

# Toolbox for Research

Stand: 01.06.2017

## Hintergrund

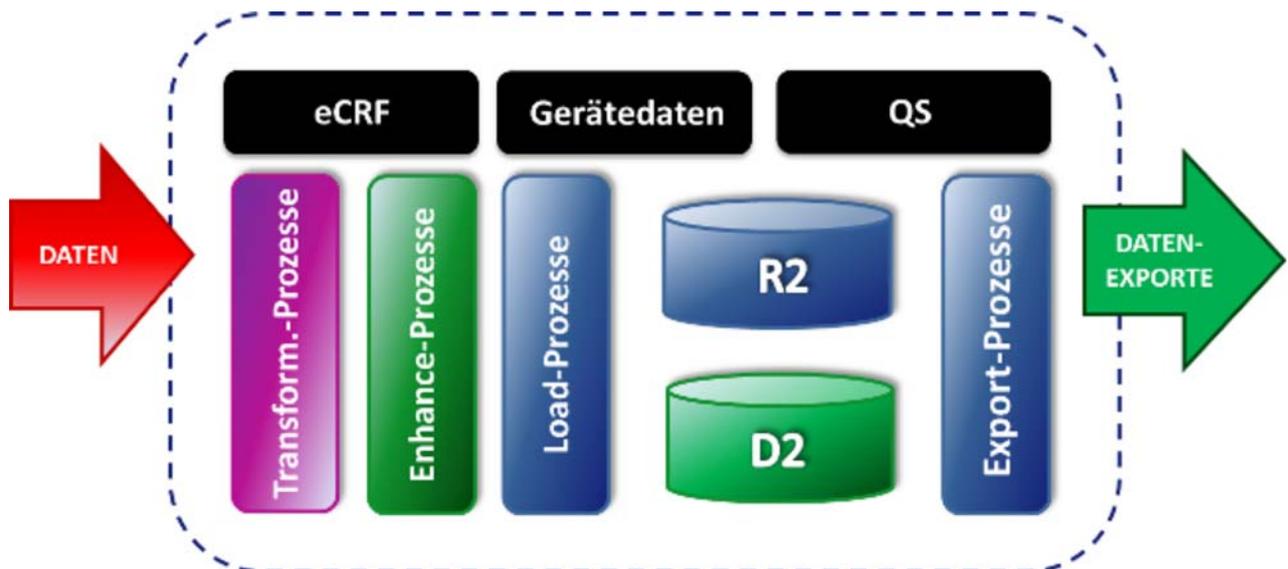
Kohortenstudien und Register unterscheiden sich in Bezug auf das Anwendungsszenario, die eingesetzte Erhebungsmethode (Formulare), das Data Dictionary, genutzte (Labor-) Geräte, QS-Methoden, Exportformate uvm.

Um vor allem kleineren Forschungsprojekten mit geringen IT-Ressourcen (Expertise, Personal, Infrastrukturen), die durchaus bisher auf Excel und Co. angewiesen waren und erhobene medizinische Daten manuell pseudonymisieren und zur Auswertung händisch zusammenführen mussten, eine einfache Möglichkeit zur web-basierten und standortübergreifende Datenerfassung zu bieten, wurde im Rahmen des MOSAIC-Projekts die TOOLBOX FOR RESEARCH entwickelt.

## Toolbox for Research

Die entwickelte Toolbox for Research stellt eine modulare und nachnutzbare Studiendatenbank sowohl für Metadaten in einem integrierten Data Dictionary (D2) als auch für Forschungsdaten in einem Research Repository (R2) bereit. Das zugrundeliegende Datenmodell fand bereits im GANI\_MED-Projekt Anwendung.

Die Toolbox vereinfacht die separate Speicherung von Forschungs- und Metadaten, da sie wesentliche technische Prozesse (Extraktion, Transformation, Aufbereitung von Forschungsprimärdaten, Export) automatisiert und beliebig um einzelne Module (z.B. zur Qualitätssicherung) als auch um mehrere Datenquellen (Formulare, Geräte, etc.) ergänzt werden kann.



Die Toolbox enthält eine direkt nutzbare OpenClinica-Installation, was die Generierung von web-basierten Fragebögen und Metadatenbeschreibungen ermöglicht. Die Definition der Metadaten ist ergänzend natürlich auch manuell (per Excel-Vorlage) möglich. Weitere von OpenClinica genutzte Vorteile sind die Verwaltung von Standorten und Nutzern, sowie die Vergabe von entsprechenden Rollen und Rechten.

Darüber hinaus wurden in die Toolbox ein Dispatcher-Modul und ein Web-Server integriert. Das Dispatcher-Modul vereinfacht das Anlegen und Suchen von Patienten und steuert die zentrale Vergabe von standort-spezifischen und projektspezifischen Pseudonymen. Der Web-Server stellt die dafür erforderlichen Web-Oberflächen bereit.

Ein beispielhaftes Exportmodul aggregiert in Intervallen Metadaten und Forschungsdaten und stellt diese derzeit im SPSS-Format für die Datenauswertung bereit. Automatisierbare Möglichkeiten zur einheitlichen Integration von Gerätedaten in den Datenbestand sind vorgesehen.

**HINWEIS: Für das Anlegen von Patienten/Studienteilnehmern sind eine PatientenID und eine Fallnummer erforderlich. Die Bedeutung von PatientenID und Fallnummer sind und bleiben allein der eingebenden Person im Studienstandort bekannt.**

## Was leistet die Toolbox?

- zentrale Vergabe standort- und studienspezifischer Pseudonyme dank gPAS-Integration
- Suche von Pseudonymen/Depseudonymisierungsfunktion für autorisiertes Personal
- vereinfachtes Anlegen und registrieren von Studienteilnehmern/Patienten dank Dispatcher-Integration
- Formulargenerierung mittels Leitfäden und OpenClinica-Mechanismen
- Web-basierte und standort-übergreifende Datenerfassung durch OpenClinica-Integration
- Getrennte Speicherung von Metadaten und Forschungsdaten
- Möglichkeit zur Gerätedatenintegration (z.B. im CSV-Format)
- Datenexport im SPSS-Format
- Verlaufsverfolgung von Studienteilnehmern (z.B. bei Standortwechsel), durch Mitführen einer „vorherigen TeilnehmerID“

## Was leistet die Toolbox NICHT?

- Keine Definition des Data Dictionary, lesen Sie hierzu bitte den Leitfaden
- Keine Definition der Fragebögen, lesen Sie hierzu bitte den Leitfaden
- Keine Dublettenerkennung, denn die Toolbox speichert KEINE personenidentifizierenden Daten (IDAT) außer Geschlecht und Geburtsdatum (Beides ist zum Anlegen von Teilnehmern in OpenClinica erforderlich.)
- Keine Treuhandstellen-Funktion
- Keine automatische Datenauswertung
- Keine Kontrolle der wissenschaftlichen Qualität des Data Dictionary, der Fragebögen, der Daten und Datenauswertung
- Keine Erstellung des notwendigen Datenschutzkonzeptes, lesen Sie hierzu bitte den Leitfaden
- Die Toolbox verwendet derzeit OpenClinica (v.3.4, neuere in Vorbereitung). Dadurch entstehen Limitierungen bei der Formularerstellung in Bezug auf Sprünge, die Validierung von Werten und Eingaben, Variablenübergreifende Plausibilitätsprüfungen, u.a.

## Welche Ressourcen sind zum Einsatz der Toolbox erforderlich?

### Personell

- Definition und Aktualisierung von Data Dictionary und Fragebogen
- Eigenverantwortlicher Betrieb der Toolbox, d.h. Systembereitstellung, Installation, Wartung, Monitoring, Pflege, IT-Sicherheit, Datensicherheit, Benutzerverwaltung, u.a.

- Weitere Infos siehe Check-Liste für Administratoren

### Technisch

- 64-Bit Hostsystem Debian8 oder Ubuntu Server 15.04 (Linuxkernel 3.10 oder höher)
- Mindestens 6GB Arbeitsspeicher
- 50GByte freier Festplattenspeicher (ohne Hostsystem)
- Docker
- Internetzugriff
- Offizielles und gültiges Server-Zertifikat für gesicherte Datenkommunikation
- Domainregistrierung
- Weitere Infos siehe Check-Liste für Administratoren

## Was muss der Betreiber der Toolbox eigenverantwortlich leisten?

---

Die relevanten Punkte wurden in einer Checkliste für Administratoren zusammengefasst. Diese finden Sie nach Registrierung im Downloadbereich.

## Installation, Einrichtung und Betrieb der Toolbox

---

Die Installation der Toolbox wurde maßgeblich durch den Einsatz von Docker vereinfacht. Dies ermöglicht die vollständig automatisierte Installation der Toolbox in Containern mittels eines einzigen Skripts. Auf diese Weise können Daten und verarbeitende System voneinander getrennt werden und zeitgleich ressourcensparend gearbeitet, sowie Wartung, Update und Wiederherstellungsprozesse vereinfacht werden.

Die Dokumente zur Installation, Einrichtung, zum Betrieb und zur Nutzung der Toolbox finden Sie nach Registrierung im Downloadbereich.

## Kontakt

---

Bitte nutzen Sie unser **Kontaktformular** für Fragen und Anregungen oder wenden Sie sich direkt an einen der folgenden **Ansprechpartner**:

*Ronny Schuldt (ronny.schuldt@uni-greifswald.de, 03834-86 7520)*

*Arne Blumentritt (arne.blumentritt@uni-greifswald.de, 03834-86 19538)*

*Martin Bialke (martin.bialke@uni-greifswald.de, 03834-86 7580)*