TSB – MHS Ontology
Methods of Creation, Maintenance and Extraction

Mariana Casella dos Santos, MD
James Flanagan, MD, PhD
Language and Computing

AHIMA 2005
Presentation Overview

- What is the TSB and Why?
- TSB mappings desiderata
- Methods of mappings creation and maintenance
- Methods of mapping extraction
Mariana Casella dos Santos

What is the TSB and Why?

- Terminology Service Bureau (TSB) is designed to meet the terminology requirements of the Composite Health Care System II (CHCS II), the computer-based Patient Record (CPR) System for the Department of Defense (DoD) Military Healthcare System (MHS).

- CHCS II collects medical data from all healthcare services within the DoD.
  - Must handle conflicting data and terminology standards.
  - Must optimize and improve the integration and transmission of medical information throughout the MHS.

- CHCS II must cope with integrations of additional software components and their associated terminologies.
  - 3M’s HDD was responsible for standardizing the terms used to capture clinical information
  - new terminology identified - mapped into the HDD - available to be used by clinical applications
    - Mappings must be performed by 3M
    - Inability to generate updates for the original terminologies mapped or components using these terminologies

**TSB Purpose:** Incorporate and relate multiple terminologies into one overarching data model that is capable of generating updates based on this integrated ontology/terminology for each of the runtime components that require terminology content.
Desiderata for TSB Mappings and Mapping-related Processes

- Ability to deliver on time and on budget
- Secure reliability of mapped information
- Mapping efforts are cumulative and publicly available
- Mappings must serve a multiplicity of disparate applications
- Mitigate the differences introduced by existing medical information sources
  - hybridism of principles or perspectives embedded in their hierarchical structure
  - divergence of meaning attributed to their language descriptions
TSB is Portable and Partitionable
Terminology Service Bureau

The Big Picture

MHS Ontology

SNOMED CT (Core)

PKC

MEDCIN

Other

RxNorm

VA Allergy Term set

DoD Allergy Term Set

TSB Content

CHCS II

TMIP

CHDR

PKC/Medcin/SNOMED CT

Others

Language and Computing
TSB Mapping Framework
The Meta and the Domain

MHS Core Ontology

SNOMED CT CONCEPT

MEDCIN CONCEPT

SNOMED CT Terms

MEDCIN Terms
TSB Mapping Framework
The Meta and the Domain

DOMIAN ENTITY

META ENTITY

ENTITY

SNOMED CT
NORMALIZED

CONCEPTS ORIGINATING FROM OTHER TERMINOLOGIES

SNOMED-CT
 оригinal structure

MEDCIN
PKC

Bi-directional mapping relation
Ontology for Alignement  
The Meta and the Domain

Dealing with Granularity

Versioning

Mariana Casella dos Santos
Ontology for Alignment
The Meta and the Domain

Reusability of Mappings

Snomed-CT: 7180009: MENINGITIS (DISORDER)

Medcin: 31192: MENINGITIS

Pkc Entity: 644: MENINGITIS

Meddra: 10027199: MENINGITIS

Language and Computing
Mapping Terminologies
Lexical Layer Issues
Lexical Variations Match with Upper Level Constraints

- Defining allowable lexical variations of a term, based on analysis of the hierarchical structures of the terminologies.

- Generation of an equivalence model for distinctive lexical representations identified as synonymous within the scope of both terminologies being mapped, utilizing regular expressions.

- Setting specific constraints to restrict realm of mappings based on terminologies hierarchical branches.

- Applying secure lexical match-based mapping algorithm utilizing the lexical variants equivalence model created.
Pattern of equivalence found:

Expression on Equivalence Model: (.*) = examination of (.*)

Setting constraint for defined expression:
Concept A with term “(.*)” MAPS TO Concept B with term “examination of (.*)” if and only if: Concept A IS-A MEDCIN : 6000 : PHYSICAL EXAMINATION and Concept B IS-A PROCEDURE (*MHS Core originating from and linked to SNCT : 71388002 : PROCEDURE (PROCEDURE))
Secure Mapping Algorithm

Invalid exact string matches

A. Meta concept parent
   Meta concept A
   term x
   term y
   match
   Domain concept parent
   Domain concept B
   term j
   term i

Constraint Violation
*Occurs for untruth variations or real homonyms.

‘Homonymous’ Term Variations
*Not disambiguated by constraint.
Occurs for cases of redundancy, concepts w/ NOS/‘other’, broad constraints.

Valid exact string matches

A. Meta concept A
   term x
   term y
   match
   Domain concept B
   term j
   term i

1 Synonymous Match

B. Meta concept A
   term x
   term y
   match
   Domain concept B
   term j
   term i

>1 Synonymous Match

C. Meta concept A
   term x
   term y
   match
   Domain concept B
   term j
   term i
   Domain concept C
   term k

>1 MHS Domain concept
*Occurs for ‘untruth’ synonyms.

Mariana Casella dos Santos
Medical Terminologies/Ontologies
Different Perspectives

 EntityState

- MEDCIN: 102935: TESTS
- MEDCIN: 23602: SYMPTOMS
- MEDCIN: 39449: DIAGNOSES, SYNDROMES AND CONDITIONS
- MEDCIN: 40011: THERAPY
- MEDCIN: 5141: HISTORY

- MEDCIN: 8080: PHYSICAL EXAMINATION
  - MEDCIN: 11454: OBJECTIVE (USE FOR S.O.A.P NOTE FREE TEXT)
  - MEDCIN: 11525: EXAMINATION OF THE RECTUM
  - MEDCIN: 202199: EXAMINATION OF THE BACK
  - MEDCIN: 208847: PHYSICAL FINDINGS (USE FOR FREE TEXT)
  - MEDCIN: 240096: STANDARD MEASUREMENTS
  - MEDCIN: 8001: VITAL SIGNS
  - MEDCIN: 5142: EXAMINATION OF THE SKIN
    - MEDCIN: 10013: SKIN GENERAL APPEARANCE
    - MEDCIN: 10014: SKIN APPEARANCE WRINKLED
    - MEDCIN: 10015: SKIN APPEARANCE FURROWED
    - MEDCIN: 11527: SKIN APPEARANCE WEATHERED
    - MEDCIN: 9380: SKIN FISSURES
    - MEDCIN: 208934: ASSESSMENT OF INDWELLING CATHETER
    - MEDCIN: 208921: ASSESSMENT OF INJECTION SITE
    - MEDCIN: 208944: ASSESSMENT OF FINGERSTICK SITE
    - MEDCIN: 248691: ASSESSMENT OF TRACHEOSTOMY SITE
    - MEDCIN: 6079: SKIN COLOR & PIGMENTATION
    - MEDCIN: 6115: SKIN MOISTURE
    - MEDCIN: 6126: SKIN TEMPERATURE
    - MEDCIN: 6141: METAL SPLINTERS IN SKIN
    - MEDCIN: 6143: SKIN LESIONS (ON EXAM)
    - MEDCIN: 6305: SKIN MOBILITY
    - MEDCIN: 6309: SKIN TEXTURE

- SNOMED-CT: 853456008: FOREIGN BODY IN SKIN (DISORDER)
- SNOMED-CT: 953196004: SKIN NODULE (DISORDER)
- SNOMED-CT: 95322002: EDEMATOUS SKIN (DISORDER)
- SNOMED-CT: 95336000: DISTICHIASIS (DISORDER)
- SNOMED-CT: 95347000: SKIN NECROSIS (DISORDER)
- SNOMED-CT: 95094004: ECTOPIA CILIUM OF EYELID (DISORDER)
- SNOMED-CT: 95546008: PERINEAL SKIN TAGS (DISORDER)
- SNOMED-CT: 95638002: SKIN LESION (DISORDER)
- SNOMED-CT: 95321009: FISSURE IN SKIN (DISORDER)
- SNOMED-CT: 367523002: PELVIC (DISORDER)
- SNOMED-CT: 10276007: FACTITIOUS PURPURA (DISORDER)
- SNOMED-CT: 105966009: TUMOR-LIKE LESION OF SKIN (DISORDER)
- SNOMED-CT: 103600008: CONGENITAL CONMISURAL P (DISORDER)
- SNOMED-CT: 103957002: DRUG-INDUCED PURPURA (DISORDER)
- SNOMED-CT: 111102003: NEOVASCULARIZATION OF SKIN (DISORDER)
- SNOMED-CT: 111102007: HYPERTROPHY OF DERMAL P (DISORDER)
- SNOMED-CT: 111102003: MACULAR CUTANEOUS AMYLC (DISORDER)
- SNOMED-CT: 123657000: SEBACEOUS GLAND DUCT DILATATION (DISORDER)
- SNOMED-CT: 123659002: SWEAT GLAND DUCT DILATATION (DISORDER)
- SNOMED-CT: 123659002: SWEAT GLAND DUCT DILATATION (DISORDER)
- SNOMED-CT: 123706005: HAIR DYSTROPHY (DISORDER)
- SNOMED-CT: 125480004: NEOPLASM OF SKIN (DISORDER)
- SNOMED-CT: 1271009: KNUCKLE PADS, DEAFNESS AND (DISORDER)
- SNOMED-CT: 1515008: GORHAM'S DISEASE (DISORDER)
- SNOMED-CT: 1521007: BLISTER OF BUTTOCK (DISORDER)
- SNOMED-CT: 1525003: BLISTER OF FOOT WITHOUT INFECT (DISORDER)
- SNOMED-CT: 182762007: DERMATITIS (DISORDER)
- SNOMED-CT: 18505002: FERNAL CUTANEOUS HEMORRHAGE (DISORDER)
- SNOMED-CT: 18748008: CONGENITAL KOILONYCHIA (DISORDER)
• Axiom: If concept A is-child-of (IS-A) concept B then every instance of A is a instance of B.

Every instance of viral meningitis is an instance of meningitis.
IS-A Link
Common Uses in Medical Terminologies

Causality

Parthood

ABSCESS

PASTEURELLA INFECTION

PASTEURELLA ABSCESS

LIVER

LIVER PART

L & C

Language and Computing

Mariana Casella dos Santos 17
Motivation: When new terminology is mapped many concepts are introduced in the MHS core ontology ‘hanging loose’ (w/o an identified parent concept) and structuring is lacking. Identifying and transferring these smaller hierarchies onto the domain can drastically reduce the manual modeling effort.

• Analysis of the terminology for the identification of certain ‘cutting points’ – concepts that are at the top of a small hierarchy with ontologically valid IS-A relations.

• Analysis of the terminology for the identification of hierarchical levels where IS-A relations are always valid.

• Transposition Procedure utilizing ‘cutting points’ and hierarchical levels identified.
Transposing Hierarchies from Mapped Terminologies

‘Cutting Point’ Methods:

Transpose to existing map n levels

Transpose to existing map 1 level

Transpose through levels of granularity

Mariana Casella dos Santos
Transposing Mapped Terminologies Hierarchies

Mariana Casella dos Santos
Transposing Mapped Terminologies Hierarchies

**Transposition in groups – Valid Hierarchical levels:**

* Transposition always to n levels down
### Transposing Mapped Terminologies Hierarchies

<table>
<thead>
<tr>
<th>EDCIN CATEGORY</th>
<th>EDCIN ENTITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDIN CATEGORY</td>
<td>MEDIN ENTITY</td>
</tr>
<tr>
<td>MEDIN: 6000: PHYSICAL EXAMINATION</td>
<td>MEDIN: 7843: EXAMINATION OF THE MUSCULOSKELETAL SYSTEM</td>
</tr>
<tr>
<td>MEDIN: 8072: MUSCULOSKELETAL EXAM - ANKLE</td>
<td>MEDIN: 8073: ANKLE APPEARANCE</td>
</tr>
<tr>
<td>MEDIN: 187609: ANKLE ERYTHEMA</td>
<td>MEDIN: 187608: ANKLE ERYTHEMA RIGHT</td>
</tr>
<tr>
<td>MEDIN: 187621: ANKLE ERYTHEMA RIGHT ANTERIOR</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170731: ANKLE ERYTHEMA RIGHT ANTEROLATERAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170732: ANKLE ERYTHEMA RIGHT MEDIAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170733: ANKLE ERYTHEMA RIGHT POSTERIOR</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170734: ANKLE ERYTHEMA RIGHT POSTEROLATERAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170735: ANKLE ERYTHEMA RIGHT POSTEROLATERAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170736: ANKLE ERYTHEMA RIGHT LATERAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170737: ANKLE ERYTHEMA RIGHT ANTEROLATERAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 170700: ANKLE ERYTHEMA LEFT</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 187701: ANKLE ERYTHEMA BILATERAL</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 16707: ANKLE DEFORMITY</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 175582: ANKLE INFLAMATION</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 175559: ANKLE DEVICATED ACHILLES TENDON</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 17768: ANKLE TOE-OUT ANGLE (DEG)</td>
<td></td>
</tr>
<tr>
<td>MEDIN: 17768: ANKLE ALIGNMENT</td>
<td></td>
</tr>
</tbody>
</table>

#### Mapping relationship

- **SNCT: 247441603: ERYTHEMA (FINDING)**
  - **CRITERIA**
    - **IS: CCC-OF ERYTHEMA**
    - **IS: CCC-OF ERYTHEMA**
  - **FULL DEFINITIONS**
Exporting the Mappings

- Configurable Exports:
  e.g. **MEDCIN ID** <> SNOMED CT PT <> SNOMED CT FSN <> SNOMED CT ID
  or
  **SNOMED CT PT** <> PKC ID <> MEDCIN ID

- Extracting ‘One to One’, ‘One to Many’ and/or ‘Many to One’ mappings by traveling the domain hierarchy.
*As MHS domain modeling evolves post-coordinated mappings extraction becomes also possible.
Mapping Extraction
‘One to Many’

PKC: DIABETES MELLITUS TYPE II

ICD-9-CM: 250:
DIABETES MELITUS

ICD-9-CM: 250.02:
DIABETES MELLITUS WITHOUT MENTION OF COMPLICATION, TYPE II

ICD-9-CM: 250.12:
DIABETES WITH KETOACIDOSIS, TYPE II

DIABETES MELLITUS

DIABETES MELLITUS TYPE II

DIABETES MELLITUS TYPE II WITH KETOACIDOSIS

DIABETES MELLITUS TYPE II W/O MENTION OF COMPLICATION
Conclusions

Mapping Processes Key Words
Reusability and Flexibility
Clinical Vocabulary Mapping Methods Institute
Saturday, October 15, 2005