Integration of classifications and terminologies in metadata registries based on ISO/IEC 11179

Sylvie Ngouongo
Jürgen Stausberg

Ludwig-Maximilian-University of Munich, Germany
Institute of Medical Informatics, Biometry and Epidemiology (IBE)

Presented by Sylvie Ngouongo
Agenda

• Introduction
• Objectives
• Material
  – ISO/IEC 11179 V3
  – Classifications and Terminologies
• Methods and Results
  – Mapping
  – Import
• Discussion
• Conclusion
Introduction

- Empirical medical research
  - is based on the collection of observations stored in DBMS;
  - needs services for the maintenance of item collections and improvement of semantic interoperability:
    - Reuse of item`s definitions
    - Quality improvement through harmonization and standardization,
    - Integration and use of controlled vocabularies as value lists for items.

- 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 V3
Objectives

- Appropriateness of ISO/IEC 11179 V3 “Information technology - Metadata Registries (MDR)” part 3 “Registry Metamodel and basic attributes” for a national MDR in Germany.

- Possibilities of including well-established classifications and terminologies (controlled vocabularies) using the metamodel of ISO/IEC 11179 V3.
ISO/IEC 11179 describes the method of standardizing and registering of data elements to make data understandable and sharable.

Data Description Package with examples in grey
Material: ISO/IEC 11179 V3

Concept Metamodel Region

Concept_System Metamodel Region

Classification Metamodel Region
Material: Classifications and Terminologies

- **ICD-10-GM**: International Statistical Classification of Diseases and Related Health Problems 10th Revision German Modifications

- **OPS**: German procedure classification


- **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, a multilingual clinical healthcare terminology, essential for electronic health records.
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.

```xml
<?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="SRFSF to be added later"/>
  <Title value="2008-05-13" name="ICD-10-2008-EN" version="2008"/>
  <Class code="B27" kind="category">
    
    
  </Class>
</ClaML>
```

ClaML and its structure's elements
Methods

- Mapping
  - Approach and results
- Import
  - Approach and results
Mapping: the approach

- Mapping of ClaML Elements to elements of ISO/IEC 11179 V3 Metamodel.
- Mapping of vocabularies structures to elements of the ISO/IEC 11179 V3 metamodel.
Mapping: the results

Concept_System Metamodel Region

Classification Metamodel Region

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

Classification’s semantic

ClaML-Classes / Vocabularies-Classes

Example of ICD-10

Intestinal infectious Diseases (A00-A09)

Cholera (A00)

Hierarchy

specialization

generalization

ClaML-Classes / Vocabularies-Classes

The source of a Concept may be the registry, just as the context of a Designation or Definition may be the registry.
Import: approach and results

- MDR_schema.xsd
- Used as Namespace
- DB: any database system of choice (Oracle, MySQL, MS Access, ...)

1. Direct import from native database structures (manual reorganization of classifications’ data needed to meet MDR-Structure)
2. XSLT
3. MDR valid Interface for XML, CSV or SQL Import

DB
MDR Prototype based on ISO 11179 V3

Direct import from native XML-Structures (discrepancies in XML-Schemas)

ICD-10/OPS in Intermediate format (ClamL, XML)

ICD-10-GM / OPS
Discussion

- Representation of classifications and terminologies as sources for value lists of items is quite simple in ISO/IEC 11179 V3, however...
  - Terminologies (MedDRA, SNOMED-CT) could be easily represented both in the Concept Metamodel Region and the Data Description Package of ISO/IEC 11179 V3,
  - Representation of Classifications (ICD-10, OPS, TNM) in the Data Description Package is not satisfactory yet,
  - Classifications’ rules (inclusions, exclusions...) and exceptions could not be properly mapped in the Concept Package yet.

- Full coverage of the item semantic is not reached yet, Implementation on logic level of application needed.
Example of discrepancies between WHO and DIMDI regarding ClaML

WHO applies `<Modifier>` (subclassifications’ coding) for 4-digit ICD-Classes to their superclasses (3-digit classes) whereas DIMDI applies them to the corresponding 4-digit classes to build the 5-digit subclassifications. M01 (3-digit) → M01.0 (4-digit) → M01.00 (5-digit)
Conclusion

- ISO/IEC 11179 V3 is powerful and contributes many useful ideas for the definition of a national MDR.
- An extension of ISO/IEC 11179 V3 metamodel to meet the predefined needs of a national MDR might be necessary.
- A unique interpretation of ClaML by Healthcare organizations (WHO, DIMDI, ...) is essential to continue work on the ClaML import interface.
Acknowledgements

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

- Support of empirical medical research through Maintenance of item collections:
  - Reuse of item’s definitions,
  - Reviews of item collections through harmonization and standardization,
  - Integration and use of controlled vocabularies as value lists for items.
- ISO 11179 is a method of standardizing and registering of data elements to make data understandable and sharable. It is powerful and contributes many useful ideas for the definition of a national MDR.
- Representation of classifications and terminologies as sources for value lists of items is quite simple in ISO 11179 V3. Problems arise when trying to represent classification’s rules (inclusions, exclusions...).
- An extension of ISO 11179 V3 metamodel to meet the predefined needs of a national MDR might be necessary.
- A unique interpretation of the ClaML-Standard by healthcare organizations is required to implement a common ClaML import interface.