

Management von Datenelementen mit dem nationalen Metadata Repository (MDR)

Standardisierte Datenelemente zur Dokumentation in klinischen Studien und epidemiologischen Registern

Matthias Löbe¹, Johannes Drepper²

¹Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE), Leipzig

²TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., Berlin

Motivation

- IMISE
 - Verbundprojekte „Familiärer Darmkrebs“ und „Familiärer Brust- und Eierstockkrebs“ der Deutschen Krebshilfe
 - Integriertes Forschungs- und Behandlungszentrum: Register für Adipositas (ca. 100 Datenelemente)
- Zentrum für Klinische Studien Leipzig
 - 100 laufende klinische Studien mit 200-1000 Datenelementen
- LIFE-Forschungszentrum für Zivilisationserkrankungen
 - Große epidemiologische Kohorte (26.000 Probanden/Patienten)
 - Befragungen, Selbstausfüller, Messungen mit Geräten, Laborbestimmungen, ca. 17.000 Datenelemente (und steigend)

Großer Aufwand bei der Spezifikation und Wartung des Dokumentationskonzeptes/der enthaltenen Datenelemente!

- *Neue Mitarbeiter erschaffen immer wieder neue Formulare (unterschiedlicher Qualität)*

Wie entstehen Dokumentationskonzepte heute?

- Unstrukturiert im Umlaufverfahren
- In diesem Bild ist ein Verbesserungsbedarf versteckt!

Vorbekannte Tumorerkrankungen, ggfs mit ICD-10: Suchmöglichkeit ICD-10 C00-D48

Chemotherapie in Anamnese: J/N/weiß nicht

Strahlentherapie in Anamnese: J/N/weiß nicht

[Sonstige Tumorbehandlungen: J/N, bei J Freitext zum Spezifizieren **!zunächst rauslassen!**]

f Relevante Begleiterkrankungen: s. Datei „relevante Begleiterkrankungen“ (oder *minimale Organe* → *Spezifität*)

[Tumor-Histologie n. ICD O.3: Freitext mit genauer Bezeichnung / ~~zunächst ersetzen durch~~

Tumor-Grading: G 1-4

↘ *Art des Primärtumors*

Rezeptorstatus, HER2-Status

[Patholog. Institut + Einsendenummer)???) **zunächst rauslassen**

Klin. Bzw. Pathologischer TNM-Status: Pulldown für T - N - M

Wo ist das Problem?

- Datenelemente werden nicht detailliert genug spezifiziert!
 - Fehlende Maßeinheiten
 - Fehlende Formatierungen (Datum)
 - Fehlende Genauigkeiten (Nachkommastellen?)
 - Unklare Kategorien (Mehrfachauswahl zulässig?)
 - Fehlende Angaben zu Pflichtfeldern
 - Undurchsichtige logische Verzweigungen beim Ausfüllen
 - Unklare Messmethoden (Liegend? In Ruhe?)
 - Fehlende Normbereiche
 - Unterschiedlich geeichte Laborgeräte und Analyseverfahren
 - Viel Freitext („Andere“, „Sonstige“)



Bsp: ADT-Datensatz der Tumorzentren

Arbeitsgemeinschaft
Deutscher Tumorzentren

Kooperationsverbund Qualitätssicherung
durch Klinische Krebsregister (KoQK)

Version 6.0 Stand 02/08/Basis
DIAGNOSEDATEN

DEUTSCHE
KREBSGESELLSCHAFT E.V.

Gesellschaft d. epidemiologischen
Krebsregister in Deutschland

Deutsche
Krebshilfe

CCC
Forum

Einwilligung zur Meldung an das Klinische und Epidemiologische Krebsregister

Liegt vor Liegt nicht vor Abgelehnt

Patientenidentifikationsnummer: Krankenkasse:

Telefon: Name:

Meldende Institution: Geburtsname: Geschlecht: w/m

Geschlecht aus
PDF-Formular

```

<xs:element name="Geschlecht">
  <xs:simpleType>
    <xs:restriction base="xs:token">
      <xs:pattern value="[MWU]" />
    </xs:restriction>
  </xs:simpleType>
</xs:element>
<xs:element name="Geburtsdatum" type="Datum_Geburt_Typ"/>
<xs:element name="Strasse" type="xs:string"/>
    
```

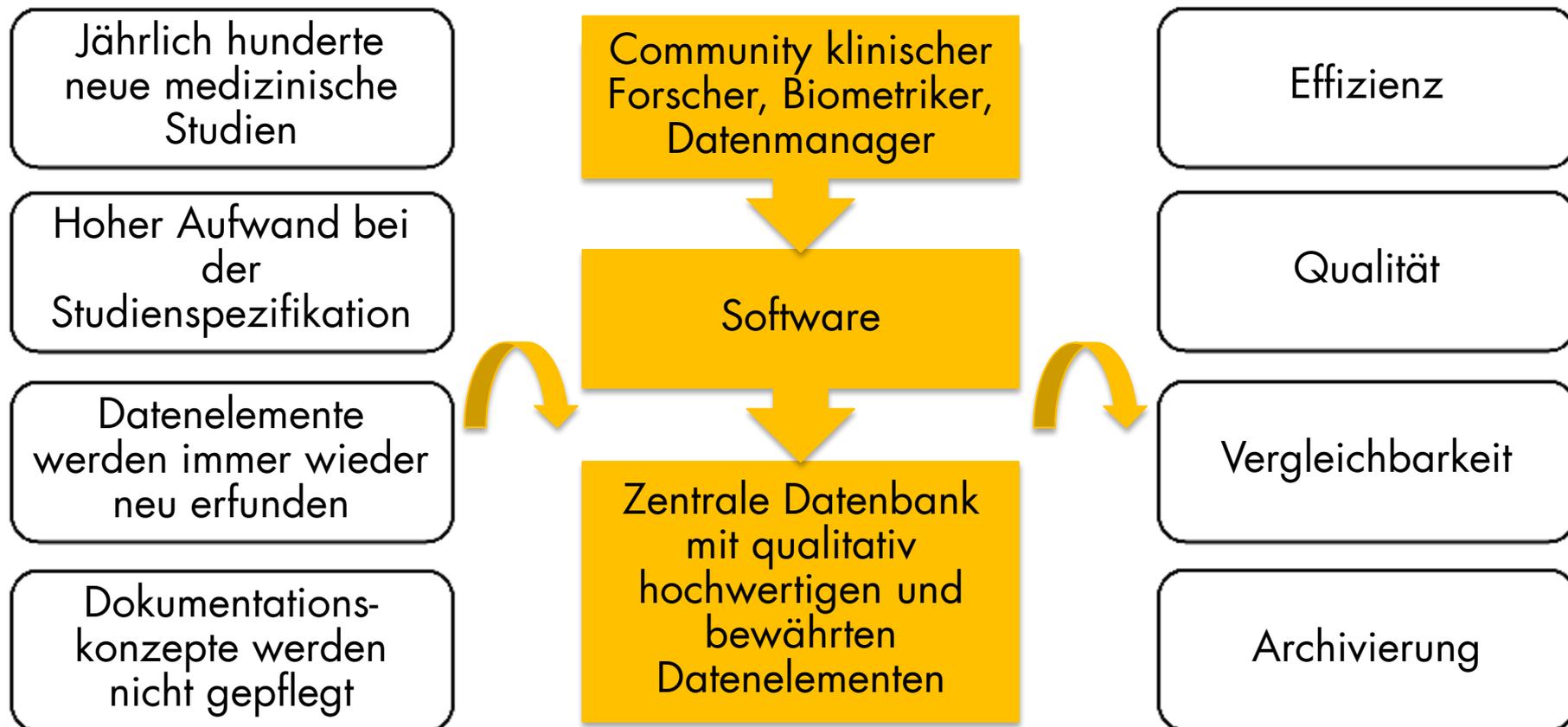
Geschlecht aus
XML-Schema

Bezeichnung	Codierung	Bedeutung / Erläuterung	Anwendung
Stammdaten		Einmal pro Patient vorzuhaltende Informationen	
Name, Vorname, Geburtsname, Adresse	Geschlecht aus Handbuch	Identifikation, aktuelle Adresse, für EKR eigentlich die Adresse zum Zeitpunkt der Erkrankung zur korrekten Einordnung der Inzidenz	Anschreiben, EKR-Meldung
Geschlecht	M=Männlich W=Weiblich X=Unbekannt		Stratifikation und Selektion von Daten, EKR-Meldung

Das nationale Metadata Repository (MDR)

5

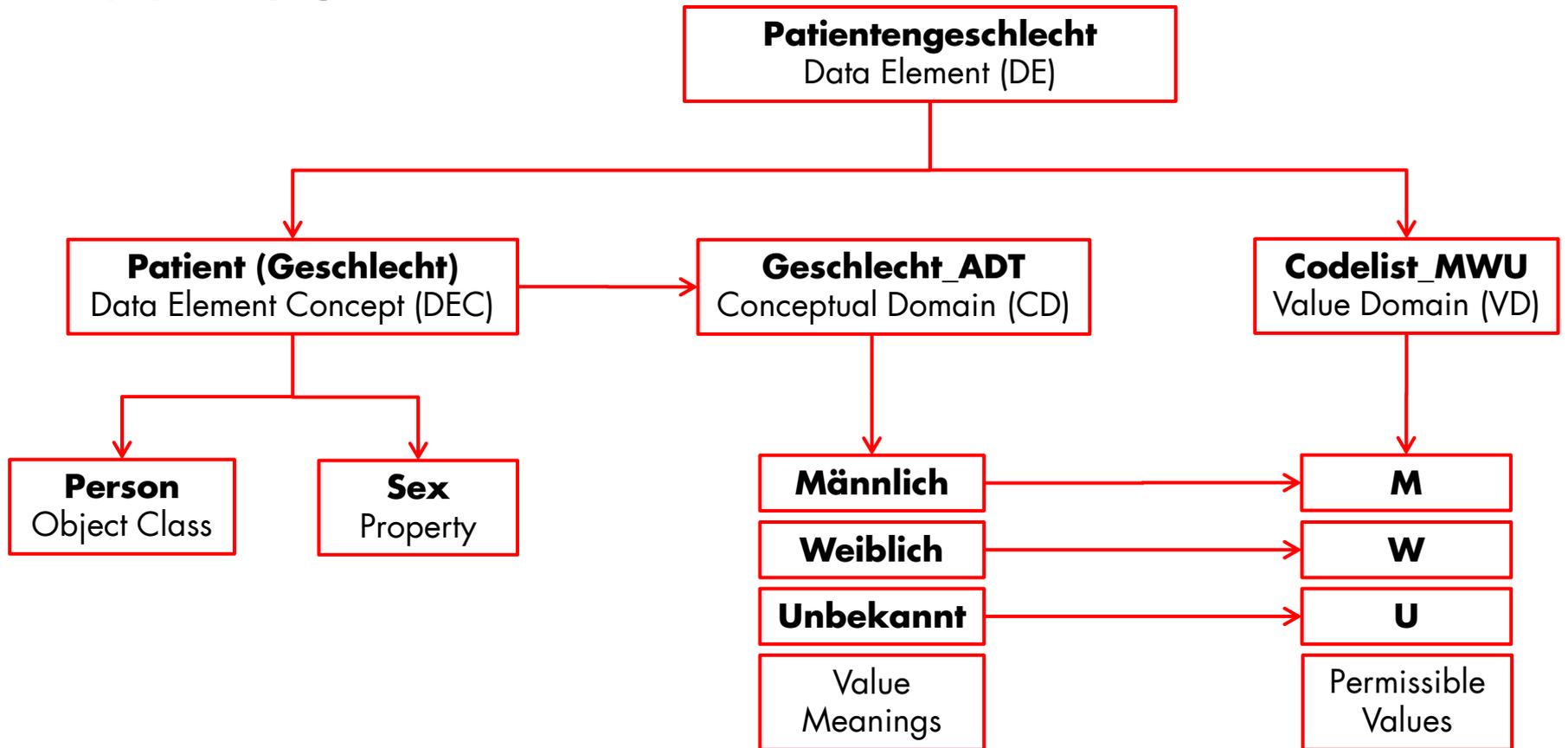
Ziele des BMBF-Projekts „Metadata Repository“



Zentrales Verzeichnis von Datenelementen (MDR) – Aktuelle Arbeiten

1. Kann man ein **Lexikon** aller wichtigen Datenelemente für klinische Studien, aber auch für Register oder Patientenakten anlegen („Biometrie-pedia“)?
2. Kann man so ein System in einem Studienzentrum **effektiv** nutzen?
3. Kann man Datenelemente in einer **Community** austauschen? Kann man „gute“ Datenelemente markieren?
4. Kann man **Versionen**, Varianten, Alternativen und Abbildungen verwalten?
5. Kann man vorhandenes Wissen aus medizinischen **Terminologien** (ICD-10, OPS, LOINC, MedDRA) nutzen?
6. Kann man damit genormte **Dokumentationskonzepte** für Forschungsprimärdatensammlungen erstellen?
7. Kann man Vergleiche oder **Metaanalysen** unterstützen?

Beispiel für ein kategoriales Datenelement nach ISO 11179



Konzeption
(Was wird gemessen)

Repräsentation
(Wie wird gemessen)

{ Object Class + Property = Data Element Concept } + Value Domain = Data Element

Stand Softwareprototyp

Web-Applikation

The screenshot shows the MDR Design Studio web application interface. The browser address bar displays `mdr.imise.uni-leipzig.de/Workshop/`. The application has a menu bar with 'New' and 'Edit' options, and a toolbar with buttons for 'Save', 'Home', 'Copy', 'Paste', 'Remove', 'Simple', 'ISO11179', 'Browser', 'Preview', 'Source', and 'Version'. The 'Preview' button is currently selected.

The main workspace is divided into three panels, all titled 'Geschlecht':

- Left Panel (Properties):** Shows metadata for the 'Geschlecht' element.
 - Parent: BGS
 - Language: de-DE
 - Designation:
 - Definition:
 - Value Domain: `Geschlechturn:mdr:ValueDomain@65544`
- Middle Panel (Value Domains):** Shows the configuration for the value domain.
 - Parent: Value Domains
 - Language: de-DE
 - Designation:
 - Definition:
 - Data Type: `Zahlurn:mdr:DataType@32774`
 - Format:
 - Length:
 - Unit of Measure: (no value)
 - Table:**

Code	Text
1	Männlich
2	Weiblich
 - Members: (empty list)
- Right Panel (Data Element Preview):** Shows a preview of the data element.
 - Label: Geschlecht
 - Radio buttons:
 - 1: männlich
 - 2: weiblich

On the far right, there is a sidebar with a search filter '(All)' and a list of related elements:

- `gesc`
- `Geschiedenurn:mdr:ValueMeaning@6573`
- `Angestellter mit umfassenden Führu...`
- `Geschlechturn:mdr:ValueDomain@65544`
- `Geschlechturn:mdr:DataElement@65868`
- `Geschwisterurn:mdr:DataElement@32844`
- `Geschlechturn:mdr:DataElement@32839`

Datenelemente aus CRF

View Section Data Entry for RE-T0 V1.1

CRF Info

Press the little flag icon beside an input to enter discrepancy notes, please note that you can only save the notes if CRF data entry has already started.

Exit

Fokus (0/16)

Patient... (0/13)

Nebendi... (0/22)

Title: Patientendaten

Page: 4

Alter (Jahre)

Geschlecht keine Angabe

Gewicht (kg)

Größe (cm)

Ethnische Zugehörigkeit keine Angabe

Aufnahme im Klinikum (Datum)

Aufnahme auf ITS (Datum)

Aufnahme auf ITS (Uhrzeit) (hh:mm)

Art der Zuweisung keine Angabe

Aufnahme auf ITS von keine Angabe

OpenClinica View Data Entry - Mozilla Firefox

uni-leipzig.de https://www.cscce.imise.uni-leipzig.de/OpenClinica/ViewSectionDataEntry?crfVersionId=818...

OpenClinica View Data Entry

View Section Data Entry for RE-T0 V1.1

CRF Info

Press the little flag icon beside an input to enter discrepancy notes, please note that you can only save the notes if CRF data entry has already started.

Exit

Nebendi... (0/22)

Klinisc... (0/37)

Glasgow... (0/4)

-- Select to Jump --

Title: Klinische Daten

Page: 6

Werte des 24h-Zeitintervalls nach dem Beginn schwere Sepsis/septischer Schock

Falls nur ein Wert erhoben wurde, diesen bitte in Feld "min." eintragen.

Körpertemperatur unbekannt min. (°C)

Mittlerer art. Blutdruck unbekannt min. (mmHg)

Herzfrequenz unbekannt min. (1/min)

Atemfrequenz (spontan oder beatmet) unbekannt min. (1/min)



Anamnese Kinderklinik		ANM KIK 01		Dokumentdaten	
Patientendaten		erw. Falldaten		erstellt	Z1WASCH
Name, Vorname	M107093587_Detlef	pfl. OE	KAID1	erstellt am	18.01.2012
GebDat	10.10.2009	Alter	2	Fachl. OE	KAID
Pat-ID	4545680	Fall-ID	12861759	geplante Entlassung	
		Geschl.	M	zuletzt geände	Z1WASCH
				zuletzt geändert am	18.01.2012

- Patientenorganizer
- Labordokumente
- Fallübersicht

Datenelemente aus KIS

aktuelle Anamnese	aktueller Besuch	Eigenanamnese	Familienanamnese	Medikamente	Anamnesen
Datum	18.01.2012				
Anamnese erfragt durch:	rsthbsyrhb				
Anamnese nach Angaben von:	szhrbsrhb				
Aktuelle Anamnese:	<p>zshjnsrthrhb rtthg</p> <p>rtah bare htgb rhb</p>				
HINWEIS: bei stat. Aufnahme evtl. Infektionen in der Umgebung des Kindes/ Inkubationen					
Aktuelle Therapie					
Nahrung					Übernahm.
Medikamente					
Physio-, Ergother. etc.					
Unverträglichkeiten/ Intoleranzen					
Aktuelle Messwerte					
Körpergewicht (kg)	0,00	Taille (cm)	0		
Körpergröße (cm)	0	Hüfte (cm)	0		
Kopfumfang (cm)	0	Blutdruck (mmHg)	0 / 0		
Sitzhöhe (cm)	0	Herzfrequenz (b/min)	0		
Halsumfang (cm)	0	Knochenalter (Jahre)	0		

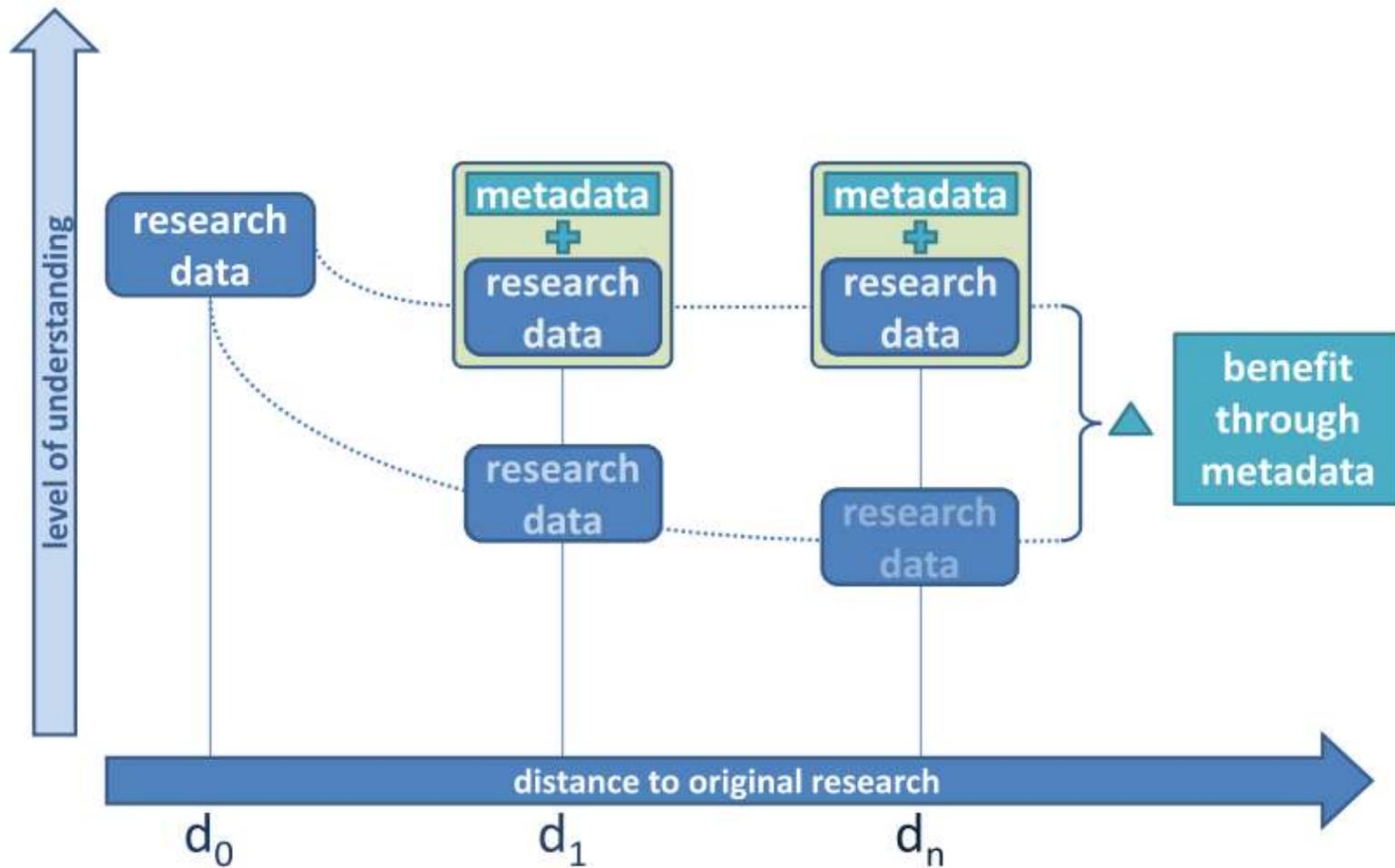
... werden in das MDR importiert, geglättet und harmonisiert

The screenshot shows the MDR Design Studio interface with five data item panels. Colored arrows point from above to each panel, indicating the import and harmonization process. The panels are:

- RE_Alter**: ID: RE_Alter um nr: DataElement@13248. Domain: (no value). Meaning: (no value). Precision: 0. Designations: RE_Alter um nr: Designation@490, Alter um nr: Designation@491.
- RE_Gewicht**: ID: RE_Gewicht um nr: DataElement@332. Domain: (no value). Meaning: (no value). Precision: 0. Designations: RE_Gewicht um nr: Designation@494, Gewicht um nr: Designation@495.
- RE_Groesse**: ID: RE_Groesse um nr: DataElement@11. Domain: (no value). Meaning: (no value). Precision: 0. Designations: RE_Groesse um nr: Designation@487, Größe um nr: Designation@488.
- RE_HerzfrequenzMin**: ID: RE_HerzfrequenzMin um nr: DataElement@572. Domain: (no value). Meaning: (no value). Precision: 0. Designations: RE_HerzfrequenzMin um nr: Designation@572, Min um nr: Designation@573.
- RE_HaematokritMin**: ID: RE_HaematokritMin um nr: DataElement@614. Domain: (no value). Meaning: (no value). Precision: 0. Designations: RE_HaematokritMin um nr: Designation@614, Min um nr: Designation@615.
- Data Item**: Parent: CED-Register. Language: en-US. Designation: Knochenalter. Definition: Value Domain: VD_Jahr um nr: ValueDomain@...

Why Metadata

Research data are only interpretable by those who made them



Quelle: Dickmann, Frank and Grütz, Romanus (2011): LABIMI/F - Digital Preservation of Biomedical Research Data, Knowledge Exchange: Workshop Research Data Management – Activities and Challenges, 14.-15. November 2011, Bonn, Poster, Access date: 2012.01.12, URL: http://www.labimi-f.med.uni-goettingen.de/Publications/Poster_LABIMI-F_v6.pdf.

3