Integrating MyHarmony
(Harmonization & Codification of Unstructured Data using SNOMED CT) and Malaysian Health Data Warehouse

Transforming unstructured to structured data using SNOMED CT

Authors
Md. Khadzir Sheikh Ahmad¹, 'Ismat Mohd Sulaiman¹, Mohd Syazrin Mohd Sakri¹, Muhammad Aiman Mazlan², Abdul Aziz Latip², Omar Ismail¹

¹Ministry of Health Malaysia, ²MIMOS
OVERVIEW of MyHDW

OPERATIONAL SYSTEMS - for primary use

Government Clinics
Government Hospitals
Private Hospitals
Pollution & Weather
Meteorology & Environment

MyHDW: MALAYSIAN HEALTH DATA WAREHOUSE

Data collection systems

Sistem Maklumat Rawatan Pelanggan (SMRP)
Patient Registry Information System (PRIS)

Analytical systems – for secondary use

To support evidence based decision making for the effective Management of the Health System

Reporting, Dashboard
Ad-hoc Analytics
Geographic Information System (GIS)
Research & Statistical Analysis
SDG, UHC, Performance Indicators
Text analytics for unstructured data

MyHDW version 1.0
Generate reports required by the stakeholders

More time-efficient
- Information timeliness
- Continuous monitoring

New questions can be answered
- Reuse available data – saves cost and avoid new system development

No Duplication of Data Entry - Data from the source
- Reduce workload and transcribing error
Structuring the unstructured (SNOMED CT codified) data

1. Create codified enumeration dimension in MyHarmony
2. Create datamart (OLAP) to be consumed by MyHDW
3. Create Dashboard and ad-hoc queries
1. Codified Enumeration Dimension

• “Codified Enumeration” dimension is a **virtual dimension** to abstract the complexity of the terminology from the BI.

• MyHarmony expands the codified enumeration into a set of Codified Enumeration Items, which are SNOMED CT concepts defined by the Refset owner.

• As a result, BI could be used to analyse codified data without the need to address the complexity of SNOMED CT and the refset used to codify it.

**Codified Enumeration Dimension**

E.g. Filtering data based on SNOMED CT-codified diagnosis

**Codified Enumeration Items**

- The codification engine expands each item when query is made.
- It allows generalizability and specificity.
- E.g. “IHD” is expanded to include STEMI, UA, ACS as defined in SNOMED CT relationship structure and the Refset used.
2. Create datamart in Mi-BIS

Fact table and Dimension table in a Star Schema model are created.

Codified data in MyHarmony are mined to generate Dimension tables.
 Codified Enumeration Dimension Transformed into Dimension Table

Data marts provides all possible combination of dimensions i.e. enumeration values for each diagnosis. Because of that, to do aggregate count (on OLAP CUBE), a “DISTINCT” keyword must be introduced. SQL: COUNT (DISTINCT(primary_key)) instead of COUNT(primary_key).

A patient can have multiple diagnosis and enumeration (codified data)

Unique ID - MRN is unique per patient, but not unique per visit
3. Unified Analytics Using Mi-BIS

Mi-BIS: Business Intelligence platform in MyHDW developed by MIMOS

- Data Mart based on harmonised and codified data, and optionally joined with other datasets

- Drag-and-drop dimensions and measures:
  - To be displayed as rows and columns
  - To be applied as fixed filters
  - To be used as dynamic filter selected by report user

- Dimensions, consisting of:
  - Normal datamart dimensions
  - “Codified Enumeration” dimensions

- Measures, where the following could be applied:
  SUM, AVG, COUNT, MIN, MAX, COUNT-DISTINCT
Dashboard

CARDIOLOGY DEPARTMENT
ST Elevation Myocardial Infarction (STEMI) Without Shock Case Fatality Rate

Month: JUL
Undergo Cardiac Catheterization: 44
Total Patients: 75
Percentage (%): 58.67
Achievement: Below Target < 90.00

Report date: 21-12-2018 03:05:49 PM
IMPLEMENTATION & VALIDATION

Generating National Cardiology Key Performance Indicator (KPI)

\{ 1 KPI = 1 Datamart \}
National Cardiology Key Performance Indicator (KPI) No. 4: Non-ST Elevation Myocardial Infarction (NSTEMI) Case Fatality Rate

MyHarmony: Numerator

4
5
10

MyHarmony: Denominator

15

Mi-BIS

3.80% 96.2% 100%

NSTEMI Case Fatality Rate (Hospital X), 2017:

MyHarmony / Mi-BIS : 3.8%
BENEFIT OF USING MyHARMONY in generating KPI in MyHDW

CONVENTIONAL

- Multiple data entries at the health facilities
  - Prone to transcribing error & increase workload

- Manual data entry and manual submission (even for IT hospitals)
  - Aggregated data submitted monthly to CKPP
  - Figures can be manipulated
  - Result published yearly or upon request
  - Prone to delayed reaction and action

MYHARMONY

- Single entry at electronic discharge summary in free text
  - Requires proper documentation and awareness

- Semi-automated submission, automated codification, and automated KPI generation
  - Able to filter & drill down for every KPI due to granular data
  - More timely results and auto-generated results
### Future works

#### SNOMED CT
- Managing SNOMED CT versioning for consistent results
- More efficient manner to develop SNOMED CT Refsets
- Use of SNOMED CT attribute relationship

#### Improve MyHarmony
- Additional context and word sense disambiguation
- Codification processing time for partial matched terms

#### Expand MyHarmony
- To other clinical specialties
- In Electronic Medical Record (front-end) – automated codification during documentation
Thank you

‘ISMAT MOHD SULAIMAN
Health Informatics Centre, Planning Division, Ministry Of Health Malaysia
drismat@moh.gov.my