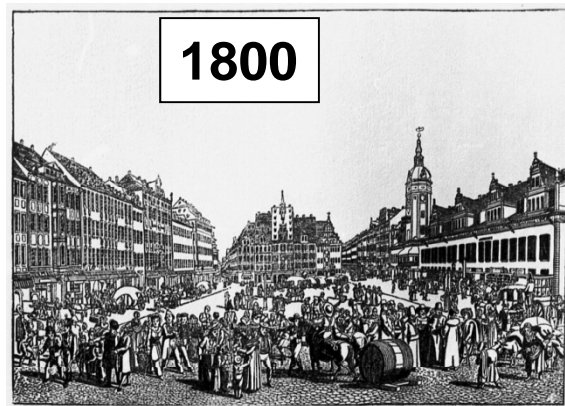

From metabiobanks to translational research platforms: Integrating Big Data through CRIP Tools

4th Munich Biomarker Conference, November 26, 2014

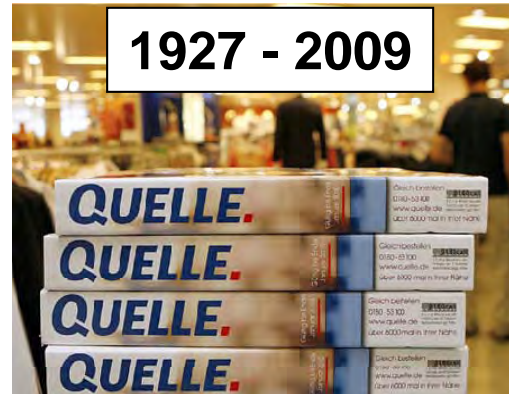
christina.schroeder@izi-bb.fraunhofer.de

Access to goods...



1800

Leipzig Trade Fair



1927 - 2009

Quelle Versandhaus © dpa



2014

“We seek to be Earth’s most customer-centric company for **four** primary customer sets: consumers, sellors, enterprises, and content creators.”

...and to biomedical research-related information

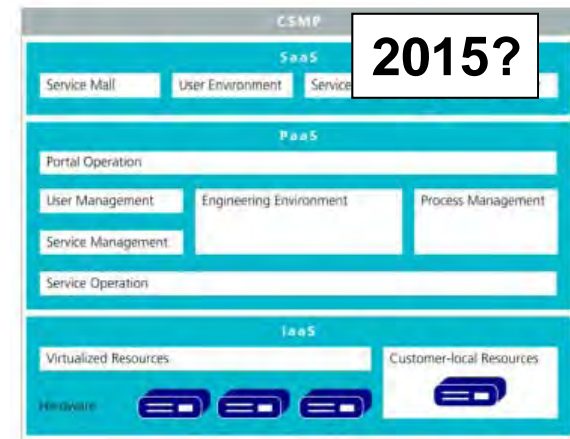


1900

www.nhm-wien.ac



2000



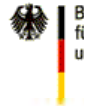
2015?

© Fraunhofer ISST 2011



Projekt-Portal im

p-BioSPRE



- Diagnosis
- Specimen
- Patient
- Annotation
- Informed Consent

+ Blood

dpkk Biobank

- Diagnose
- Therapie
- Nachsorge
- Patient
- Probe
- Institut

ICD-10-Code:
ICD-0-3-Code:
Befallene Lymphkno...:
Digital-rektale Unte...:

- Extracts
- Other

strata pools

Anzahl

Treffer: 0

strata pools

- Your Selection:
- ✗ DiagnosisC00
 - ✗ DiagnosisC64
 - ✗ DiagnosisC64
 - ✗ DiagnosisC65
 - ✗ DiagnosisC81
 - ✗ DiagnosisC91
 - ✗ SpecimenBloc
 - ✗ SpecimenBloc
 - ✗ SpecimenBloc
 - ✗ SpecimenBloc
 - ✗ SpecimenBloc
 - ✗ SpecimenBloc

- [KompNet HIV/AIDS](#)
- [Pediatric Diabetes Biobank](#)
- [PopGen](#)

Metabiobanks / CSMPs

- **Whitepaper:**

http://www.isst.fraunhofer.de/content/dam/isst/de/documents/Technologiefelder/CloudComputing/Fraunhofer-ISST_CSMP-Whitepaper_www.pdf

- **Demonstrators:**

- P-BioSPRE



<https://preview-crip.fraunhofer.de/intern/demo/searchtool/search/p-biospre.cgi>

Weiler G, Schröder C et al.: „p-BioSPRE – An ICT Framework for Transnational Biomaterial Sharing and Access“
ecancer 2014, **8**:401 (www.ecancer.org)

- DPKK-Biobank



<https://preview-crip.fraunhofer.de/intern/demo/dpkk>

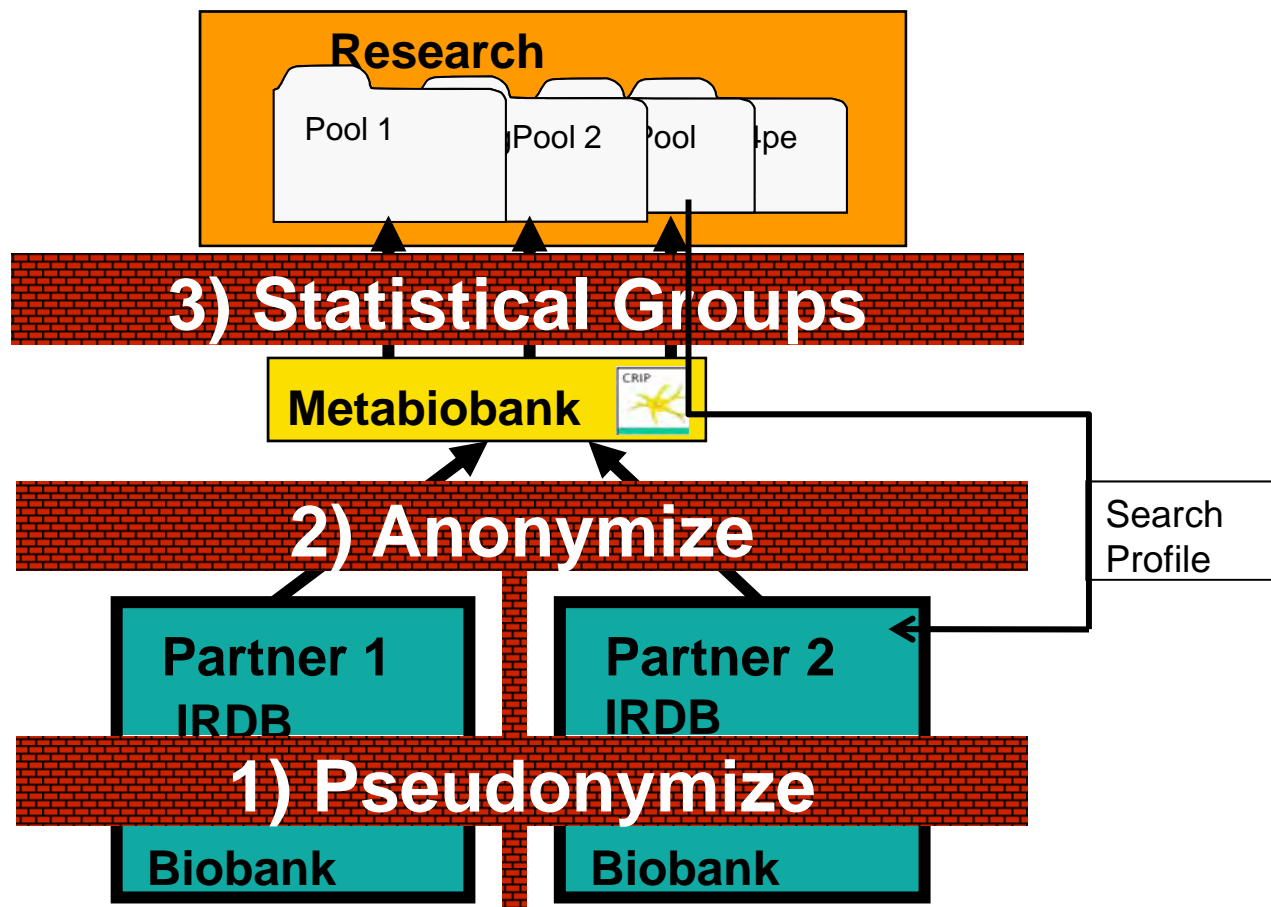
User: user

PW : dpkk

CRIP Privacy Regime



...beispielhaft!



C. Schröder C, Heidtke KR, Zacherl N, Zatloukal K, Taupitz J: *Safeguarding donors' personal rights and biobank autonomy in biobank networks: the CRIP privacy regime*. Cell Tissue Bank **12** (2011), 233 – 240.
<http://www.springerlink.com/content/257x56488841215k/fulltext.pdf>

Features of our metabiobanks

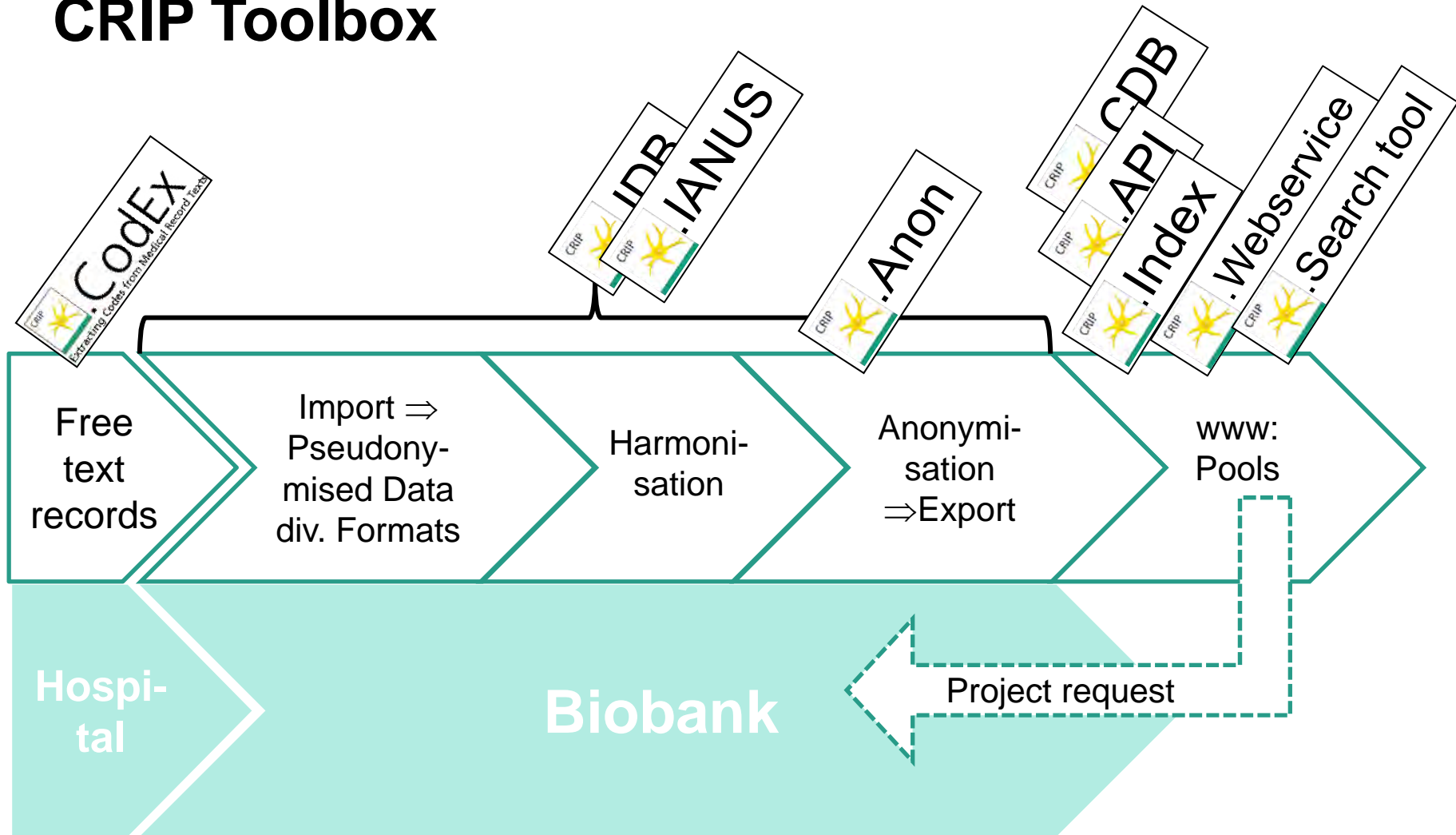
- Web-based parameterized queries (case-by-case)...
- ...across the entire spectrum of diseases...
- ...on anonymized data only: Privacy preserving
- Fulfill BBMRI requirements for networked ICT infrastructure
- Utilize all pre-existing information management systems
- No ICT skills required
- Semi-automated / controlled data export
- Project request:
 - Establish „Biobanking on demand“?

Biobank Partners

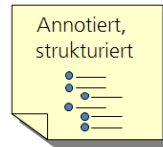
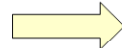
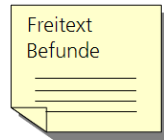
Biobank Data Transfer Agreements with ...

- Charité Universitätsmedizin Berlin
 - Medizinische Universität Graz
 - Klinikum Rechts der Isar / TU München
 - Universitätsklinikum Ulm
 - Zentralinstitut für Seelische Gesundheit, Mannheim
 - Norddeutsches Darmkrebsnetzwerk / UKSH Lübeck
 - Kinderonkologie und PopGen , UKSH Kiel
 - Kompetenznetz HIV/AIDS, Bochum / Köln
 - Universitätsklinikum des Saarlandes, Homburg
 - Deutsches Prostatakarzinom-Konsortium, Düsseldorf,
- CRIP website coupled over web services with...
- biobank-suisse

CRIP Toolbox



<http://www.crip.fraunhofer.de/en/toolbox>



Befundtext:

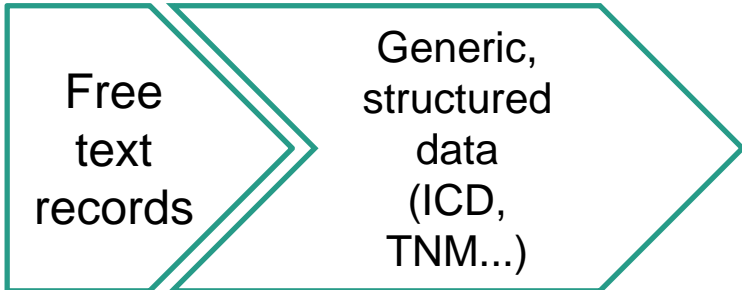
„**SIGMA**. DICKDARMRESEKTAT MIT EXULCERIERTEM, MITTELGRADIG DIFFERENZIERTEM, **TUBULAEREM ADENOKARZINOM** DES DICKDARMES (...)“

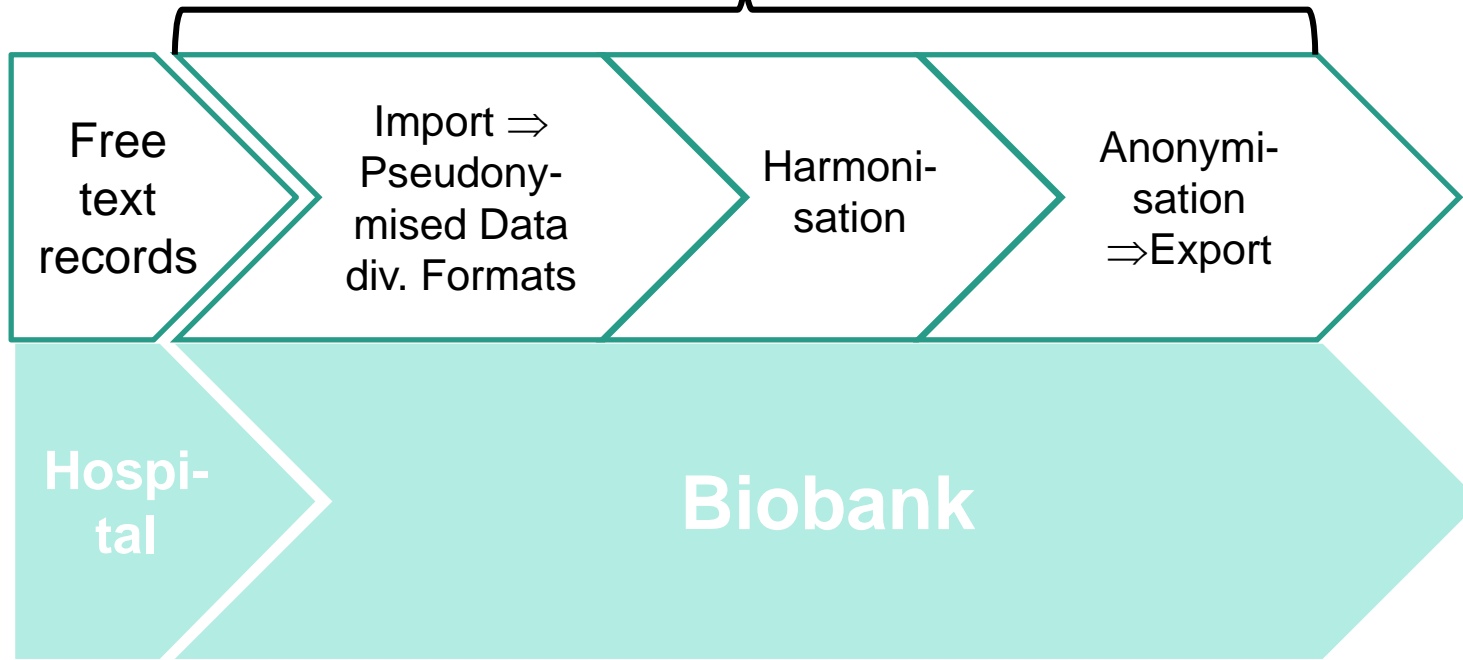
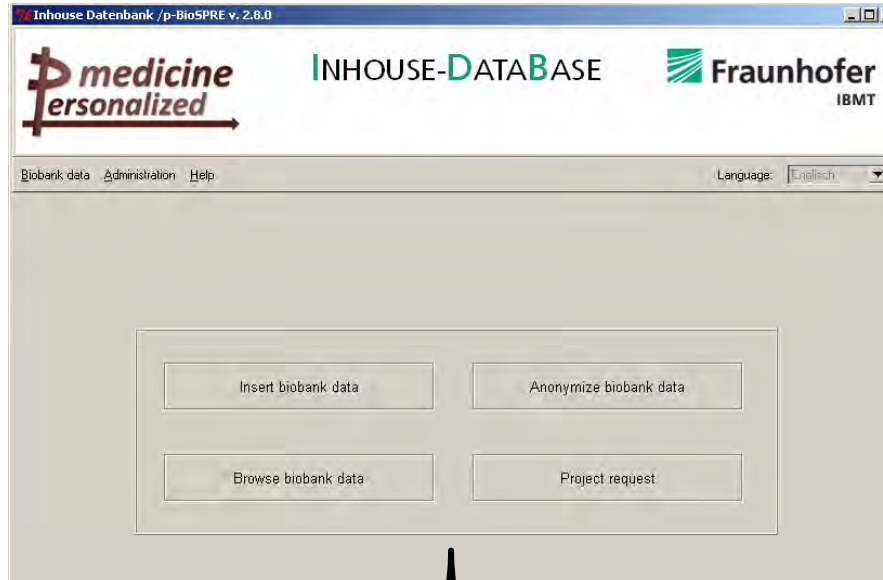


ICD-Codes:

C18.7 (ICD-10)
 8211/3 (ICD-O-3)

- ✓ Import of Coding Guide (e.g. ICD)
 ⇒ no training set and no manual entry of rules required
- ✓ Negations interpreted correctly
- ✓ Quick-and-easy amendment of dictionaries and rules
- ✓ multiple languages
- ✓ under clinical evaluation with Inst. of Pathology, Univ. Bern





- **Harmonize data**
 - Checking of data (text, value, date)
 - Catalog values, calculation of values, regular expressions
 - Pre-processing (prior to data import)
 - Analysis of errors
 - Mapping of parameters

- **Anonymize data**
 - No tracing back, no reference list
 - Singular random codes

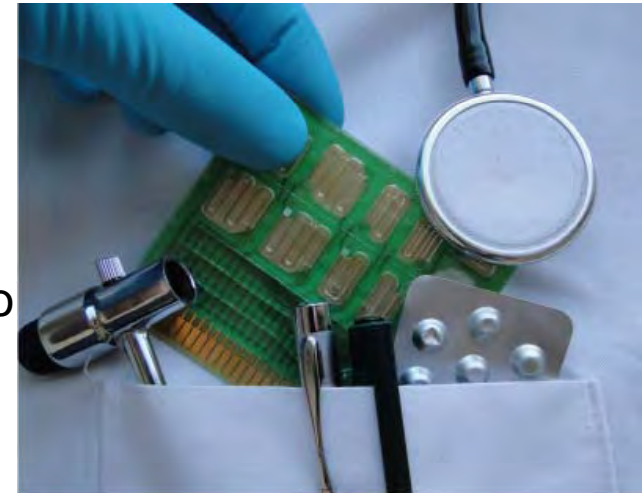
- **Control data export**
 - Review export files; no hidden data transfer

- **Metadata repository**
 - Adopt biobank partner's needs without programming efforts

- **Plug-in external data sources**

Data source: ivD platform

- **Multimarker Lab-On-a-Chip System**
(up to 500 parameters / test)
- Pluggable into existing CRIP architecture and too
- Standards-compliant (CDISC, HL7, LOINC, POCT1-A)



<http://www.ivd-plattform.fraunhofer.de/>

Scientific community



Projekt-Portal im Deutschen Biobanken-Register

P2B2 Home Zugang **Biobankpartner** Projektpartner FAQ DBR

GEFÖRDERT VOM
 Bundesministerium für Bildung und Forschung

BETRIEBEN DURCH
 **Fraunhofer**



dpkk Biobank

DPKK Members / biobanks



Funktionalities DPKK-Biobank

The screenshot displays the CRIP.IANUS v. 0.5 application window. The title bar reads "CRIP.IANUS v. 0.5". The main header features the CRIP logo (a yellow starburst) and the text ". IANUS" in large letters, followed by the Fraunhofer IZI logo and the tagline "Integrate, Access, Navigate, Unite, Stratify". Below the header, there are menu items: "Biobankdaten", "Verwaltung", and "Hilfe". A language dropdown menu is set to "Deutsch".

Overlaid on the main window is a smaller window titled "Institut anlegen". This window contains a form with the following fields:

- Institutsname: [text input]
- Kontakt:**
- Vorname: [text input]
- Nachname: [text input]
- E-Mail: [text input]

Below the form are two buttons: "Speichern" and "Abbruch".

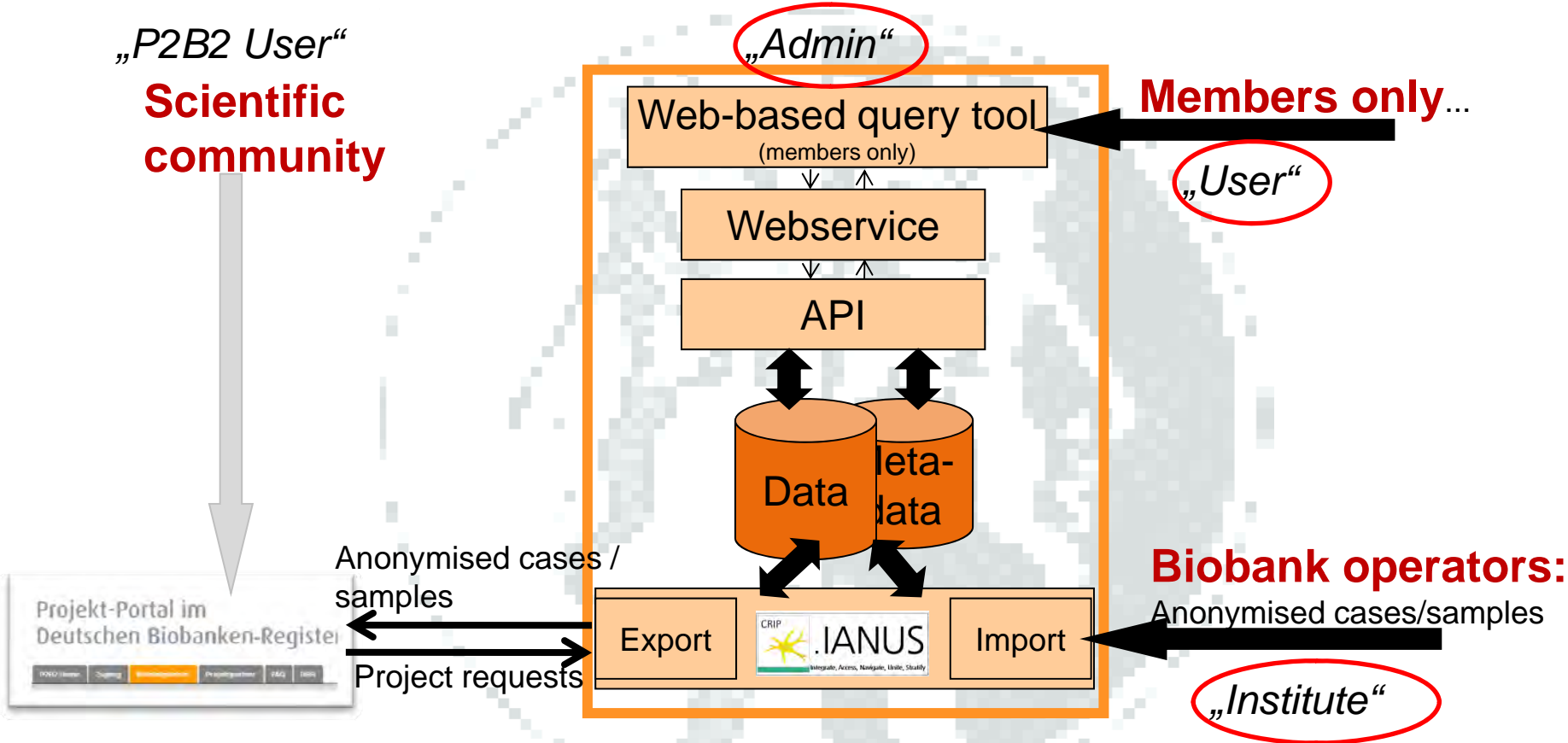
Underneath the form is a section titled "Institut bearbeiten / löschen:" containing a list of radio buttons:

- Uniklinikum des Saarlandes (ID-39)
- Uniklinikum Erlangen (ID-38)

At the bottom of this window are three buttons: "Bearbeiten", "Löschen", and "Abbruch".

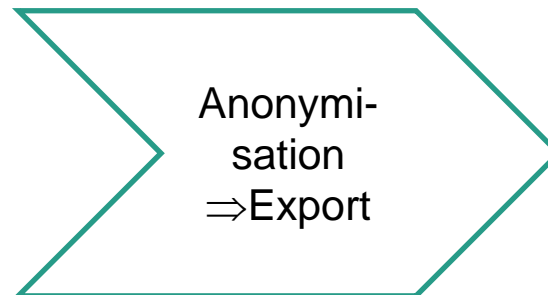
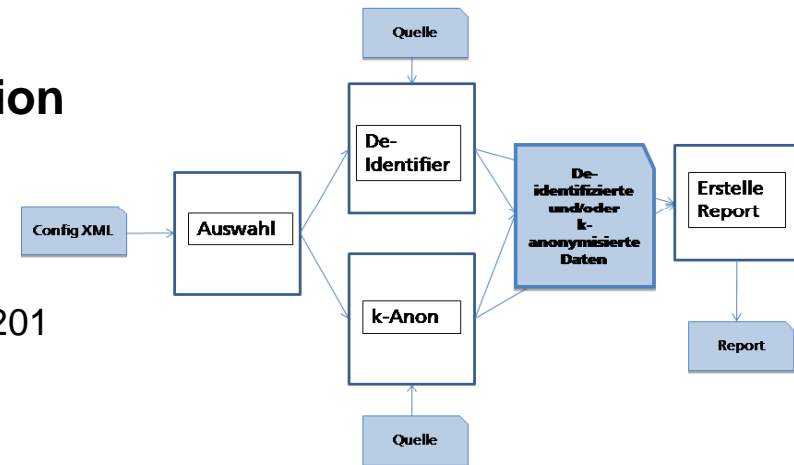
- Integrate
- Access
- Navigate
- Unite
- Stratify

Role Model CRIP.IANUS / DPKK-Biobank

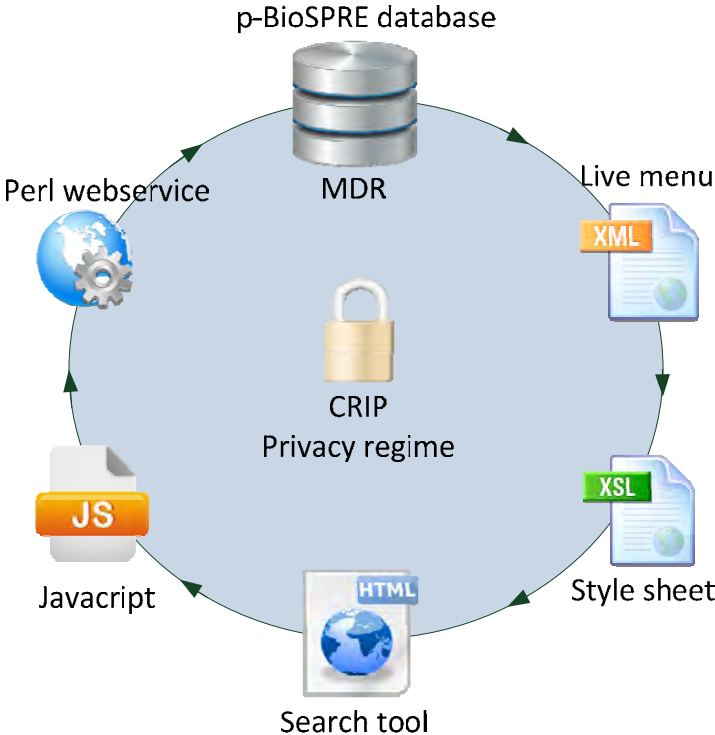


Configure and document data anonymization

- De-Identification: Identifiers replaced by anonymized „brackets“ ...
- ...and / or k-anonymized through <http://www.tmf-ev.de/Produkte/Uebersicht.aspx#P100201>
- Report:
 - Time stamp
 - De-Identifiers: Name(s), status, file
 - K-Anon: Report file



CRIP Search tool



- ✓ Dynamic generation of query interface
- ✓ Corporate design



Technologies

- Server infrastructure: Linux, Apache, PostgreSQL



- Web Standards (HTML, XML, XSL, CSS, Javascript)



Open standards
and technologies

- Programming platform: Perl, Webservices, jQuery



Big Data

■ Velocity

Enhance performance: CRIP.Index

■ Variety

Integrate Diagnoses: ICD, SNOMED, disease ontology

■ Variability

Restrict Minimal Dataset to a minimum!

■ Volume

Molecular biology / Omics data

Process multi-parameter analyses (ivD platform)

<http://www.bigdata.fraunhofer.de/de/big-data.html>

Thank you!



...Matthias Dobkowicz, Rukeia El-Athman, Thomas Finkenwirth, Harry Freitas da Cruz, Oliver Gros, Karsten R. Heidtke, Tim Jüttner, Sascha Kunze

