

Dockerbank

Container-basiertes Deployment von biomedizinischen IT-Lösungen

Block 3: Vorstellung von Containern biomedizinischer IT-Lösungen

Christian Bauer¹, Dr. med. Thomas Ganslandt²

¹ *Universitätsmedizin Göttingen*

² *Universitätsklinikum Erlangen*

Rahmen

- ▶ Untersuchung beim TMF-Deployment-Workshop am 17.01.2016 in Berlin

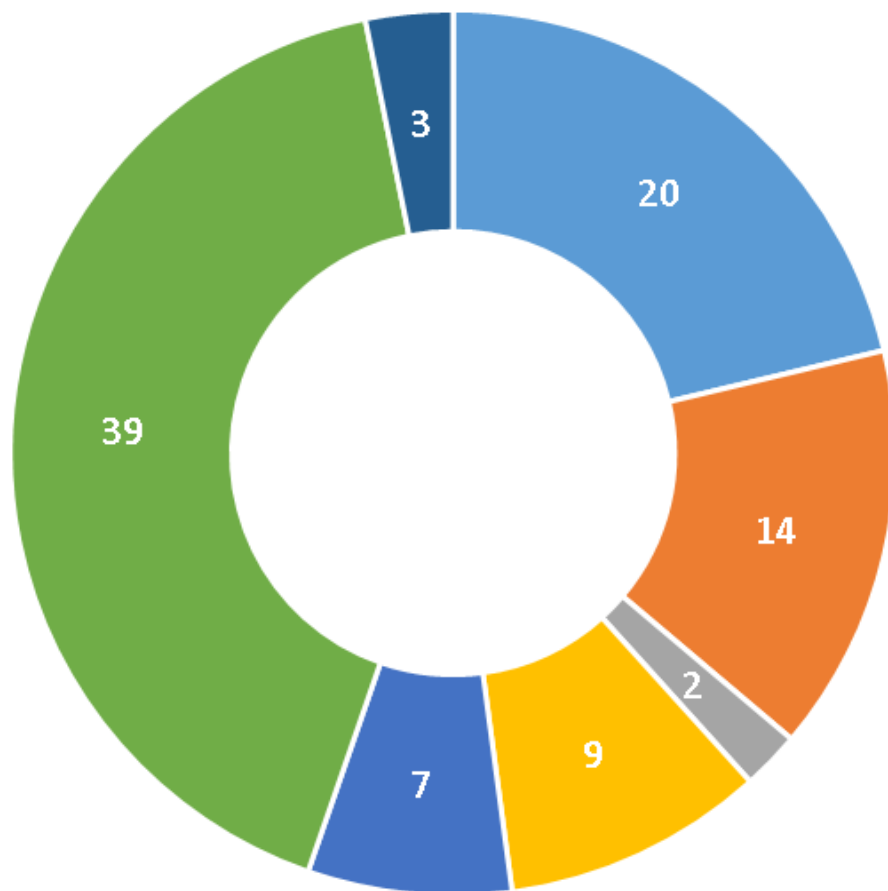
Einschlusskriterien

- ▶ Produktkatalog der TMF
- ▶ (damals) laufende DFG-Anträge
- ▶ internationale Tools mit bekannter Nutzung bei TMF-Mitgliedern

Untersuchte Eigenschaften

- ▶ Typ des Produkts, aktuelle Relevanz
- ▶ Eignung zur Containerisierung
- ▶ mögliche Synergien

Typ der untersuchten Produkte

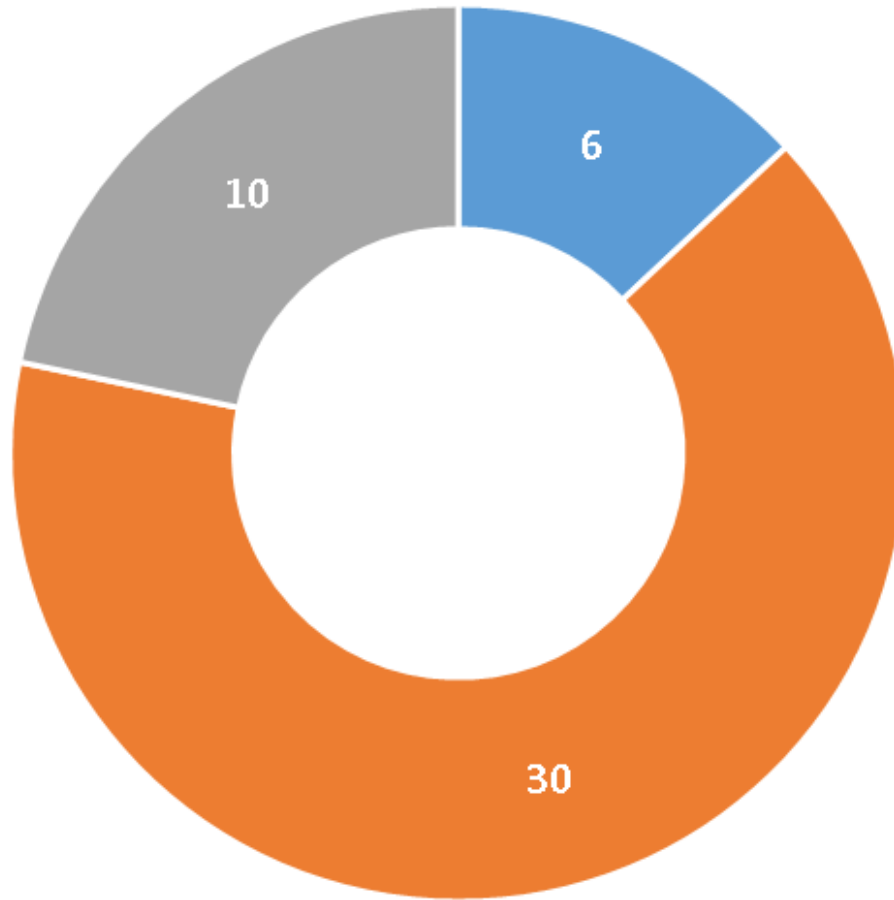


Typ

- Arbeitsmaterial
- Band der TMF-Schriftenreihe
- Beraterung
- Bericht oder Gutachten
- IT-Infrastruktur
- Software
- Sonstiges

} 46 Produkte
näher untersucht

Deploymentrelevanz der IT-Produkte



Deploymentrelevanz ▼

- Betrieb
- Dissemination
- nein

Nicht containerisierbare IT-Produkte

- ▶ Lokal installierbare Software mit GUI
 - ▶ ARX
 - ▶ Quiz-Software

- ▶ Innerhalb von anderer Software zu nutzende Pakete, Bibliotheken
 - ▶ Narrative-Skript für Pharmakovigilanzsoftware
 - ▶ SAS-Makros für klinische Studien
 - ▶ Report-Vorlagen für Pharmakovigilanzsoftware
 - ▶ Software-Bibliothek zu Datenschutzkonzept A

- ▶ Abgelehnte DFG-Anträge
 - ▶ tranSUSTAIN

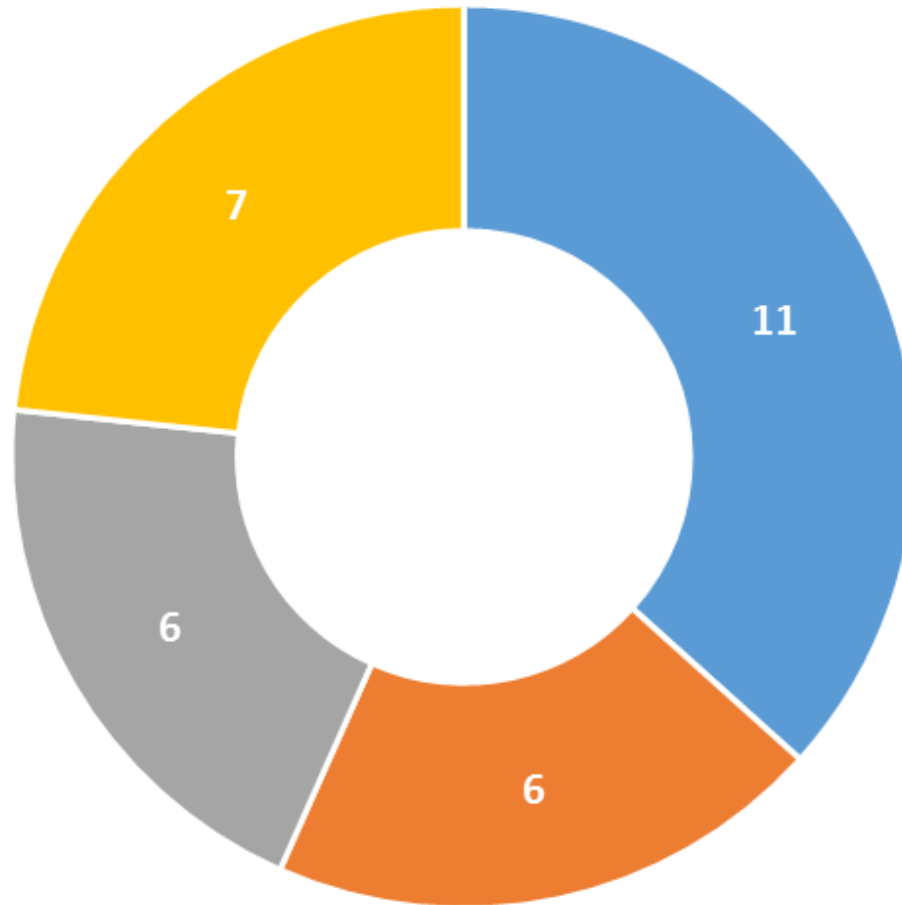
- ▶ Veraltete Produkte
 - ▶ PID-Service
 - ▶ Projektportal im Deutschen Biobanken-Register
 - ▶ Referenzimplementation Datenschutzkonzept A

Zentral betriebene IT-Produkte

- ▶ werden in einzelner Instanz zentral von TMF betrieben
 - ▶ daher keine Relevanz für Dissemination
 - ▶ Containerisierung ggf. für Betrieb denkbar

- ▶ Produkte
 - ▶ Autorenportal zur Leitlinienentwicklung
 - ▶ Deutsches Biobanken-Register
 - ▶ Hosting für SAE-Software
 - ▶ Informed Consent - Online-Assistent
 - ▶ IRene-Tool – Interaktiv Rechtsgrundlagen erkunden
 - ▶ MetaReg

Containerisierbare Produkte



Empfehlung ▼

- denkbar
- empfohlen
- existiert
- Umsetzung im Workshop

Containerisierte Produkte

- ▶ bereits Images verfügbar
 - ▶ Galaxy
 - ▶ MirthConnect
 - ▶ OpenClinica (im Workshop aktualisiert)
 - ▶ OpenMRS
 - ▶ REDCap (lizenzbedingt nur Dockerfile, im Workshop aktualisiert)
 - ▶ tranSMART (im Workshop aktualisiert)
 - ▶ XNAT

- ▶ im Workshop + Folgetermin 06/2016 containerisiert
 - ▶ i2b2
 - ▶ Mainzliste
 - ▶ Mainzer MDR (Bestandteil OSSE-Plattform)
 - ▶ Greifswalder MOSAIC-Plattform (EPIX, gPAS, gICS)
 - ▶ PID-Generator

- ▶ empfohlen
 - ▶ MAGIC
 - ▶ Mikrobiologische Primärdatenbank
 - ▶ MOSAIC
 - ▶ Nationales Metadata Repository
 - ▶ OSSE

- ▶ denkbar
 - ▶ Anonymisierungs-Tool
 - ▶ Greifswalder R-Paket zur Prüfung der Datenqualität
 - ▶ NLP-Analyse
 - ▶ Phase-II-Paket
 - ▶ Pseudonymisierungs-Software
 - ▶ RADAR
 - ▶ repliCheckSNP
 - ▶ R-Programmpaket zur flexiblen Anpassung von Mehrstufen-Designs für GWAS
 - ▶ SDTM-Wandler
 - ▶ Werkzeuge Datenqualität

Containerisierbare Produkte

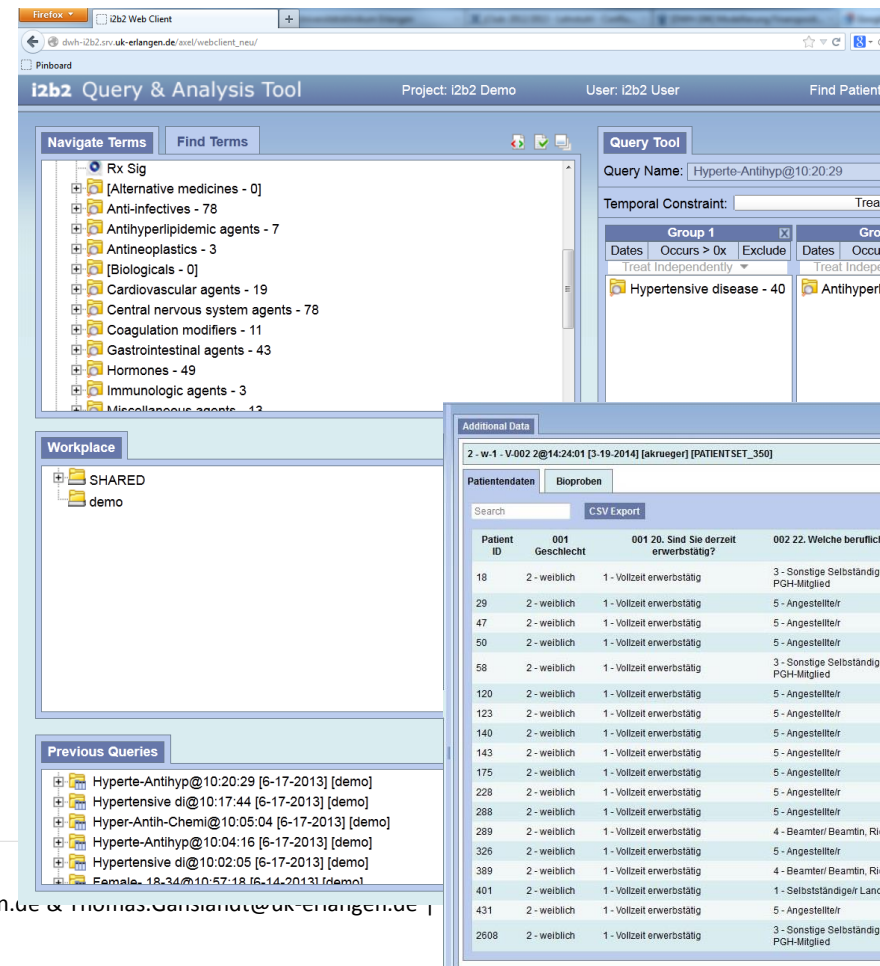
- ▶ empfohlen
 - ▶ MAGIC
 - ▶ Mikrobiologische Primärdatenbank
 - ▶ MOSAIC
 - ▶ Nationales Metadata Repository
 - ▶ OSSE

- ▶ denkbar
 - ▶ Anonymisierungs-Tool
 - ▶ Greifswalder R-Paket zur Prüfung der Datenqualität
 - ▶ NLP-Analyse
 - ▶ Phase-II-Paket
 - ▶ Pseudonymisierungs-Software
 - ▶ RADAR
 - ▶ repliCheckSNP
 - ▶ R-Programmpaket zur flexiblen Anpassung von Mehrstufen-Designs für GWAS
 - ▶ SDTM-Wandler
 - ▶ Werkzeuge Datenqualität

Vorstellung der Werkzeuge

i2b2: Informatics for Integrating Biology & the Bedside – Abfrage und Analyse von Daten unterschiedlicher Quellen

- ▶ Open Source Serveranwendung
- ▶ Datawarehouse zur Ablage von Phentypdaten
- ▶ Ad-hoc Kohortenerstellung und Export von Daten
- ▶ <https://hub.docker.com/r/tmfev/i2b2/>



The screenshot displays the i2b2 Query & Analysis Tool interface. The main window is titled "i2b2 Query & Analysis Tool" and shows a "Navigate Terms" panel on the left with a tree structure of medical terms. The "Query Tool" panel on the right shows a query named "Hyperte-Anthyp@10:20:29" with a temporal constraint. Below the query tool is a "Workplace" panel showing a file structure with "demo" and "Previous Queries" panel listing several queries. At the bottom right, an "Additional Data" panel displays a table of patient data.

Patient ID	001 Geschlecht	001 20. Sind Sie derzeit erwerbstätig?	002 22. Welche berufliche Tätigkeit haben Sie?
18	2 - weiblich	1 - Vollzeit erwerbstätig	3 - Sonstige Selbstständig PGH-Mitglied
29	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
47	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
50	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
58	2 - weiblich	1 - Vollzeit erwerbstätig	3 - Sonstige Selbstständig PGH-Mitglied
120	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
123	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
140	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
143	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
175	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
228	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
288	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
289	2 - weiblich	1 - Vollzeit erwerbstätig	4 - Beamter/ Beamtin, Ri
326	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
389	2 - weiblich	1 - Vollzeit erwerbstätig	4 - Beamter/ Beamtin, Ri
401	2 - weiblich	1 - Vollzeit erwerbstätig	1 - Selbstständig/er Lanc
431	2 - weiblich	1 - Vollzeit erwerbstätig	5 - Angestellter
2608	2 - weiblich	1 - Vollzeit erwerbstätig	3 - Sonstige Selbstständig PGH-Mitglied

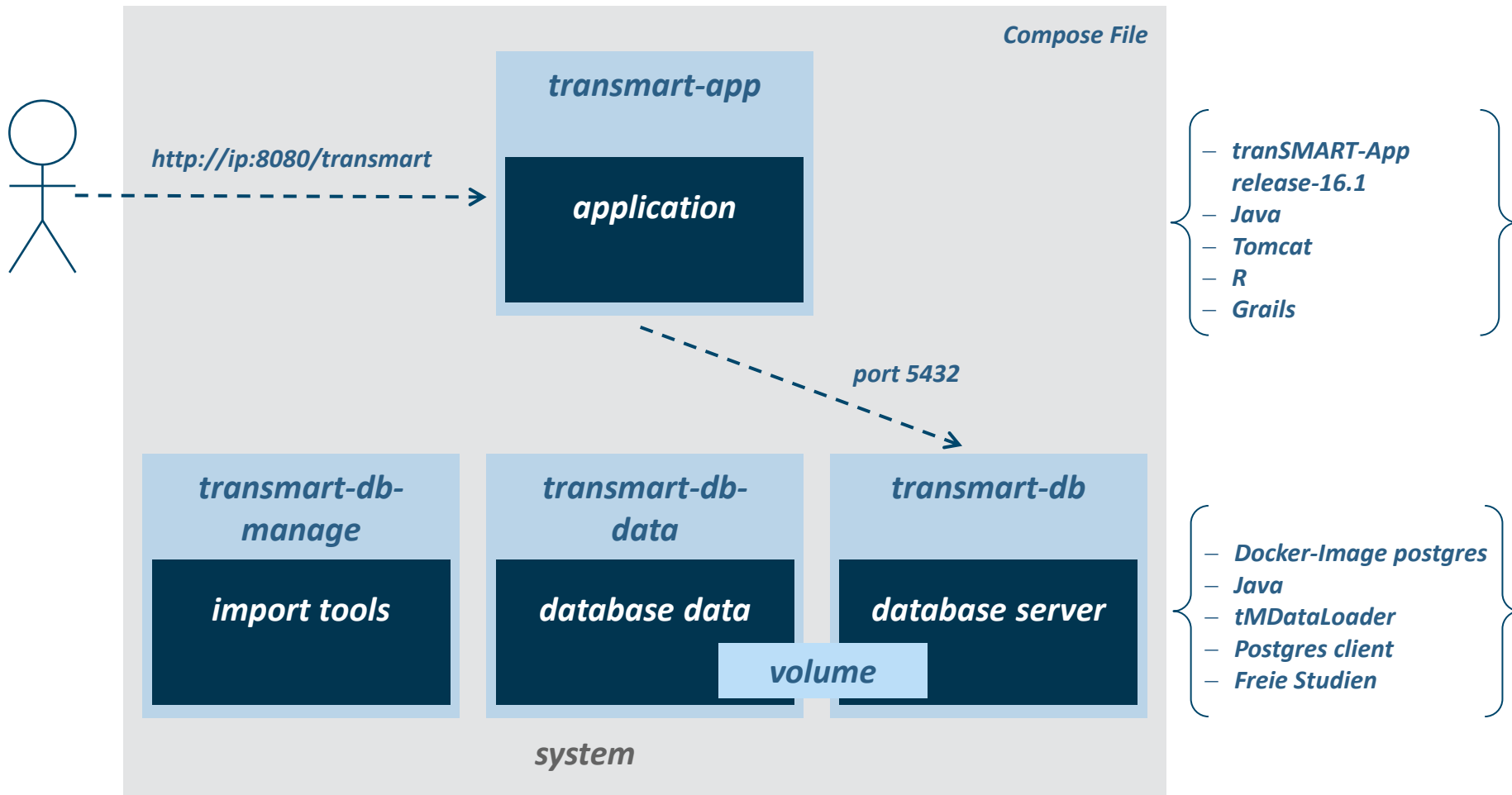
Vorstellung der Werkzeuge

tranSMART – Abfrage und Analyse von Daten unterschiedlicher Quellen

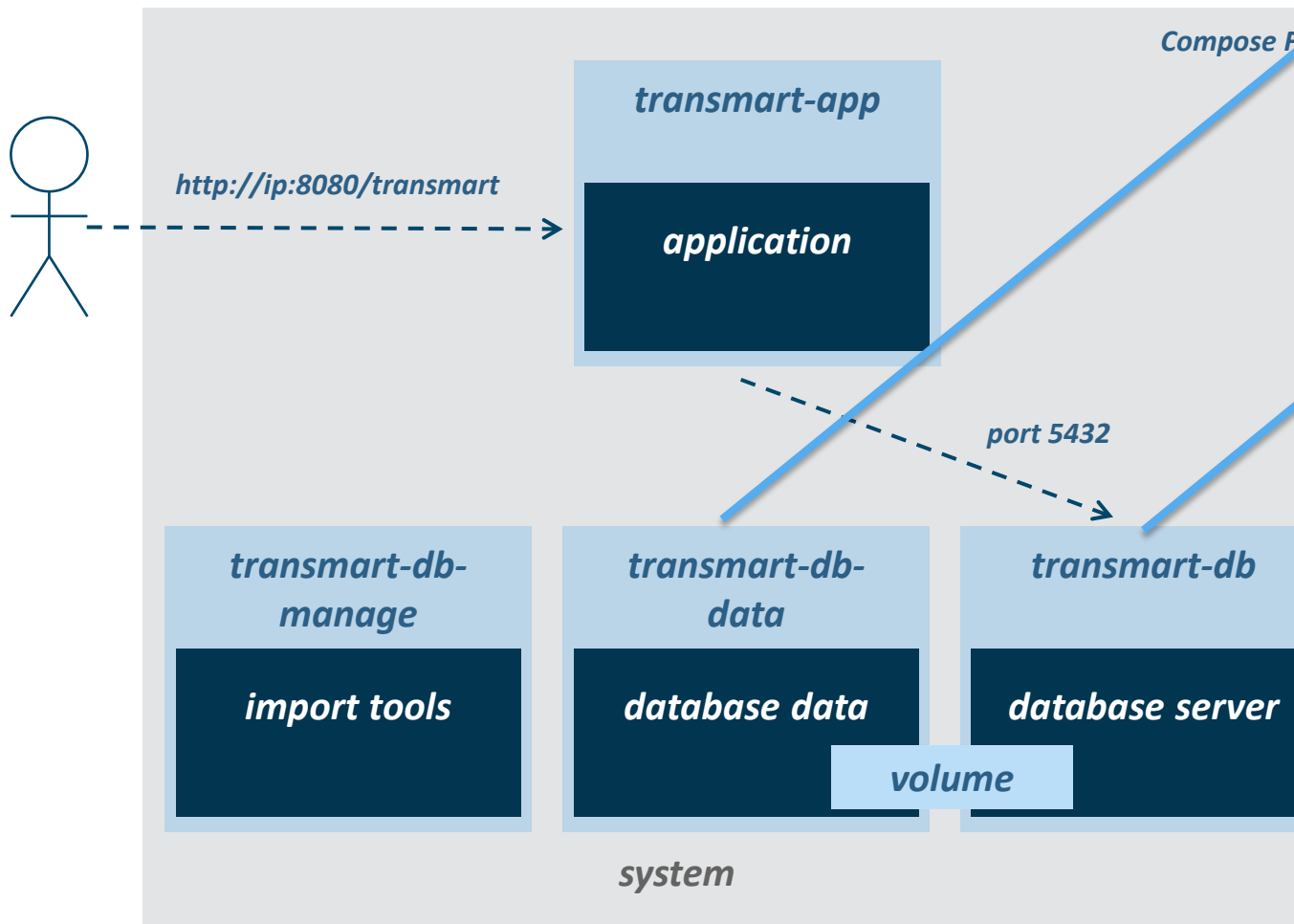
- ▶ Open Source Serveranwendung basierend auf i2b2
- ▶ Datawarehouse zur Ablage von Phenotyp- und NGS-Daten
- ▶ Build in: u.a. Genexpressionsdaten, SNP, VCF, mehr durch Community-Erweiterungen
- ▶ Ad-hoc Kohortenerstellung, Export und Analyse von Daten
- ▶ <https://hub.docker.com/u/tmfev/>



tranSMART Container-Lösung



tranSMART Container-Lösung

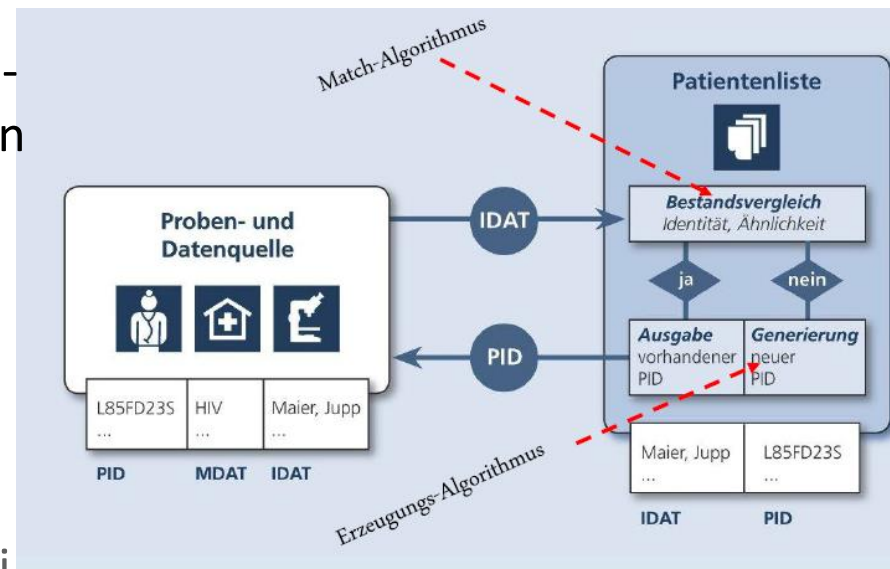


```
1 version: '2'
2 services:
3   transmartdbdata:
4     build:
5       context: ../tr
6       dockerfile: Do
7     volumes:
8       - /opt/pg
9   transmartdb:
10    build:
11      context: ../tr
12      dockerfile: Do
13    environment:
14      - PGDATA:/opt/p
15      - POSTGRES_PAS
16    ports:
17      - "5432:5432"
18    depends_on:
19      - transmartdbda
20    volumes_from:
21      - transmartdbda
22  transmartdbmanage:
23    build:
24      context: ../tr
25      dockerfile: Do
26    depends_on:
27      - transmartdb
28    links:
29      - transmartdb:
```

Vorstellung der Werkzeuge

TMF-PID-Generator – Erzeugung nichtsprechender Identifikatoren

- ▶ Pseudonymisierungstool zur Erstellung von PID, SIC, Forschungs-ID nach TMF-Datenschutzrichtlinien
- ▶ Auffangen von Fehlern in IDAT
- ▶ Einbindung in verschiedene medizinische Forschungsszenarien als eine zentrale Komponente des ID-Managements*
- ▶ <https://hub.docker.com/r/tmfev/pidgen/>

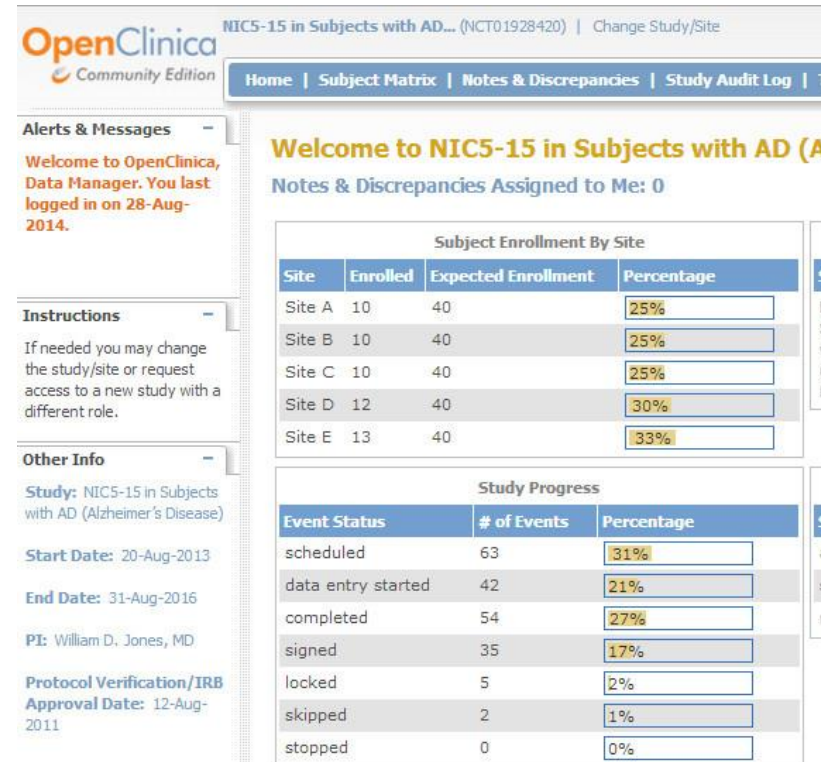


* Aus: Prof. Dr. Klaus Pommerening, Dr. Murat Sariyar: „Der PID-Generator der TMF“ - TMF-Workshop „Tools zum ID-Management in der klinischen Forschung“. Foliensatz Berlin, 24. September 2010

Vorstellung der Werkzeuge

OpenClinica – Freie Software für Electronic Data Capture

- ▶ Community und Enterprise Serveranwendung
- ▶ eCRFs, Rules/Checks
- ▶ Audit trail, Role-based access
- ▶ <https://hub.docker.com/r/tmfev/openclinica/>



The screenshot displays the OpenClinica Community Edition interface for a study titled "NIC5-15 in Subjects with AD... (NCT01928420)". The page includes a navigation bar with links for Home, Subject Matrix, Notes & Discrepancies, and Study Audit Log. A welcome message indicates the user is logged in on 28-Aug-2014. The main content area is divided into two sections: "Subject Enrollment By Site" and "Study Progress".

Subject Enrollment By Site

Site	Enrolled	Expected Enrollment	Percentage
Site A	10	40	25%
Site B	10	40	25%
Site C	10	40	25%
Site D	12	40	30%
Site E	13	40	33%

Study Progress

Event Status	# of Events	Percentage
scheduled	63	31%
data entry started	42	21%
completed	54	27%
signed	35	17%
locked	5	2%
skipped	2	1%
stopped	0	0%

Other Info

- Study:** NIC5-15 in Subjects with AD (Alzheimer's Disease)
- Start Date:** 20-Aug-2013
- End Date:** 31-Aug-2016
- PI:** William D. Jones, MD
- Protocol Verification/IRB Approval Date:** 12-Aug-2011

Vielen Dank für Ihre Aufmerksamkeit!