

A National Metadata Repository for Empirical Research in Germany

15th International Open Forum on Metadata



Jürgen Stausberg¹, Johannes Drepper², Matthias Löbe³, Sylvie MN
Nguongo¹, Philippe Verplancke⁴

¹IBE, Ludwig-Maximilians-Universität München, Germany

²TMF - Technology, Methods, and Infrastructure for Networked Medical
Research, Berlin, Germany

³Institute for Medical Informatics (IMISE), Universität Leipzig, Germany

⁴XClinical, München, Germany



Agenda

- + Project
- + Core packages of ISO/IEC 11179 Ed. 3
- + Information models and ISO/IEC 11179 Ed. 3
- + Controlled vocabularies and ISO/IEC 11179 Ed. 3
- + Pilot implementation

Objectives, scope

- + Empirical medical research
 - + is based on the collection of observations stored in DBMS
 - + needs services for the maintenance of item collections and improvement of interoperability:
 - + Reuse of item`s definitions
 - + Quality improvement through harmonization and standardization
 - + Integration and use of controlled vocabularies as value lists for items
 - + Use of terminologies as buildings blocks for data elements
- + MDR-Project in Germany launched in 2009 by the Federal Ministry of Education and Research to set up a national metadata repository for the support of empirical research.
 - + ISO/IEC 11179 Ed. 3 "Information technology - Metadata Registries (MDR)" part 3 "Registry Metamodel and basic attributes"

TMF – Technology, Methods, and Infrastructure for Networked Medical Research



Synergy, quality & standards – shared solutions for health



Oktober 2011: 89 Member organisations

- Disease-oriented Competence Networks in Medicine
- Coordinating Centres for Clinical Trials
- Networks for Rare Diseases
- German Clinical Trials Register
- German National Genome Research Network (NGFN)
- Fraunhofer Institute for Toxicology and Experimental Medicine
- Patient Organisations
- German National Cohort Pre-Test Phase
- Centralized Biobanks
- German Centres of Health Research
- ...



in der Rehabilitation und für Behinderte*

lant Erworbene Pneumonie

ng Charité

onotische Infektionen beim Menschen

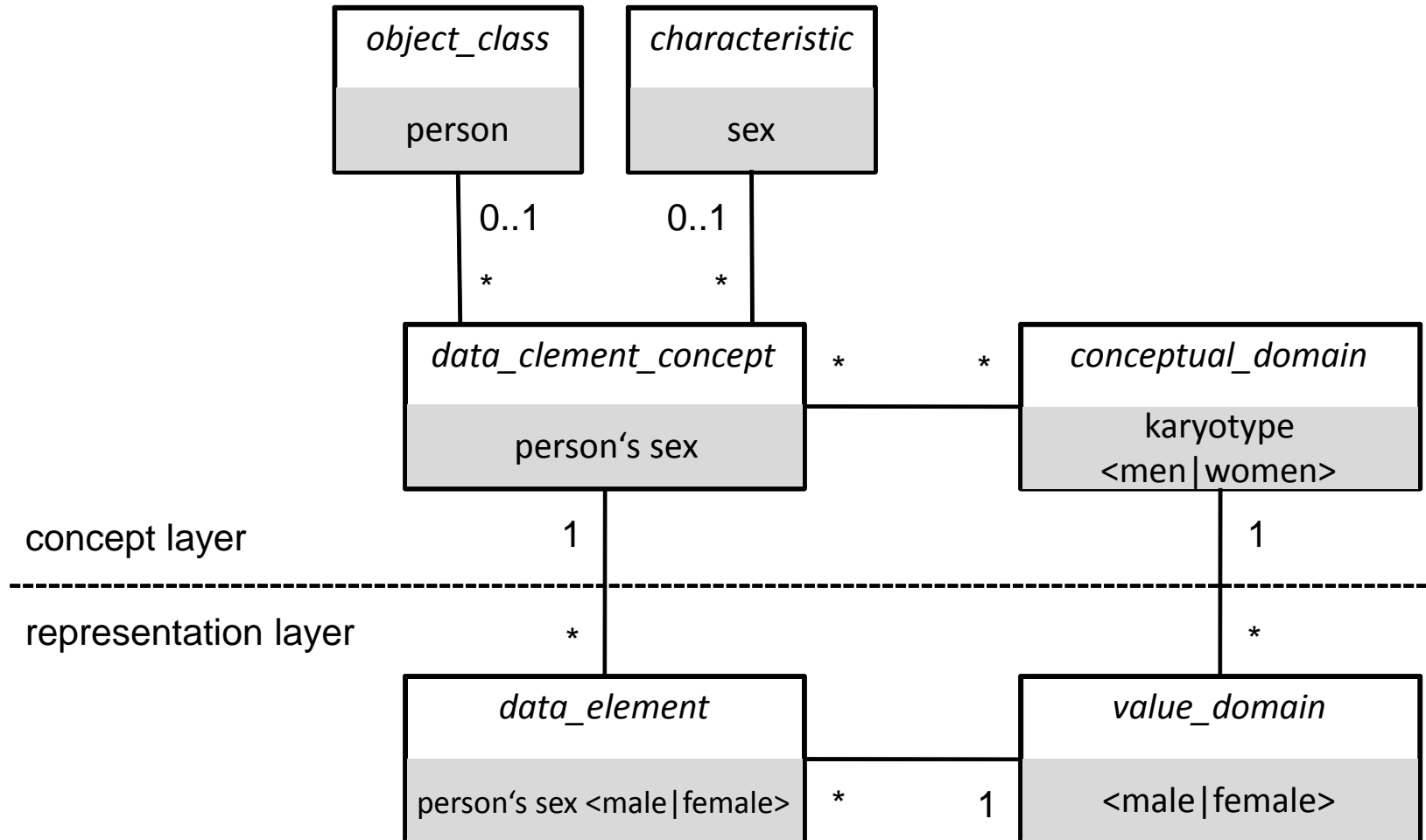
ologie und Hämatologie

che Studien Berlin

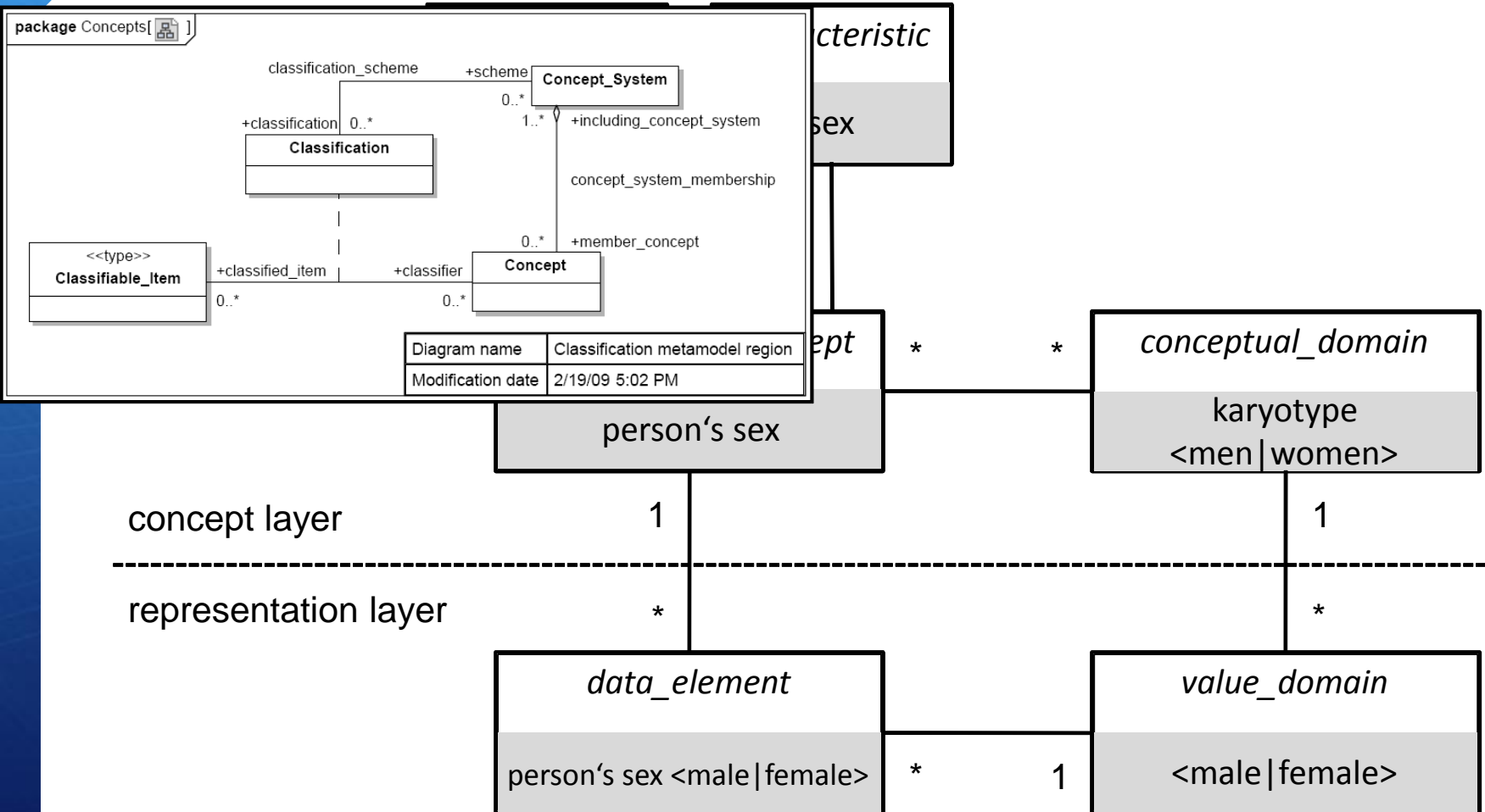
Use cases

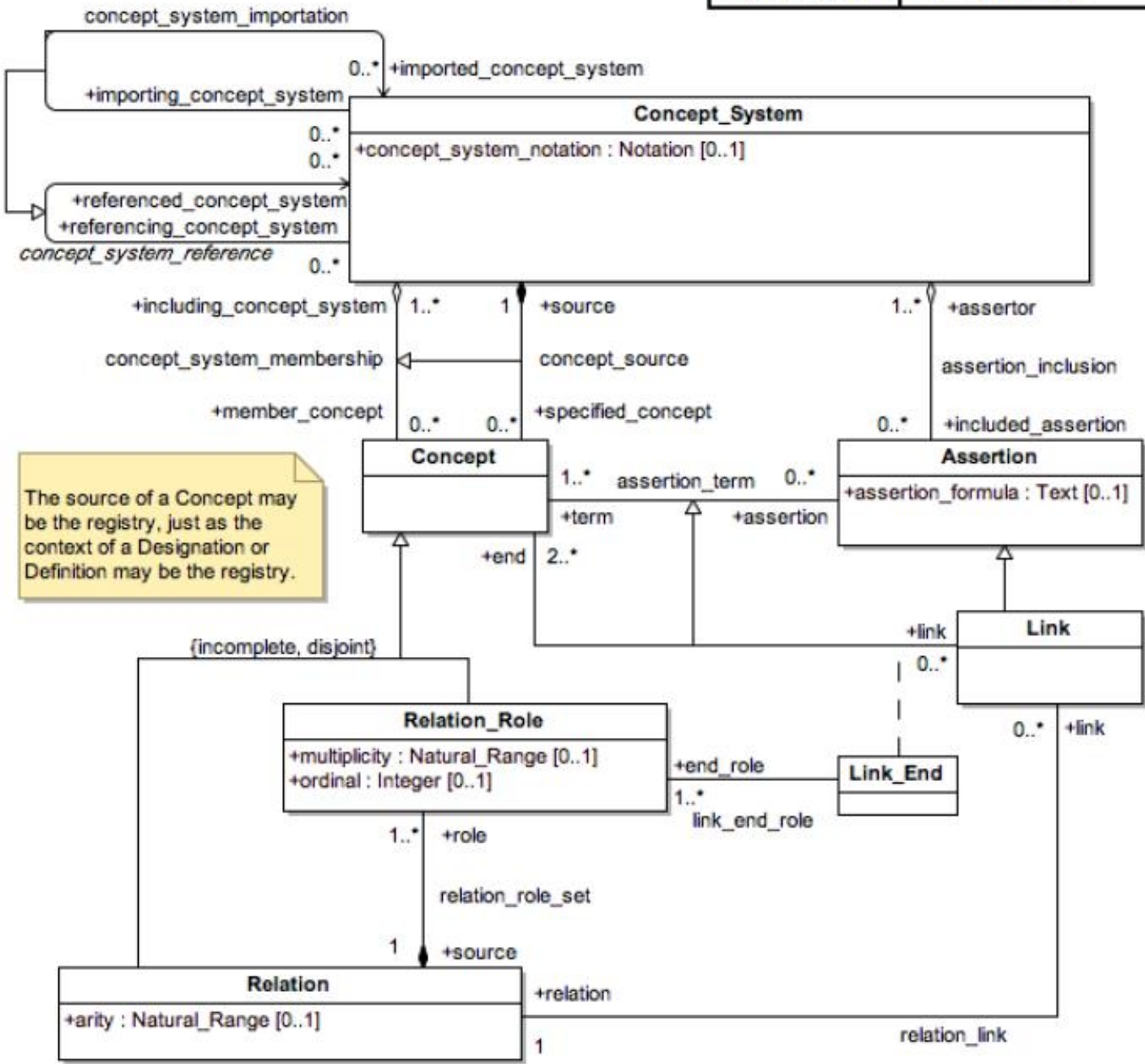
- + Single organization
 - + Maintain metadata - to specify a CRF or a catalogue of variables
 - + Import metadata - to establish a pool of metadata
 - + Mapping of metadata - to transform data from one model to another
 - + Queries for multiple versions of metadata - to support study series
- + Multiple organizations
 - + Queries for multiple metadata - to support pooling of data
 - + Select projects - to identify projects covering specific variables
 - + Review of metadata - to harmonize variables and value lists
- + National level
 - + Queries for multiple metadata - to support pooling of data
 - + Select projects - to identify projects covering specific variables
 - + Maintain standards - to harmonize CRFs and databases
 - + Improve standards - to identify gaps in data standards or terminologies

Core of the metamodel of ISO/IEC 11179 Ed. 3



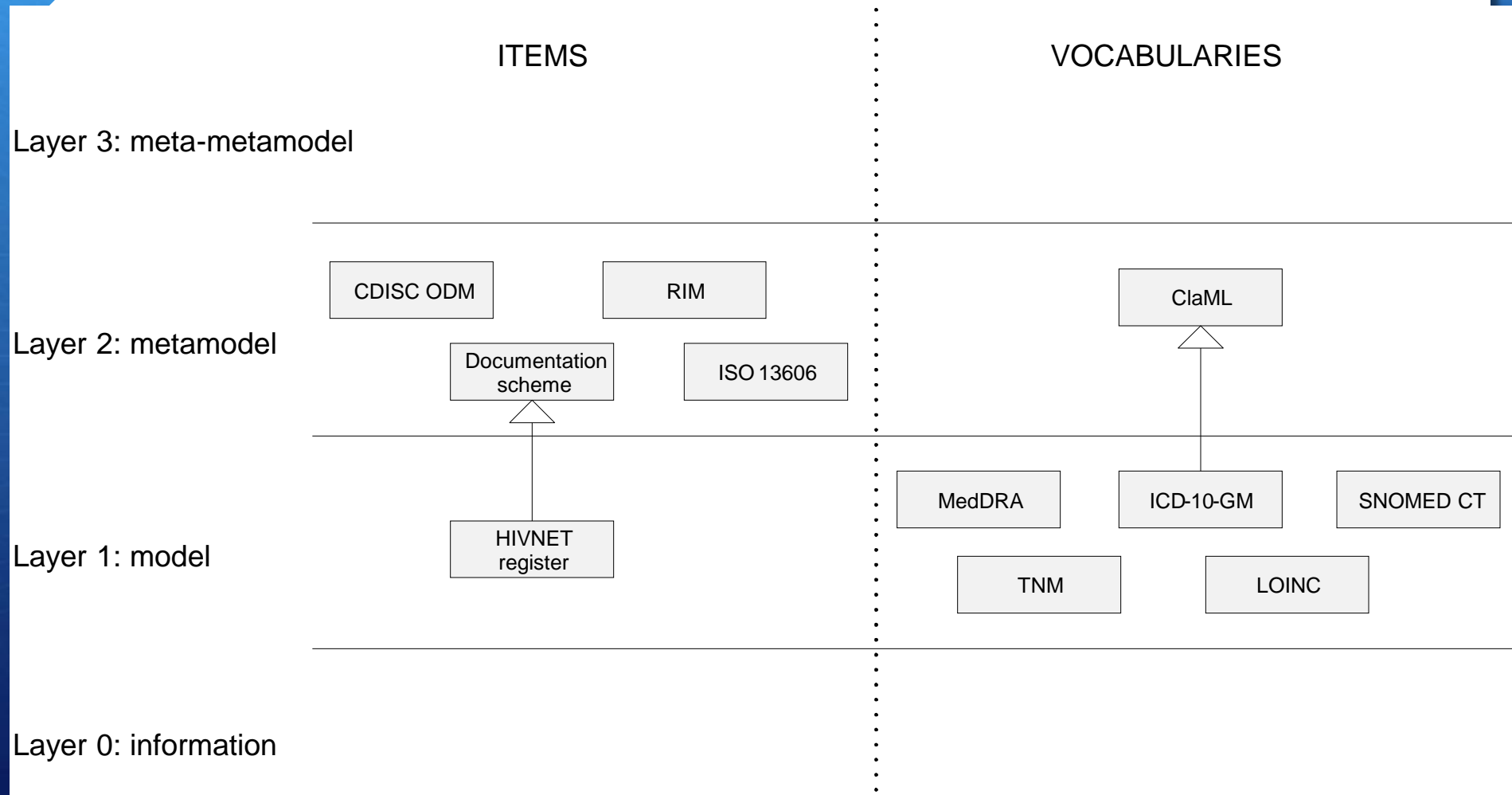
Core of the metamodel of ISO/IEC 11179 Ed. 3



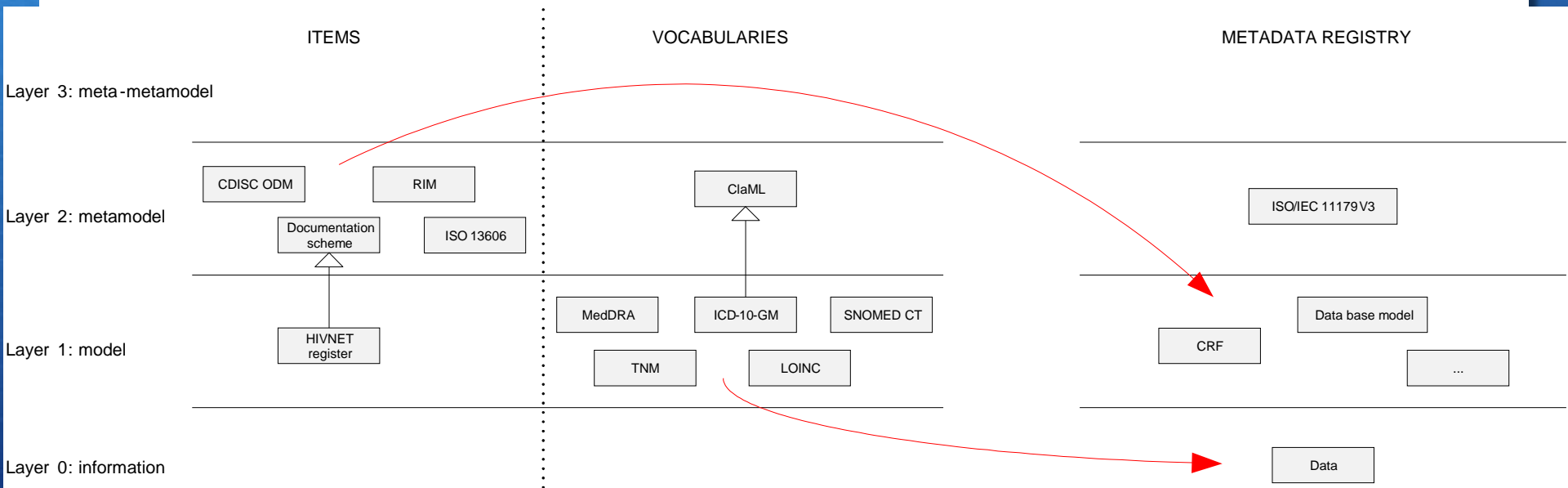


The source of a Concept may be the registry, just as the context of a Designation or Definition may be the registry.

Tested material in the metadata architecture



Mapping to ISO/IEC 11179 V3



Methods

- + Mapping of model elements (metamodels)
- + Import of models into an MDR-database

Documentation scheme/EAV-model

- Disease registers
- + no problem in general
- + but in particular
 - + OIDs
 - + layout information of CRFs
 - + relationships between data_elements

Dokumentationseinheit

Merkmal/Wertebereich/Beschreibung

Vitalstatus

Stammdaten Vitalstatus s. Tabelle [RKI]

Patient

Berufsausbildung, höchste

Stammdaten Berufsausbildung s. Tabelle [BGS98 F082]

Datum der ersten Vorstellung

Monat/Jahr

Datum der letzten Vorstellung

Monat/Jahr

Erklärung Einverständnis

Datum

ethnische Zugehörigkeit

<kaukasisch (weiss)|schwarz|asiatisch|unbekannt>

Geburtsdatum

Monat/Jahr

Geschlecht

Stammdaten Geschlecht s. Tabelle [RKI]

Reference models

➤ EHR

- + RIM of HL7

- + ISO 13606 Health informatics - Electronic health record communication

- + **no problem in general**

- + but

- + inheritance

- + complex information objects not explicitly mentioned

Reference models

➤ EHR

+ RIM of HL7

+ ISO 13606 Health informatics - Electronic health record communication

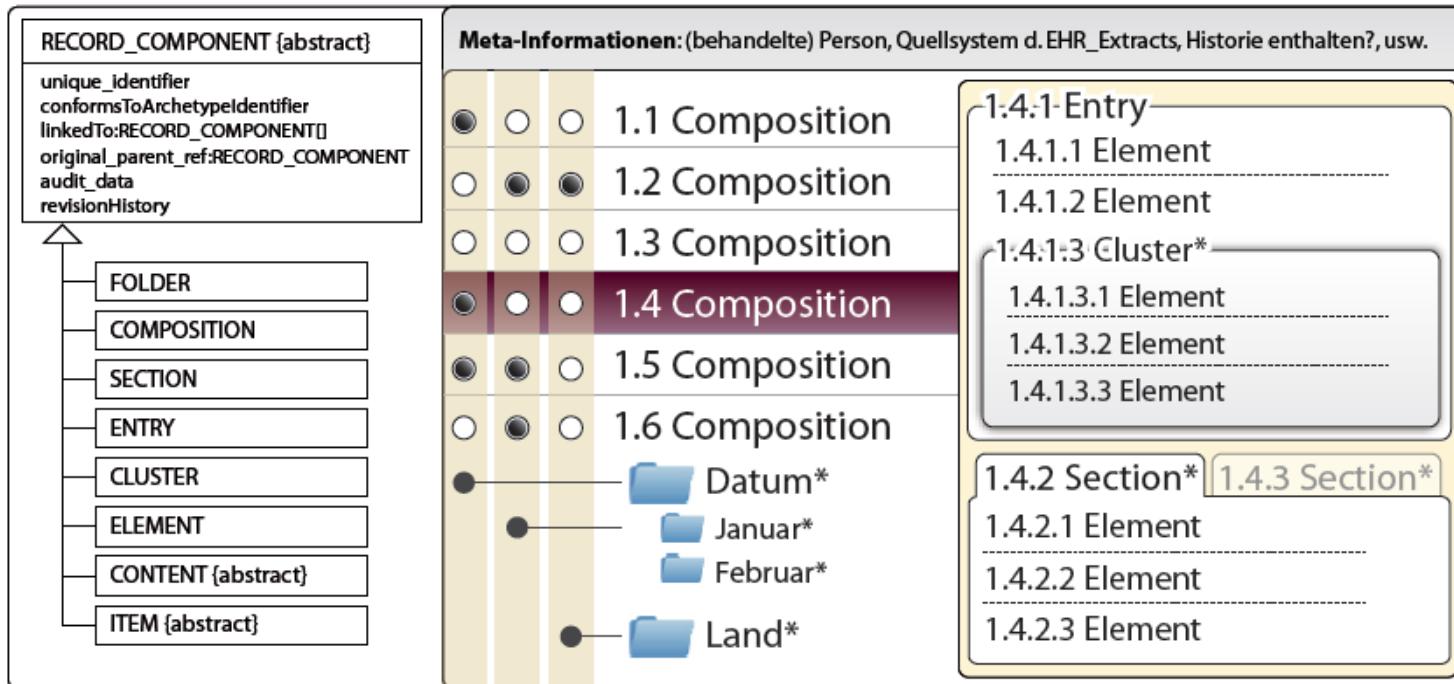
EHR_Extract

+ no prob

+ but

+ inher

+ comp



„CRF“-model 2

- + 1 Context
- + 3 Data_Elements
- + Object_Class: hepatitis B
- + 3 Characteristics: date of diagnosis, reason for diagnosis, first date of raised liver values
- + 3 Value_Domains: 2 date, value list for reason for diagnosis
- + 3 Permissible_Values, 3 Value_Meanings, 3 Conceptual_Domains
- + 22 Designations, 1 Definition
- + Not covered: position (4), constraint between „reason for diagnosis“ and „first date of raised liver values“, order of the Data_Elements, order of the Permissible_Values

4. HBV-Erstdiagnose (HBs-Antigennachweis)

M M * J J J J *

Erstdiagnose durch

- Abklärung von Symptomen oder Leberwerten
- Einstellungsuntersuchung
- andere Untersuchungen

erhöht seit

M M * J J J J *

Mapping CDISC ODM to ISO 11179 Ed. 3

- + CDISC Operational Data Model (ODM): XML-based standard that contains all metadata and clinical data of a clinical trial
- + ODM item definitions are composed of meaning and representation
- + Approach included the basic ODM build blocks for the specification of clinical items (ItemDef, MeasurementUnit, and CodeList)
- + Results: Most elements could be mapped
 - + ODM doesn't have a counterpart for Data Element Concept
 - + Described_Conceptual_Domain/-Value_Domain is implicit in ODM
 - + ISO 11179 permits only a single Measurement Unit where ODM items can have many

Classifications and Terminologies

- + ICD-10-GM: International Statistical Classification of Diseases and Related Health Problems 10th Revision German Modifications
- + OPS: German procedure classification
- + The TNM classification of malignant tumors
 - + A well-established standard for the classification of malignant tumors in oncology. the TNM-system classification provides three axes for post-coordination (T: local tumor, N: regional lymph nodes, M: metastasis).
- + MedDRA: Medical Dictionary for Regulatory Activities
 - + The international medical terminology used to classify adverse events associated with the use of biopharmaceuticals and other medical products.
- + SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms

Material: Classifications and Terminologies - ClaML

- + Classification Markup Language (ClaML) is attractive as interface standard for the import of classifications (ICD-10-GM, OPS) into the MDR.
- + ClaML: EN 14463:2007, Standard for electronic representation of medical classification systems.

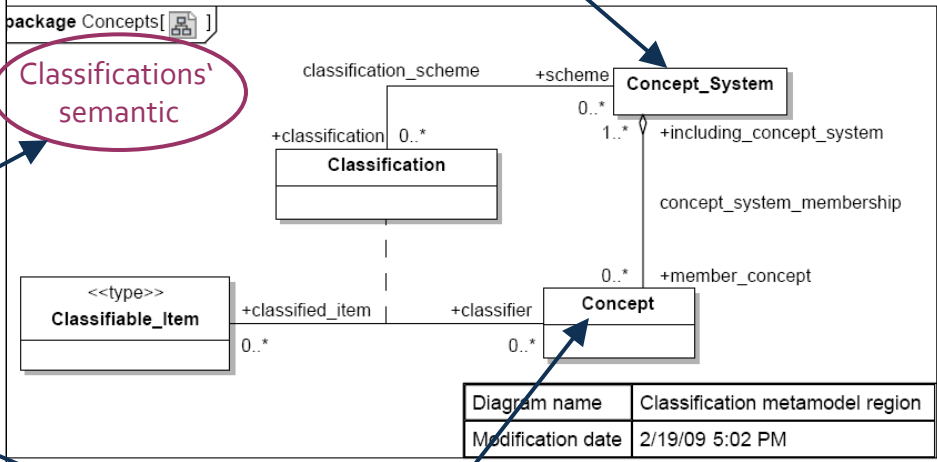
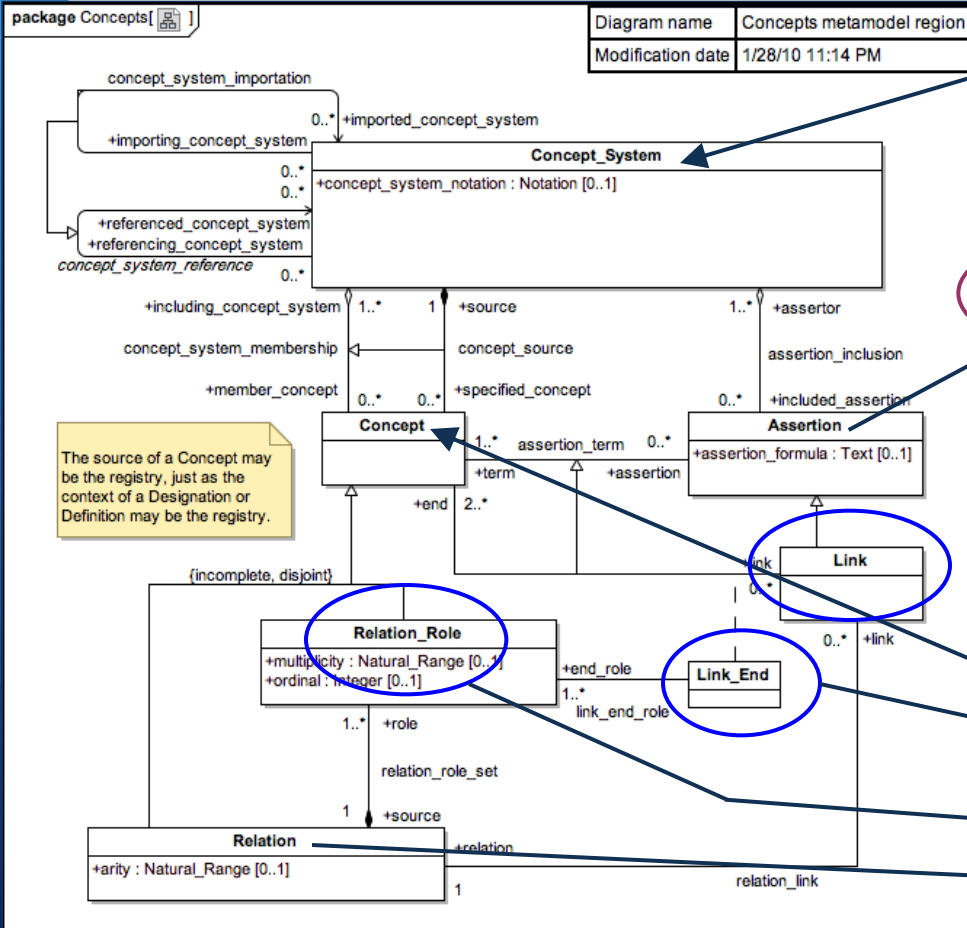
```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE ClaML SYSTEM "claml.dtd"[>
<ClaML version="2.0.0">
  <Meta name="TopLevelSort" value="I II III IV V VI VII VIII IX X XI XII
    XIII XIV XV XVI XVII XVIII XIX XX XXI XXII" />
  <Meta name="lang" value="en" />
  <Meta name="titleLong" value="International Statistical Classification
    of Diseases and Related Health Problems 10th Revision"/>
  <Identifier authority="WHO" uid="SRFSFto be added later" />
  <Title date="2008-05-13" name="ICD-10-2008-EN" version="2008" />
  <Class code="B27" kind="category">
    <SuperClass code="B25-B34" />
    <SubClass code="B27.0" />
    <SubClass code="B27.1" />
    <Rubric id="D0000588" kind="preferred">
      <Label xml:lang="en" xml:space="default">Infectious
        mononucleosis</Label>
    </Rubric>
  </Class>
</ClaML>
```

Mapping : the results

ClAML.root; ICD-10-GM; OPS;
TNM; MedDRA; SNOMED-CT
(Vocabularies)

Classifications'
semantic

The source of a Concept may be the registry, just as the context of a Designation or Definition may be the registry.



ClAML-Classes /
Vocabularies-Classes

Example of ICD-10
Intestinal infectious
Diseases (Aoo-Aog)

generalization
Hierarchy
specialization

ClAML-Classes /
Vocabularies-Classes

Cholera (Aoo)

MIE 2011, Sylvie

Classifications, ClaML

- + some problems:
 - + representation of the code
 - + rules not covered
 - + coding standards
 - + exclusions, inclusions, special codes
 - + reconstruction not possible
 - + intermediate levels not supported
 - + direct use of classes not possible, doubling required

SNOMED CT – core tables

SNOMED CT	ISO 11179 E3	Comment
	concept_system	Value is the release of SNOMED CT.
	context	a) Value is the release of SNOMED CT. b) Value is “Read Codes CTV3” as context for Read Codes.
Concept	concept	
SNOMEDID	definition	Definition is associated with the context of the release of SNOMED CT.
CTV3ID	definition	Definition is associated with the context “Read Codes CTV3”.
IsPrimitive	assertion	
Description	designation	Designation is associated with the context of the release of SNOMED CT.
Term	designation_sign	
DescriptionType	designation_class	Designation_class is an association nclass between designation and context.
LanguageCode	designation_language	
Relationships	link	
RelationshipType	relation_role	The reverse relationship must be generated artificially.
CharacteristicType	assertion_formula	
Refinability		Refinability is not covered by ISO/IEC 11179 E3.
RelationshipGroup		RelationshipGroup is not covered by ISO/IEC 11179 E3.

Terminological systems

- + no problem in general
- + reconstruction to some extent possible
- + but
 - + SNOMED CT elements not covered
 - + relationship_type
 - + refinability for data entry
 - + Medical Dictionary for Regulatory Activities (MedDRA)
 - + versioning of coding systems and their relationship to the data description package
 - + TNM
 - + composition of concepts

Implementation

- + ISO/IEC 11179 Ed. 3 view
- + Data manager view

- Simple
- ISO11179
- Browser
- Source

Soziodemographische Ang

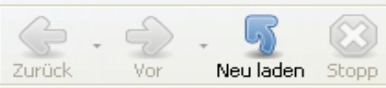


ID: [Soziodemographische A](#) [Context@32781](#)

- Designations
- Definitions
- Relevant Designations

Anzahl noch lebender leiblicher Angehöriger	urn:mdr:Context@32798
Geburtsdatum	urn:mdr:Context@32789
SETHNO	urn:mdr:DataElement@32797
SSEX	urn:mdr:DataElement@32785

- (All)
- characteristic_designation
 - Papierkorb
 - Blutdruck
 - Blutdruck
 - HERZ
 - HERZ
 - HERZ
 - HERZINSUFFIZIENZ
 - Study study1-name
 - Measurement Units
 - Code Lists
 - Code List Item
 - e1-name
 - p1-name
 - m1-name
 - m2-name
 - e2-name
 - p2-name
 - null
 - HERZINSUFFIZIENZ



http://mdr.xclinical.net:8081/



Sign out sn

- Simple
- ISO11179
- Browser**
- Source

Soziodemographische Angaben

Anzahl noch lebender leibliche Geschwisterum:mdr:DataElement@32786	SELTERNum:mdr:DataElement@3
Geburtsdatumum:mdr:Context@32789	SGESCHWum:mdr:DataElement@32790
SETHNOum:mdr:DataElement@32797	SKINDum:mdr:DataElement@32800
SSEXum:mdr:DataElement@32785	



ID: **SGESCHWum:mdr:DataElement@32801**
 Domain: Anzahlum:mdr:ValueDomain@32799
 Meaning: (no value)
 Precision: 0

Designations Definitions

Geschwisterum:mdr:Definition@110

- (All)
- characteristic_designationum:mdr:Characteristic@109
- Papierkorbum:mdr:Context@109
- Blutdruckum:mdr:Context@110
- Blutdruckum:mdr:Context@111
- Herzum:mdr:Context@118
- Herzum:mdr:Context@119
- Herzum:mdr:Context@120
- HERZINSUFFIZIENZum:mdr:Context@1
- Study study1-nameum:mdr:Context@13
- Measurement Unitsum:mdr:Context@13
- Code Listsum:mdr:Context@131107
- Code List ItemsSum:mdr:Context@131108
- e1-nameum:mdr:Context@131113
- p1-nameum:mdr:Context@131114
- m1-nameum:mdr:Context@131115
- m2-nameum:mdr:Context@131119
- e2-nameum:mdr:Context@131122
- p2-nameum:mdr:Context@131123
- nullum:mdr:Context@131124
- HERZINSUFFIZIENZum:mdr:Context@1
- Herzum:mdr:Context@16
- Testbogenum:mdr:Context@18
- Testbogenum:mdr:Context@19
- Testbogen eNum:mdr:Context@20
- Test1um:mdr:Context@21
- IMISEum:mdr:Context@24
- IMISE:DataTypesum:mdr:Context@25



- Simple
- ISO11179
- Browser
- Source

Code Lists: cl1-name

cl1-nameum:mdr:ValueDomain@131109



ID: cl1-nameum:mdr:ValueDomain@131109

Meaning: (no value)

Data Type: integerum:mdr:DataType@131083

Format:

Maximum Character Quantity:

Unit of Measure: (no value)

Designations Definitions Members

- 0um:mdr:PermissibleValue@79
- 1um:mdr:PermissibleValue@80
- 2um:mdr:PermissibleValue@81

- (All)
- characteristic_designationum:mdr:Char
- Papierkorbum:mdr:Context@109
- Blutdruckum:mdr:Context@110
- Blutdruckum:mdr:Context@111
- Herzum:mdr:Context@118
- Herzum:mdr:Context@119
- Herzum:mdr:Context@120
- HERZINSUFFIZIENZum:mdr:Context@1
- Study study1-nameum:mdr:Context@13
- Measurement Unitsum:mdr:Context@13
- Code Listsum:mdr:Context@131107
- Code List ItemsSum:mdr:Context@131108
- e1-nameum:mdr:Context@131113
- p1-nameum:mdr:Context@131114
- m1-nameum:mdr:Context@131115
- m2-nameum:mdr:Context@131119
- e2-nameum:mdr:Context@131122
- p2-nameum:mdr:Context@131123
- nullum:mdr:Context@131124
- HERZINSUFFIZIENZum:mdr:Context@1
- Herzum:mdr:Context@16
- Testbogenum:mdr:Context@18
- Testbogenum:mdr:Context@19
- Testbogen eNum:mdr:Context@20
- Test1um:mdr:Context@21

- Simple
- ISO11179
- Browser
- Source

cl1-name



Parent: Code Lists

Language: en-US

Designation: cl1-name

Definition:

Data Type: integerum:mdr:DataType@131083

Format:

Length: 0

Unit of Measure: (no value)

Members:

2	c	
0	a	
1	b	

- (All)
- characteristic_designationum:mdr:Cha...
 - Papierkorbum:mdr:Context@109
 - Blutdruckum:mdr:Context@110
 - Blutdruckum:mdr:Context@111
 - HerZum:mdr:Context@118
 - HerZum:mdr:Context@119
 - HerZum:mdr:Context@120
 - HERZINSUFFIZIENZum:mdr:Context@1...
 - Study study1-nameum:mdr:Context@13...
 - Measurement Unitsum:mdr:Context@13...
 - Code Listsum:mdr:Context@131107
 - Code List ItemsSum:mdr:Context@131108
 - e1-nameum:mdr:Context@131113
 - p1-nameum:mdr:Context@131114
 - m1-nameum:mdr:Context@131115
 - m2-nameum:mdr:Context@131119
 - e2-nameum:mdr:Context@131122
 - p2-nameum:mdr:Context@131123
 - nullum:mdr:Context@131124
 - HERZINSUFFIZIENZum:mdr:Context@1...
 - HerZum:mdr:Context@16
 - Testbogenum:mdr:Context@18
 - Testbogenum:mdr:Context@19
 - Testbogen enum:mdr:Context@20
 - Test1um:mdr:Context@21

Conclusions

- + ISO/IEC 11179 Ed. 3 is powerful and contributes many useful ideas for the definition of a national MDR.
- + An extension of ISO/IEC 11179 Ed. 3 metamodel might be necessary to meet the predefined needs of a national MDR.
- + The relationship between data elements and concepts should be further elaborated.
- + The implementation for data managers requires a simplified view around the ISO/IEC 11179 Ed. 3 model.
- + For clinical studies, ODM could be the simplified model.

Acknowledgements

- + The presented work is part of the project MDR - Metadata Repository funded by the German Federal Ministry of Education and Research (BMBF).