

# MOSAIC: Werkzeuge aus der Praxis und für die Community

Dr. Martin Bialke

Institut für Community Medicine  
Abt. Versorgungsepidemiologie und Community Health  
Universitätsmedizin Greifswald

9. TMF Jahreskongress  
Göttingen, 15. März 2017

- DFG-gefördertes Einzelprojekt der Ausschreibung „*Informationsinfrastrukturen für Forschungsdaten*“
- Ziel:
  - Datenmanagement in der epidemiologischen Forschung durch praxisorientierte Werkzeuge vereinfachen
- Zielgruppe:
  - Kohorten und Register (mit minimaler IT-Unterstützung)
- Kostenfreie Bereitstellung der Werkzeuge
  - Software: AGPLv3
  - Dokumente & Co. : CC-BY 3.0

## TMF

- Vorstellung ausgewählter Werkzeuge in den Arbeitsgruppen *Datenschutz* und *IT-Infrastruktur und Qualitätsmanagement*
- Workshop-Unterstützung
- Kooperationsprojekte „Dockerbank“ und MAGIC (DFG)

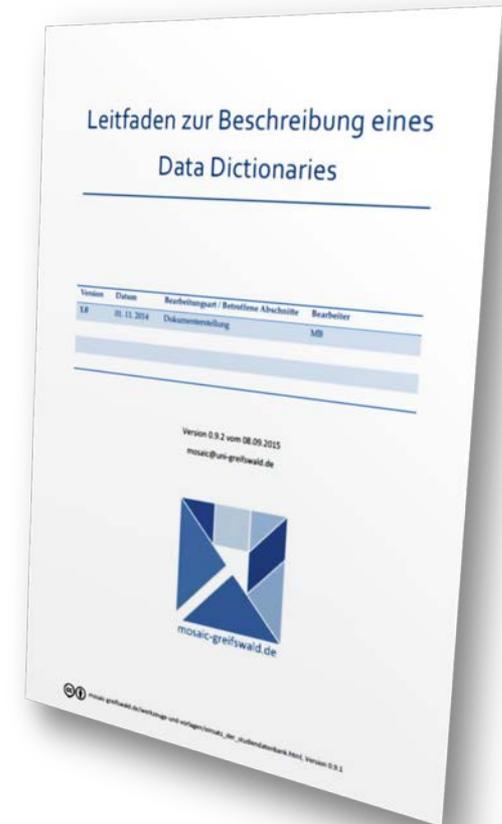
## Aktuelle Anwender-Projekte



Praktische Hilfestellung durch

# VORLAGEN, LEITFÄDEN UND EMPFEHLUNGEN

- Variablennamen und Ausprägungen
- Typische Datentypen und Wertebereiche
- Codierung gültiger Werte und Missings
- Empfehlungen aus der Praxis
- Abstimmungsbedarf
- Vorlagen und Beispiele
- Sprachen: deutsch, englisch



<http://www.toolpool-gesundheitsforschung.de/produkte/leitfaden-dd-ecrf>

# Leitfaden Erstellung von eCRF

- Fragenformulierung und –struktur
- Tipps zur Bestimmung der Eingabeelemente
- Wahl der eCRF-Lösung
- Formulardefinition am Beispiel von OpenClinica
- Tipps zur Fragebogengenerierung
- Empfehlungen aus der Praxis
- Weiterführende Literatur und Beispiele



<http://www.toolpool-gesundheitsforschung.de/produkte/leitfaden-dd-ecrf>

# Musterdokument Datenschutzkonzept

- praxisnaher Schnelleinstieg
- „Checkliste und Erinnerungshilfe“:
- vorgegebene Struktur führt durch typische Problemfelder des Datenschutzes in der medizinischen Forschung
- Flexibel je Szenario anpassbar durch fragenbasierten Ansatz



<http://www.toolpool-gesundheitsforschung.de/produkte/vorlage-datenschutzkonzept>

## Guide

- Erläuterung möglicher Schwächen bei der Datensicherung / –wiederherstellung
- Übersicht typischer Sicherungsarten und Einflussfaktoren

## Vorlage

- Vereinfacht Planerstellung
- Unterstützt bei Einrichtung von Sicherungsmaßnahmen
- Handlungsunterstützung im Schadensfall



<http://www.toolpool-gesundheitsforschung.de/produkte/vorlage-datensicherung>

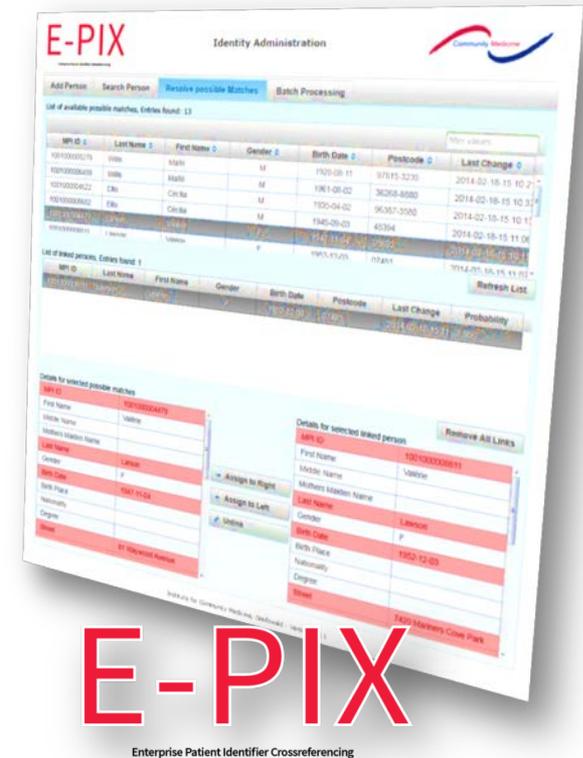
Unterstützung durch

# KOSTENFREIE SOFTWARE-WERKZEUGE

# ID-Management mit E-PIX

- Verwaltung von Personen und Identitäten
- Datenzusammenführung aus div. Quellen
- Umgang mit fehlerhaften bzw. unvollständigen Daten
- Unterstützung bei der Auflösung von möglichen Synonymfehlern (Dublettenerkennung)
- Konfigurierbare Matching-Parameter
- Import und Export (CSV-Format)

-> Optimal für die zentrale Datenhaltung mit Datenübermittlung



<http://www.toolpool-gesundheitsforschung.de/produkte/e-pix>

# Pseudonymverwaltung mit gPAS

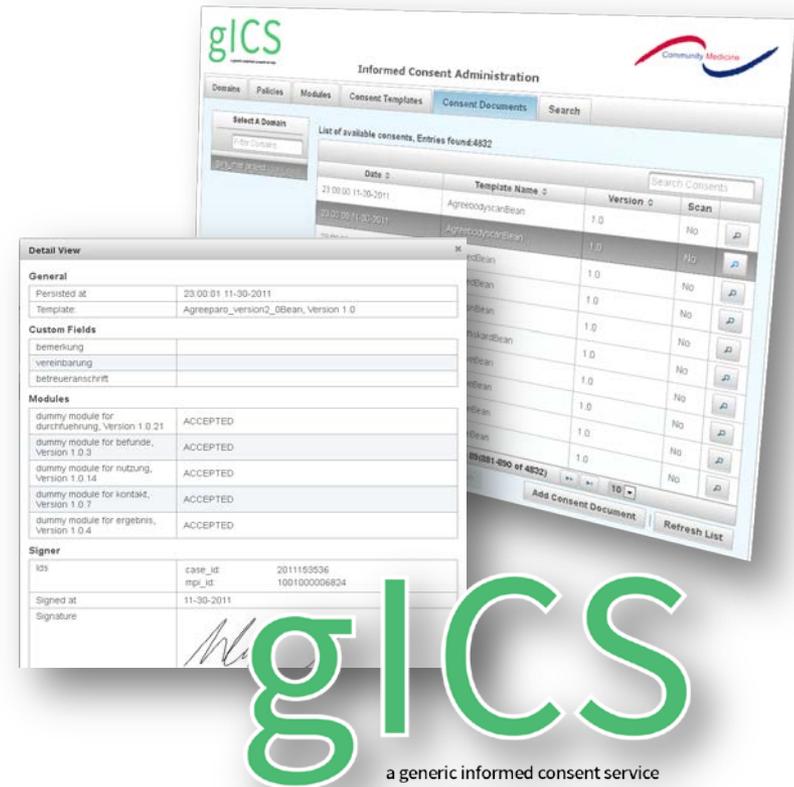


- Pseudonyme generieren und verwalten
- Verarbeitung beliebiger Zeichenfolgen
- Mehrfach-Pseudonymisierung
- De-Pseudonymisierung und (faktische) Anonymisierung
- Flexibel konfigurierbar
- Integration von Altpseudonymen
- Anzeige von Pseudonym-Hierarchien
- Erstellung temporärer Pseudonyme

<http://www.toolpool-gesundheitsforschung.de/produkte/gpas>

# Einwilligungsmanagement mit gICS

- Verwaltung modularer Einwilligungen und Widerrufe
- policy-spezifische Abfragen des Einwilligungsstatus
- flexibel kombinierbare Policies und Module
- Individuell anpassbar



**gICS**  
a generic informed consent service

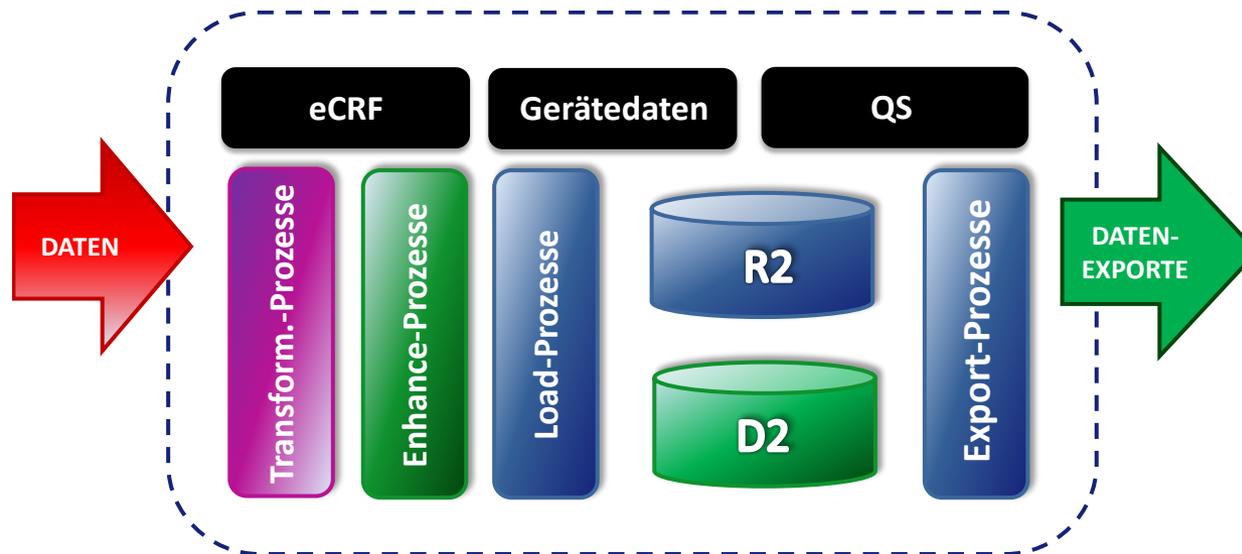
<http://www.toolpool-gesundheitsforschung.de/produkte/gics>

## Modulares Einwilligungskonzept



## Ziel

- kleinere Forschungsprojekte mit geringen IT-Ressourcen (Expertise, Personal, Infrastrukturen) eine vereinfachte Möglichkeit für web-basierte und standortübergreifende Datenerfassung bieten



<http://www.toolpool-gesundheitsforschung.de/produkte/toolbox-for-research>

## Leistungsumfang

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

## Ziel: Qualität metrischer und kategorialer Daten grundlegend prüfen ohne Kenntnisse von R vorauszusetzen

```
# specify the csv import file with metric data, use one column per variable  
metric_datafile='c:/mosaic/sample_data/metric_single_var.csv'
```

```
#specify output folder  
outputFolder='c:/mosaic/output/'
```

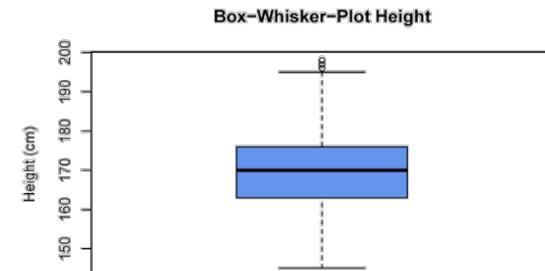
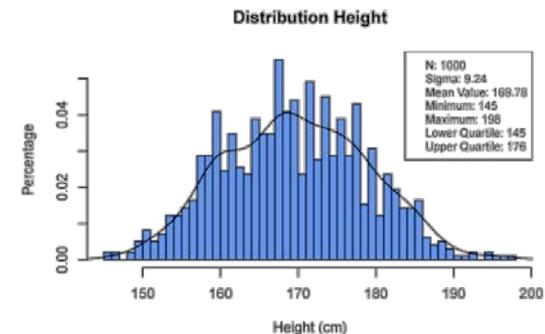
```
# load mosaicQA package  
library('mosaicQA')
```

```
#set missing threshold, optional, default is 99900  
mosaic.setGlobalMissingTreshold(99900)
```

```
#set variable unit, optional  
mosaic.setGlobalUnit(' (cm) ')
```

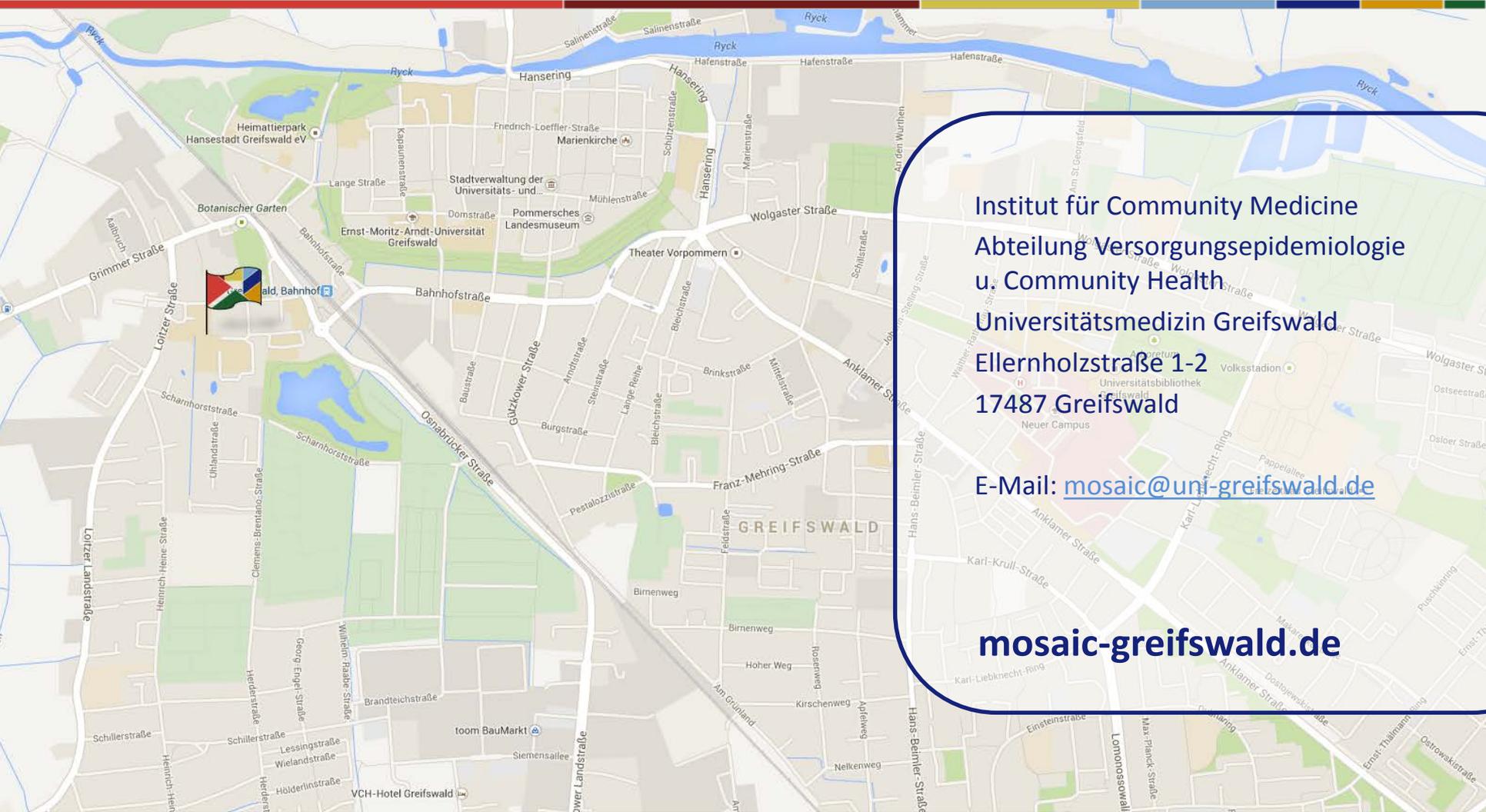
```
#set variable description, optional  
mosaic.setGlobalDescription('Height')
```

```
#create PDF-report  
mosaic.createSimplePdfmetric(metric_datafile, outputFolder)
```



<http://www.toolpool-gesundheitsforschung.de/produkte/mosaicQA>

# Vielen Dank für Ihre Aufmerksamkeit



**Institut für Community Medicine  
Abteilung Versorgungsepidemiologie  
u. Community Health  
Universitätsmedizin Greifswald  
Ellernholzstraße 1-2  
17487 Greifswald**

**E-Mail: [mosaic@uni-greifswald.de](mailto:mosaic@uni-greifswald.de)**

**[mosaic-greifswald.de](http://mosaic-greifswald.de)**

