingestellt und muss nicht manuell erfasst werden. Die Kennung sowie deren Einrichtung für Ihre Proben ertragen Sie bitte in der Biobank. [Short ID of the department that provides the Sample]	KMTA	Alphanum

MIABIS 2.0

- Bereitstellung der „Core Components“?
 - Ja, ist möglich, Informationen vorhanden
 - Derzeit manuelle Eingabe z.B. im BBRI
- Bereitstellung „Additional Components“?
 - Teilweise, aber eher nicht geplant
- Planung
 - Ausbau/Definition Kerndatensatz + „Registerdatensatz“ (MIABIS)

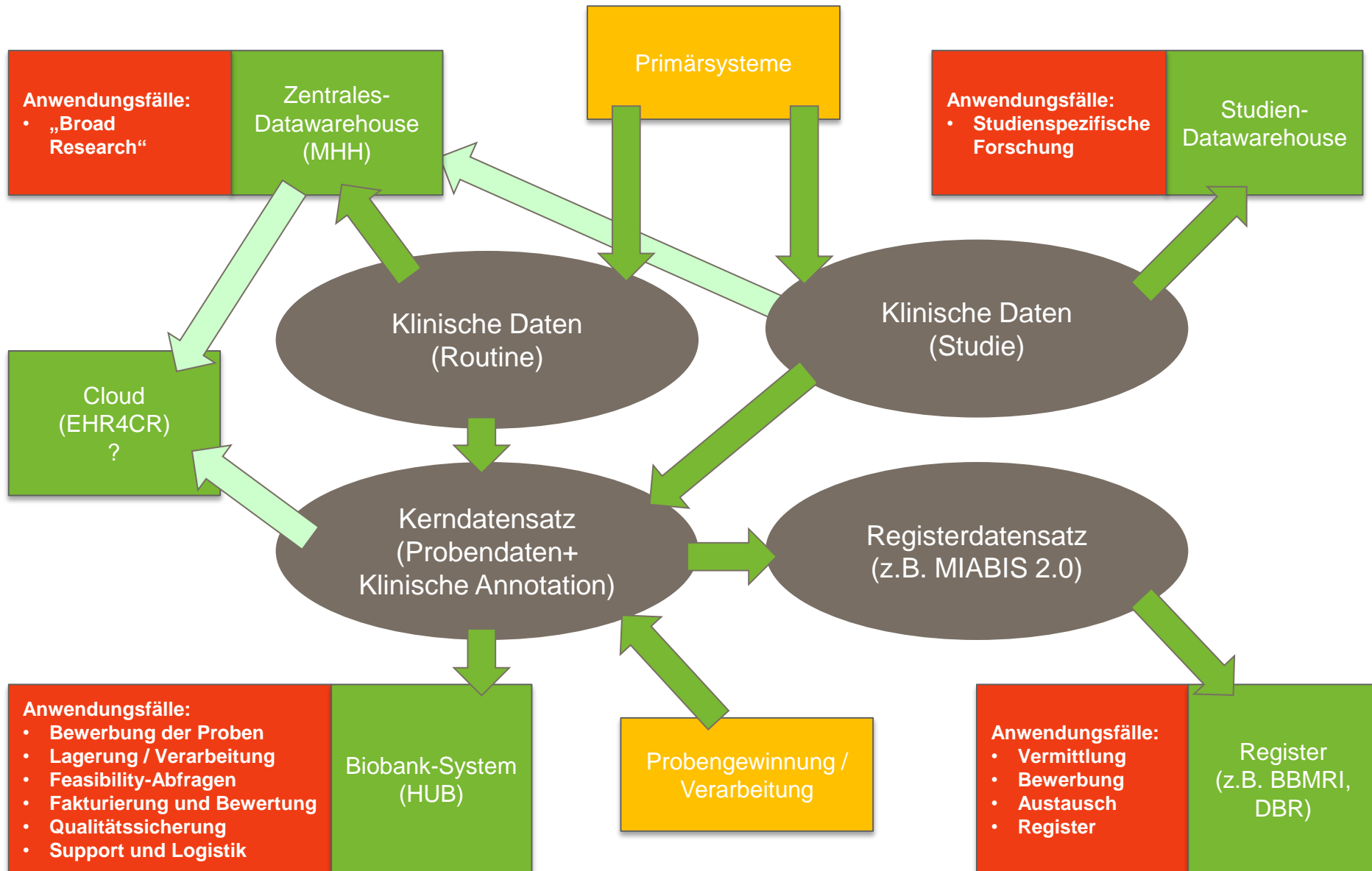
Kerndatensatz, Ausgangsbasis



Ziel:

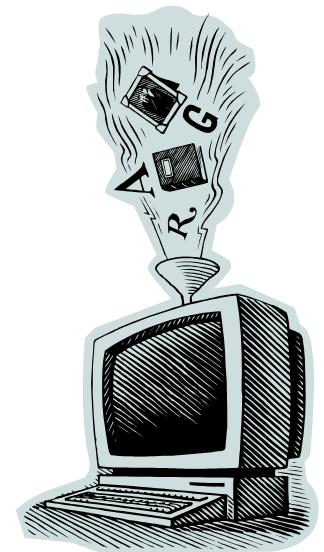
- Vorgabe/Definition durch die HUB oder Biobank-Community
- Anbieten von Schnittstellen (API)
- Definition von Nachrichtenformaten

The Big (Data) Picture...



Fazit

- Klinische Daten gehören in das (ein) **DataWarehouse**
- Ein **gemeinsames Konzept** für Datenschutz und Zugriff für DataWarehouse und Biobank ist notwendig.
- Biobank benötigt einen **Kerndatensatz** für ihr operatives Geschäft
- Kerndatensatz ist Proben- und Probandenbezogen
- **Registerdatensatz** ist Untermenge (Aggregat) von Kerndatensatz
- MIABIS ist guter Start für einen „Registerdatensatz“
- Definition eines Kerndatensatzes durch die Community ist wünschenswert
- Verständnis der **Biobank als Dienstleister** mit Schnittstellen!



Ende

*„Wer bewirkt, dass
nur **noch eine Dateneingabe erfolgt**, wo früher
noch **zwei redundante Eingaben** notwendig waren,
der ist größer als jeder Feldherr.“*

Frei nach „dem Alten Fritz“