Smart Open Services for European Patients

Open eHealth initiative for a European large scale pilot of

Patient Summary and Electronic Prescription

Work Package 3.5 - Semantic Services

Appendix D - epSOS Master Value Set Catalogue

D3.5.2

February 10th, 2010

Document Version: 0.0.3

WORK PACKAGE	**WP 3.5**
DOCUMENT VERSION	**0.0.3**
DATE	**18/02/2010**

Table of contents

For:

- Document Information
- Sub-Project Identification
- History of Alteration
- Referring Documents

Please refer to the main document:

"Semantic Services Definition D3.5.2" D3.5.2_epSOS_WP3_5_v0.0.5_20100223.doc

Appendix D - epSOS Master Value Set Catalogue

1. General Description

A general description about the epSOS Master Value Set Catalogue and its relationship to the epSOS Translation/Transcoding Catalogue was given in the main document.

This document focuses on the **content** of the epSOS MVC and the characteristics of the code systems that were used.

1.1 epSOS Value Sets Catalogue Code Systems

These code systems are the ones that are used in the epSOS MVC and their characteristics.

1.1.1 EDQM

- **EDQM: Standard Terms of European Directorate of Quality in Medicine.**
  - Copyright & Issuer: EDQM
  - Languages & Localization: Albanian, Bulgarian, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Macedonian, Maltese, Norwegian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swedish and Turkish.
  - Fields of application: The List of Standard Terms covers dosage forms, routes of administration and containers used for medicines for human and veterinary use.
  - Number of entries: ~ 450
  - Structure: Part 1: Pharmaceutical dosage forms and short terms; Part 2: Routes of administration; Part 3: Containers.

1.1.2 LOINC

  - LOINC: Logical Observation Identifiers names and codes.
  - Copyright & Issuer: Regenstrief Institute, Indiana
  - Languages & Localization: English, Spanish, Chinese, German (Users guide), Estonian, Italian
  - Fields of application: The scope of the LOINC effort includes laboratory and other clinical observations. The laboratory portion of the LOINC database contains the usual categories of chemistry, hematology, serology, microbiology, toxicology; as well as categories for drugs and the cell counts and antibiotic susceptibilities. The clinical portion of the LOINC database includes entries for vital signs, hemodynamics, intake/output, electrocardiogram (EKG), obstetric ultrasound, cardiac echo, urologic imaging, gastroendoscopic procedures, pulmonary ventilator management, selected survey instruments (e.g. Glascow Coma Score, PHQ-9 depression scale, CMS-required patient assessment instruments), other clinical observations and document sections.
  - Number of entries: 53,344
  - Structure: Each LOINC record corresponds to a single test result or panel. The record includes fields for specifying:

- Component (analyte) - e.g., potassium, hemoglobin, hepatitis C antigen.
- Property measured - e.g., a mass concentration, enzyme activity (catalytic rate).
- Timing - i.e., whether the measurement is an observation at a moment of time, or an observation integrated over an extended duration of time - e.g., 24-hour urine.
- The type of sample - e.g., urine, blood.
- The type of scale - e.g., whether the measurement is quantitative (a true measurement) ordinal (a ranked set of options), nominal (e.g., E. coli; Staphylococcus aureus), or narrative (e.g., dictation results from x-rays).
- Where relevant, the method used to produce the result or other observation

- LOINC is also used for the CDA (Clinical Document Architecture) sections

1.1.3 UCUM

- UCUM: Unified Code for Units of Measure.[Datatype PQ]
- Copyright & Issuer: Regenstrief Institute, Indiana
- Languages & Localisation: English
- Fields of application: UCUM is a system of codes for unambiguously representing measurement units to both humans and machines.
- Number of entries: 556
- Structure: Each unit is defined relative to a system of base units by a numeric factor and a vector of exponents by which the base units contribute to the unit to be defined. Although we can reflect all the meaning of units covered by dimensional analysis with this vector notation, UCUM does not show vectors. Proposed definition from Regenstrief: The Unified Code for Units of Measure is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business. The purpose is to facilitate unambiguous electronic communication of quantities together with their units.

1.1.4 ATC

- ATC: Anatomical Therapeutic Chemical (ATC) classification system
- Copyright & Issuer: WHO Collaborating Centre for Drug Statistics Methodology, Norwegian Institute of Public Health
- Languages & Localization: English, Spanish, German, Italian
- Various manifestations: ATC-WHO, ATC-GM (Germany), ATC-WIDO (Germany), ATC Vet, ATC Herbal
- International non-proprietary names (INN) are used. If INN names are not assigned, USAN (United States Adopted Name) or BAN (British Approved Name) names are usually chosen.
- Fields of application: In the ATC-classification pharmacological substances are divided into different groups according to the organ or organ system which they affect and their chemical, pharmacological and therapeutic properties.
  A defined daily dose is assigned to each active substance. Defined daily doses (DDD) are the assumed average daily maintenance dose for the main indication of each sub-

stance in adults. The ATC-Classification with defined daily doses serves as an easing of comparisons between drugs and guarantees a standardized reference for the specification of daily treatment expenses. The purpose of the ATC/DDD system is to serve as a tool for drug utilization research in order to improve quality of drug use.

- Number of entries: 4067 Codes (2006)
- Structure: In the Anatomical Therapeutic Chemical (ATC) classification system, the drugs are divided into different groups according to the organ or system on which they act and their chemical, pharmacological and therapeutic properties. Drugs are classified in groups at five different levels. The drugs are divided into fourteen main groups (1st level), with one pharmacological/therapeutic subgroup (2nd level). The 3rd and 4th levels are chemical/pharmacological/therapeutic subgroups and the 5th level is the chemical substance. The 2nd, 3rd and 4th levels are often used to identify pharmacological subgroups when that is considered more appropriate than therapeutic or chemical subgroups.

  ATC assures the unified coding of active components registered. This feature, coupled with strength (dosage) and pharmaceutical form) assures the possibility to transfer the full information on a medication from Country to Country, regardless the brand name of the medication. A medicinal product can be given more than one ATC code if it is available in two or more strengths or formulations with clearly different therapeutic uses. The existence of multiple codes reduces the risk of mistakes in the specification of strength and prescription / dispensation of different medication for the two pathologies.

1.1.5 ICD

- ICD: International Statistical Classification of Diseases and Related Health Problems
- Copyright & Issuer: WHO
- Languages & Localization: ICD-10 is available in the six official languages of WHO (Arabic, Chinese, English, French, Russian and Spanish) as well as in 36 other languages.
- ICD-9 CM coding system and the application guidelines are available in Italian.
- Localizations have been made mainly for reimbursement purposes:
- Various manifestations: ICD-10-GM Version 2009 [1.2.276.0.76.5.356] icd10gm2009
- ICD-10-CM [2.16.840.1.113883.6.90]
- ICD-9-CM [2.16.840.1.113883.6.2]
- The Official Updates to the published volumes of ICD-10 are available as annual lists of changes and new versions with slight differences.
- Fields of application: The ICD is the international standard diagnostic classification of diseases (signs, symptoms, conditions) for all general epidemiological, health management and statistics (death) and clinical use (health records). These include monitoring of the incidence and prevalence of diseases and other health problems.
- Number of entries: ca. 13.000 classes
- Structure: 22 chapters, its nodes denote classes of diseases and related problems. ICD classes are arranged into up to five levels. There is one terminal class for each entity. Attributes of ICD are inclusions and exclusions and glossary-like text.

1.1.6 SNOMED CT

- SNOMED: Systematized Nomenclature of Medicine-Clinical Terms
- Copyright & Issuer: IHTSDO – International Health Terminology Standards Development Organization, a non-for-profit association based in Denmark. Products are open for researchers but for clinical coding it is restricted to its twelve country licensees and some companies/hospitals paying the license fee.
- Languages & Localization: English (US, UK), Spanish, Danish. Translations into French and Swedish are currently taking place. Translation into Lithuanian are also taking place but only on a small-scale.
- Fields of application: SNOMED CT is a comprehensive terminology, created to cover the whole patient record and medical documentation.
- Number of entries: 310.000 active concepts
- Structure: There are almost 800,000 descriptions in SNOMED CT, including synonyms that can be used to refer to a concept. In addition, there are approximately 1,360,000 links or semantic relationships between the SNOMED CT concepts. These relationships provide formal definitions and other characteristics of the concept. One type of link is the "IS_A" relationship.
- Cross-map to other international standards is not officially available.
- It has been announced the creation of a join working group between IHTSDO and WHO to develop an official cross-reference between SNOMED-CT and ICD-10 (ICD-11).

1.1.7 HL7 Code Systems

- In addition to the existing code systems, HL7 also provides different code systems such as the HL7 AdministrativeGender for example.
-

1.1.8 IHE (Integrating the Healthcare Enterprise) Vocabularies

- Since the syntax used in the pivot document is based on the IHE PCC specifications, there are some specific code systems defined by IHE such as IHEActCodeVocabulary or IHERoleCodeVocabulary. These are necessary for the implementers and are part of the coding specifications.

The epSOS MVC can be found in the file folder Supporting materials on Project Place.

1.2 epSOS Value Sets Catalogue Mapping

Based on the use cases, the appropriate value sets are chosen based from the existing code systems mentioned above. The participating member states have various degrees of formalities within their code systems which are suitable for different purposes. Parts of the epSOS Master Value Sets Catalogue can be mapped to many of the national code systems; however with the increasing number of code systems, the maintenance of this mapping becomes a challenging task.

The following general steps are employed in the mapping process, and should be kept in mind by the person(s) doing the mapping

Define the context of the concept to be mapped. The concepts being mapped do not equate exactly all the time. Whenever this happens, there is a need to make decisions about the best correspondence, and these decisions are often dependent upon the purpose and the contextual use case. Mapping may require not just the identification of the closest match, but also the rules that are to be applied when choosing terms or code systems. Accurate and consistent mapping of concepts may need to reference the context in which the term or code is used. The ability to identify context within an information model along with the purpose for which mapping the data may significantly change the concepts determined to map to each other.

- **Address difficulties associated with the level of granularity**
  If a term form a code system must be decomposed into its terminological components (reverse aggregation) so that mapping can be done, the resultant decomposed elements must be analyzed to identify how they best match the source and the target term.

- **Identification of the mapping**

  The data handled must be identified since different methods can be chosen depending if dealing with clinical or administrative data.

- **Standardized mapping**
  The requirements for standardized mapping methods are defined, including the ability to represent these relationships in a machine processable format.

1.3 Mapping Issues

Mapping features are one of essential requirements for the epSOS semantic services. The main challenge of the adoption of the epSOS Master Value Sets Catalogue is the use of legacy, non-internationally based terms tailored and optimized to meet local (national) requirements. One such example is the catalogue of procedures, provided by health care providers from various countries, which is in many countries not compatible with standardized clinical terminology. For this purpose is very important to be methodical about this activity.

Various codes systems cover the spectrum of the epSOS Master Value Sets Catalogue as defined by the use cases such as ICD 10, ATC, LOINC and SNOMED CT for example. Most of these codes systems are used in the participating countries.

As there is usually M:N relationship between concepts of the epSOS Value Sets Catalogue and the official code systems , the direction of mapping makes a difference. . The sections below describe this process.

1.3.1 Handling of 1:N and M:N relations

One-to-one mapping-strategy was chosen by WP 3.5 (one term from the epSOS Value Sets Catalogue maps only to terms from one official code system). Whenever a term can correspond to many terms from different code systems, the mapping becomes much more complicated, there is the need to chose the most expressive one among a group of terms.

While doing the mapping or more accurately, the cross-referencing between the epSOS value sets, the following points have to be considered:

- detailed review of the meaning of concepts on both source and target,
- identification of mapping problems (concepts not having unique equivalent), and the way how to handle them,
- deep knowledge of the terms and the code systems being used for those mapping
- future usage of mapping (transformation from the epSOS Value Sets Catalogue to the code system, and from the code system to epSOS Value Sets Catalogue or both)
- Evaluation and broad acceptability of all users of various code systems.

It must be noted that the mapping one-to-one or one-to-many is under the responsibility of the Member States, since there is no official mapping between codes systems. This becomes even more of an issue when national extensions of code systems are involved. There can be no risk assessment done as epSOS does not take the responsibility for mapping between the code systems, just offering the services for accessing and managing the content. The content itself is produced by the Member States.