Data models
standards to define, exchange and archive data

MedInfo 2013, Copenhagen
HL7 CDA

- CDA = Clinical Document Architecture
- Defined by HL7
- Available at http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7
- Version 1: November 2000

- Specification of structure and semantics
- XML based
- Usage of HL7 RIM (Reference Implementation Model) data types

- Purpose: exchange data between providers and patients
HL7 CDA - Structure

- CDA-Header: Meta information
- CDA-Body: Document itself
  - Level 1: textual information (discharge letters)
  - Level 2: separation and annotation of segments
  - Level 3: computable data (lab values)
Archetypes

- Mentioned by T. Beale 2002
- Core part of EN 13606
- Actively researched and published by D. Kalra [1]
- Commonly known: openEHR archetypes
- Definition in Archetype Definition Language (ADL)
- Available at http://www.openehr.org

- Purpose: Definition of small (10-50), domain specific unit
- Re-usable, combinable
- Based on reference model

Archetypes

• Commonly known: openEHR archetypes

Figure 1: Blood Pressure, Unstructured

Figure 2: Blood Pressure, Big Model

Figure 3: Blood Pressure, Generic Model
ODM

- Defined by CDISC
- Available at http://www.cdisc.org/odm
- Current version 1.3.1: 2010, only registered users
- XML schema definition (ODM1-3-1.xsd and ODM1-3-1-foundation.xsd)

- Purpose: clinical studies (Study events + patient data)
- Study Events (Visit, Follow-Up, ...) = CRFs
- Principle: Definitions ("Def") and References ("Ref")
ODM – important elements

```xml
<FormDef OID="F.0000" Name="Medical History">
  <Description>
    <TranslatedText xml:lang="en">Medical History</TranslatedText>
    <TranslatedText xml:lang="de">Anamnese</TranslatedText>
  </Description>
  <ItemGroupOID=
```
ODM – important elements

```xml
<FormDef OID="F.0000" Name="Medical History"
  Repeating="No">
  <Description>
    <TranslatedText xml:lang="en">
      Medical History
    </TranslatedText>
  </Description>
</FormDef>
```

```xml
<FormDef OID="F.0000" Name="Medical History">
  <TranslatedText xml:lang="en">
    Medical History
  </TranslatedText>
</FormDef>
```
ODM – important elements

<FormDef OID="F.0000" Name="Medical History"
Repeating="No">
  <Description>
    <TranslatedText xml:lang="en">Medical History</TranslatedText>
    <TranslatedText xml:lang="de">Anamnese</TranslatedText>
  </Description>
</FormDef>

OID=
Repeating="No"F.0000" Name="Medical History"
Description=TranslatedText

xml:lang="en"
ODM – important elements (2)

```xml
<ItemGroupDef_OID="IG.0000" Name="Admin Dates" Repeating="No">
  <ItemRef ItemOID="I.0000" Mandatory="Yes" OrderNumber="1"/>
  <ItemGroupRef ItemOID="I.0001" Mandatory="Yes" OID="IG.0000" Name Repeating="No">
    "Admin Dates"
  </ItemGroupRef>
</ItemGroupDef>
```
ODM – important elements (3)

<ItemDef OID="I.0000" Name="Admission Date"
  DataTypes="date">
  <Question>
    <TranslatedText xml:lang="en">Admission date</TranslatedText>
    <TranslatedText xml:lang="de"/>
    <Alias Context="Code-1" Name="Date of admission"/>
    = "UMLS CUI-1" Name = "C1302393"
ODM – important elements (3)

```xml
<ItemDef OID="I.0000" Name="Admission Date"
  DataType="date"
  <Question>
    <TranslatedText xml:lang="en">Admission date</TranslatedText>
    <TranslatedText xml:lang="de">Aufnahmedatum</TranslatedText>
    <TranslatedText xml:lang="en">Date of admission</TranslatedText>
    <Alias Context="Code-1" Name="Date of admission" />
  </Question>
</ItemDef>
```
Contact

Benjamin.Trinczek@uni-muenster.de
Bernhard.Breil@uni-muenster.de
dugas@uni-muenster.de

Institute of Medical Informatics
University Münster
http://imi.uni-muenster.de
Portal of Medical Data Models – An open repository of Forms

MedInfo 2013, Copenhagen
Portal of Medical Data Models

- Multilingual Medical Data Models with structured information and coded elements
  - 3,800 Forms
  - 120,000 Data elements

# Medical Data Models

## Portal of Bernhard Breil

### My observations:

<table>
<thead>
<tr>
<th>ODM Files</th>
<th>Unpublished files</th>
<th>Ratings</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### My contributions:

<table>
<thead>
<tr>
<th>ODM Files</th>
<th>Unpublished files</th>
<th>Ratings</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

### Latest comments:

- by Martin Dugas        2012-12-06
- by Martin Dugas        2012-12-03
- by Martin Dugas        2012-12-03
- by Martin Dugas        2012-12-03
- by Bernhard Breil      2012-09-21
- by Dr. Vojtech Huser   2012-09-04
- by Dr. Vojtech Huser   2012-09-04
- by Dr. Vojtech Huser   2012-09-04
- by Dr. Vojtech Huser   2012-09-04

## Portal of Bernhard Breil

### Best rated forms:

<table>
<thead>
<tr>
<th>Follow Up</th>
<th>Eligibility NCT0332621</th>
<th>Diabetic N...</th>
<th>★★★★★</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eligibility NCT00508723</td>
<td>Acute Myel...</td>
<td>★★★★★</td>
</tr>
<tr>
<td></td>
<td>CD4 discharge letter VHigh 1.50 CD...</td>
<td>★★★★★</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 blood pressure any condition</td>
<td>★★★★★</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HIS Review of Systems</td>
<td>★★★★★</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ankylosing Spondylitis NCT00530589</td>
<td>On-Study ...</td>
<td>★★★★★</td>
</tr>
<tr>
<td></td>
<td>Breast Cancer NCT00053339</td>
<td>Treatment...</td>
<td>★★★★★</td>
</tr>
<tr>
<td></td>
<td>Breast Cancer NCT0132992</td>
<td>Registration...</td>
<td>★★★★★</td>
</tr>
<tr>
<td></td>
<td>Gynecological Cancer NCT00883062</td>
<td>On Tri...</td>
<td>★★★★★</td>
</tr>
</tbody>
</table>

### Medical Data Models currently offers you:

<table>
<thead>
<tr>
<th>ODM Files</th>
<th>Forms</th>
<th>Itemgroups</th>
<th>Items</th>
<th>Keywords</th>
<th>Ratings</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3459</td>
<td>3474</td>
<td>17172</td>
<td>105108</td>
<td>21</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

### Latest forms:

- Eligibility NCT01033997 | Acute Demyelinating Spectrum Disease 2012-12-04
- Eligibility NCT01031680 | Type 2 Diabetes Mellitus 2012-12-04
- Eligibility NCT01026142 | Breast Cancer 2012-12-04
- Eligibility NCT01021670 | Sexual Dysfunction 2012-12-04
- Eligibility NCT01019473 | Huntington’s Disease 2012-12-04
- Eligibility NCT01019174 | Multiple Myeloma 2012-12-04
- Eligibility NCT01018069 | Leukemia 2012-12-04
- Eligibility NCT01016275 | Bladder Lesions 2012-12-04
- Eligibility NCT01012973 | Renal Neoplasia 2012-12-04
- Eligibility NCT01011729 | Hepatitis ... 2012-12-04
Presentation

- Overview
  - Focus on item groups, names
  - Form in a preview

- Detailed view
  - Item groups, names
  - Data types
  - Concept codes
Vocabularies

- Systematized Nomenclature of Medicine (SNOMED)
- Unified Medical Language System (UMLS)
- Logical Observation Identifiers Names and Codes (LOINC)
- International Classification of Diseases (ICD)
- International Classification of Procedures in Medicine (ICPM)
Follow Up

Follow Up Dokumentation zur automatischen Berichtigung von Kaplan-Meier Kürzen.

Form family: Follow Up
Version: 4
Copyright: none
Created at: 2012-10-05
Created by: Bernhard Breil
Change owner:
Average rating:
Your rating:
Keywords: Cancer

Comments (1)

Martin Dugas 4 languages 2012-12-03 available (english, german, italian, portuguese)

Download

Follow-Up

Identity

Name: Identity
Description: Identity

<table>
<thead>
<tr>
<th>Item</th>
<th>Datatype</th>
<th>Aliases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Last Name</td>
<td>text</td>
<td></td>
</tr>
<tr>
<td>Patient First Name</td>
<td>text</td>
<td></td>
</tr>
<tr>
<td>Patient DateOfBirth</td>
<td>partialdate</td>
<td></td>
</tr>
</tbody>
</table>

Martin Dugas, Bernhard Breil, Benjamin Trinczek, Martin Dugas
Multilingual

- Multilingual
  - System
    - English
    - German
  - Form
    - Currently up to 7 languages in use
Multilingual
Download Options

- Publication and distribution of data models in different formats:
  - SPSS
  - SQL
  - CSV
  - R
  - ADL
  - PDF
Download Options

**OP-Personal**

- **Anesthesiologist**
- **Surgeon Staff**

**Weitere Angaben für Urologie-Ops**

**Intraoperative**

- **Specimen**
- **Length of surgical procedure**
- **Volume Prostate**
- **estimated blood loss**
- **Erythrocyte Transfusion**

- **not coded**

- **Lesion of rectum**
  - Yes
  - No

- **Obturator nerve lesion**
  - Yes
  - No
From Open Source to Open Data Models

- Open data models allow
  - Reuse of forms (standard format)
  - Rating of forms
  - Discussion of forms on three levels
    - Form
    - Item group
    - Item
Comment Forms – Item Group Level

Identity

- Patient Last Name
- PatientFirstName
- PatientDayOfBirth

Comments (0)

Create Itemgroup comment
Join us and Contribute
Upload ODM-files

New form

Please select the file you want to upload. In the following steps you will be able to check the contained data and to add additional information. Please be aware, that you are only allowed to upload files of which you are the copyright owner.

Durchsuchen odm.xls

Please select the ODM-version of your file. If you experience problems, feel free to skip the validation by selecting "no check".

ODM version 1.3.1

Continue

You can create ODM-files from Excel-Templates with this webtool.
Convert Excel files into ODM

Convert Excel file into ODM format (xlsx2ODM)

Please use this Excel template to describe Your medical form / data model.
An example for a completed form can be found here.

Upload Your excel file

Durchsuchen... Keine Datei ausgewählt. Daten absenden

Goto form upload
Excel Template

<table>
<thead>
<tr>
<th>StudyOID or FormsetOID</th>
<th>Your FormID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sponsor</td>
<td>Name of Study Sponsor</td>
</tr>
<tr>
<td>Condition</td>
<td>Disease</td>
</tr>
<tr>
<td>StudyName</td>
<td>Name of Study</td>
</tr>
<tr>
<td>StudyDescription</td>
<td>description of study</td>
</tr>
<tr>
<td>Form</td>
<td>form category, for example eligibility form, EHR form, CRF</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>en</th>
<th>de</th>
</tr>
</thead>
<tbody>
<tr>
<td>itemgroup</td>
<td>Personal Information</td>
<td>Personal Information</td>
<td>Persönliche Informationen</td>
</tr>
<tr>
<td>string</td>
<td>Last Name</td>
<td>Last name</td>
<td>Nachname</td>
</tr>
<tr>
<td>string</td>
<td>First Name</td>
<td>First name</td>
<td>Vorname</td>
</tr>
<tr>
<td>date</td>
<td>Date of Birth</td>
<td>Date of Birth</td>
<td>Geburtstag</td>
</tr>
<tr>
<td>itemgroup</td>
<td>Medical History</td>
<td>Medical history</td>
<td>Anamnese</td>
</tr>
<tr>
<td>string</td>
<td>diagnosis</td>
<td>diagnosis</td>
<td>Diagnose</td>
</tr>
<tr>
<td>date</td>
<td>diagnosis date</td>
<td>diagnosis date</td>
<td>Diagnosedatum</td>
</tr>
</tbody>
</table>

Bernhard Breil, Benjamin Trinczek, Martin Dugas
## Assign Keywords

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thrombosis</td>
<td>(1)</td>
</tr>
<tr>
<td>Gynaecology</td>
<td>(0)</td>
</tr>
<tr>
<td>Orthopaedics</td>
<td>(0)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>(27)</td>
</tr>
<tr>
<td>Cancer</td>
<td>(171)</td>
</tr>
<tr>
<td>Heart Failure</td>
<td>(4)</td>
</tr>
<tr>
<td>Stroke</td>
<td>(11)</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>(12)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>(3)</td>
</tr>
<tr>
<td>Epilapsy</td>
<td>(4)</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>(15)</td>
</tr>
<tr>
<td>Heart Disease</td>
<td>(6)</td>
</tr>
<tr>
<td>COPD</td>
<td>(6)</td>
</tr>
<tr>
<td>Arthritis</td>
<td>(2)</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>(5)</td>
</tr>
<tr>
<td>Asthma</td>
<td>(5)</td>
</tr>
<tr>
<td>Dermatology</td>
<td>(14)</td>
</tr>
</tbody>
</table>
Form Comparison
Contact

Benjamin.Trinczek@uni-muenster.de
Bernhard.Breil@uni-muenster.de
dugas@uni-muenster.de

Institute of Medical Informatics
University Münster
http://imi.uni-muenster.de
Open Metadata for Medical Data Models

Martin Dugas\textsuperscript{a}, Bernhard Breil\textsuperscript{a}, Benjamin Trinczek\textsuperscript{a}, Jürgen Stausberg\textsuperscript{b}, Dipak Kalra\textsuperscript{c}

\textsuperscript{a} Institute of Medical Informatics, University of Münster, Germany
\textsuperscript{b} Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie, Ludwig-Maximilians-Universität München, Germany
\textsuperscript{c} Centre for Health Informatics and Multiprofessional Education, University College London, UK

At present, many information systems in healthcare and medical research are based on heterogeneous, secret and proprietary data models. Standards and open metadata can facilitate data exchange between systems and therefore contribute to semantic interoperability. The objective of the workshop is to discuss various state-of-the-art approaches regarding open metadata for medical data models. In particular, CDISC-ODM, which is commonly used in EDC systems, and openEHR will be used to cover routine care and research settings. In addition, this discussion will be complemented by hands-on training regarding open medical data models.
INTRODUCTION & OBJECTIVES OF THE WORKSHOP
The Current Documentation Landscape

- Electronic Health Records (EHRs)
  - 25% - 40% of physicians daily workload
  - Proprietary data models for each hospital
  - > 100 documentation forms per hospital
    with uncontrolled redundancy

- Research documentation (EDC)
  - Large and increasing amount of CRFs per study:
  - Large overlap with routine documentation,
    especially to exclude adverse events
Medical Data Models are complex and heterogeneous ...

and >99% are secret !!!
>99% of Medical forms are secret

<table>
<thead>
<tr>
<th>Differentialblutbild</th>
<th>Datum: <strong>/</strong>/___</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Haemoglobin</strong></td>
<td>mmol/l or g/dl</td>
</tr>
<tr>
<td><strong>Leukozyten</strong></td>
<td>GPl/l /μl 10^3/μl /nl</td>
</tr>
<tr>
<td><strong>Neutrophile</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Stabkernige</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Segmentkernige</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Lymphozyten</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Monozyten</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Eosinophile</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Basophile</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Blasten</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Monocytes</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Myelozyten</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Metamyelozyten</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Sonstige Zellen</strong></td>
<td>%</td>
</tr>
<tr>
<td><strong>Thrombozyten</strong></td>
<td>GPl/l /μl 10^3/μl /nl</td>
</tr>
</tbody>
</table>
How can we tackle the Global Documentation Crisis?

- **Step 1: Open Metadata**
  - Transparency of Medical Data Models
  - Free access to Medical Forms: EHR forms & CRFs

- **Step 2: Consolidation of Medical Data Models**
  - Evolving process: Discussion & Consensus
  - Sharing best practice
  - Semantic annotation
Open Metadata - What Can You Do?

- As scientist: Publish Your Metadata – make it usable for the scientific community

- As reviewer: Demand Open Metadata for scientific publications

- As customer: Demand free access to medical forms from EHR and EDC vendors

- As citizen: Claim transparency - informed consent implies access to Metadata
Workshop Objectives

- Brief Introduction to Standards
- Introduction Portal of Medical Data Models
- Hands On Training: Creation of Medical Data Model
- Discussion
INTRODUCTION OF WORKSHOP PARTICIPANTS

- BACKGROUND
- EXPECTATIONS