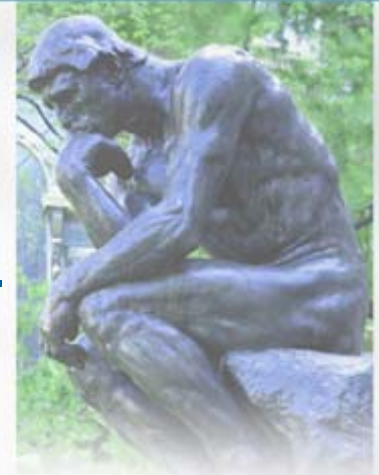


# TSB – MHS Ontology Methods of Creation, Maintenance and Extraction



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

**AHIMA 2005**

Language Computing



Language and Computing

# Presentation Overview

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



# 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.

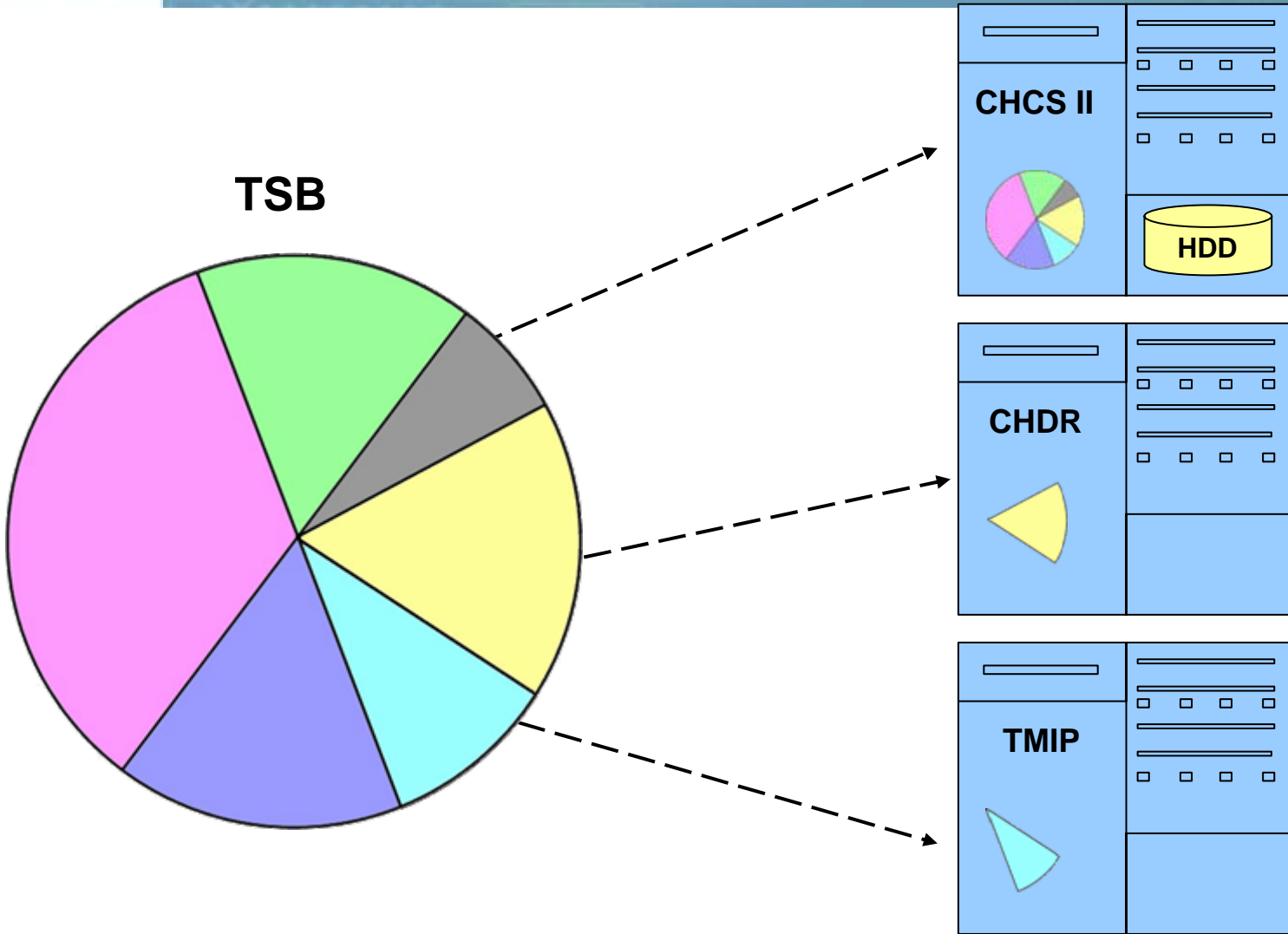


# Desiderata for TSB Mappings and Mapping-related Processes

- 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

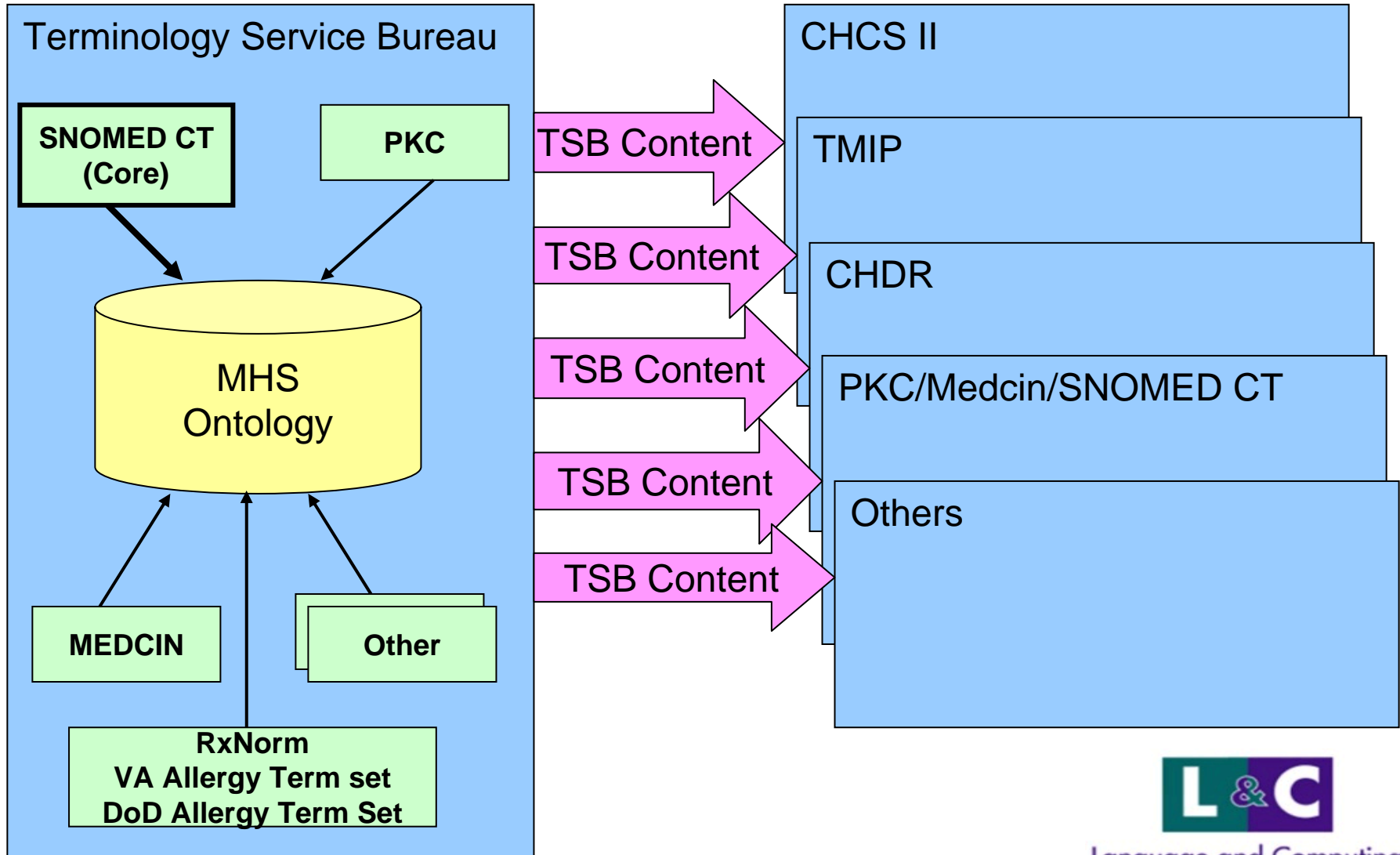


# TSB is Portable and Partitionable



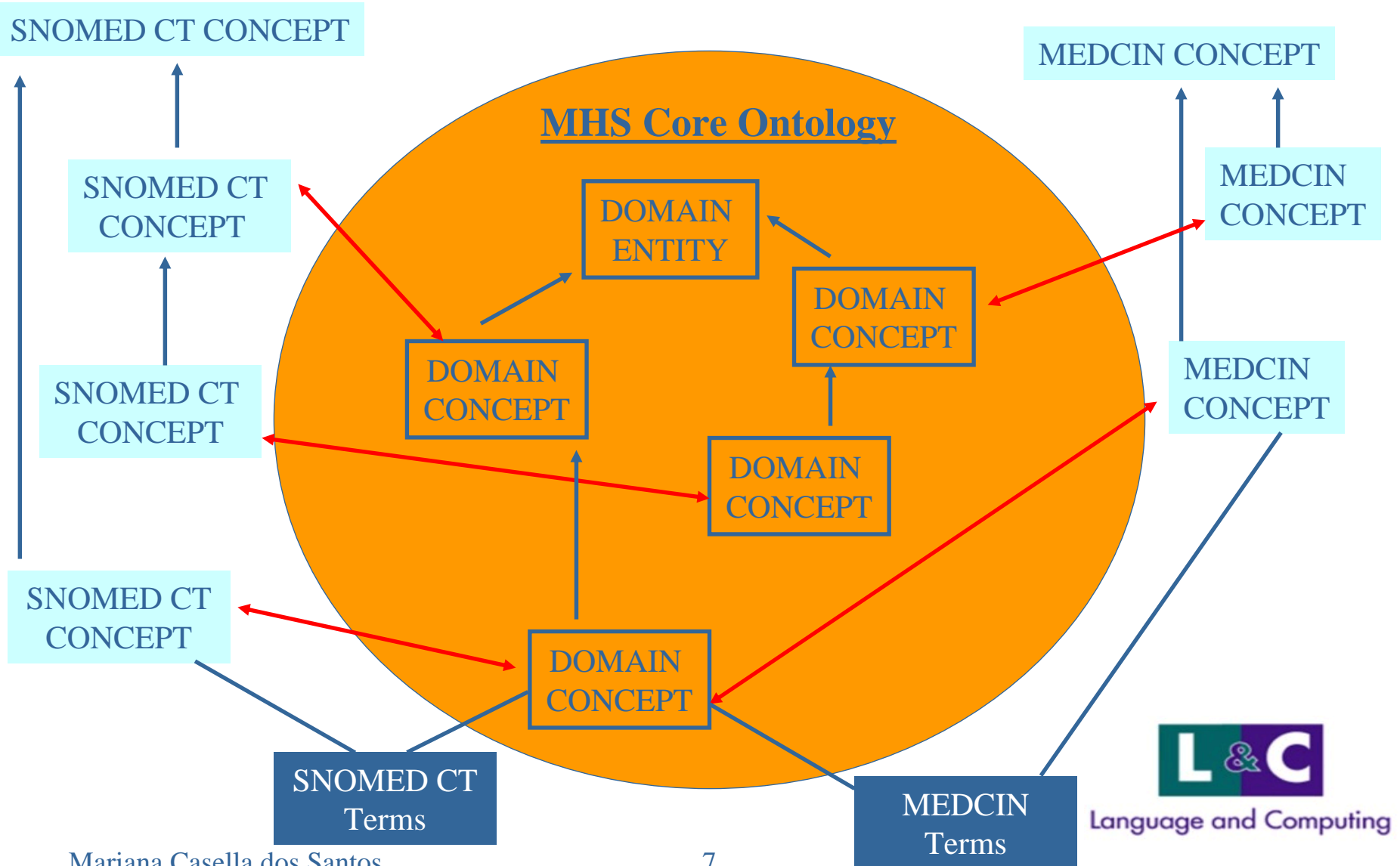
# Terminology Service Bureau

## The Big Picture



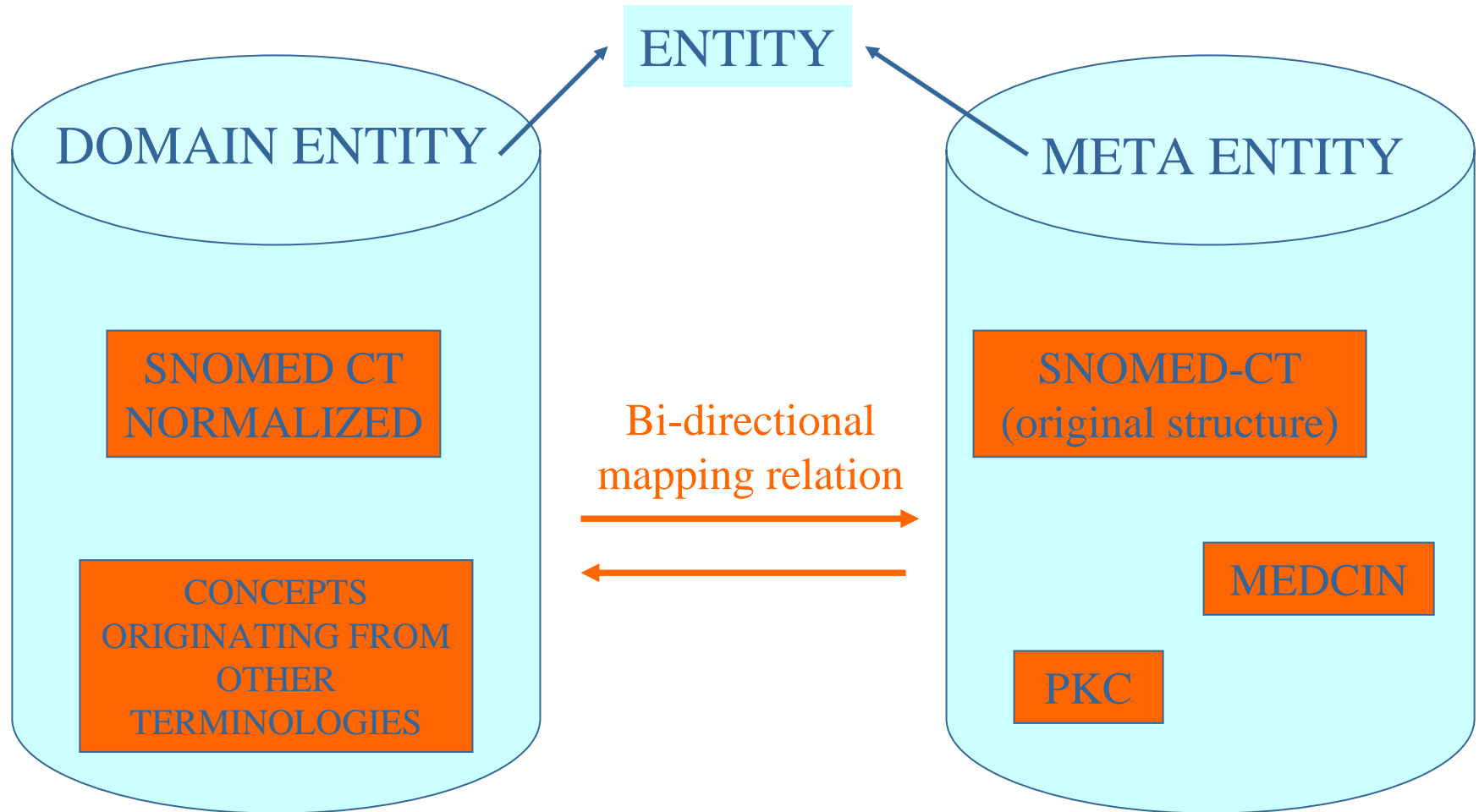
# TSB Mapping Framework

## The Meta and the Domain



# TSB Mapping Framework

## The Meta and the Domain

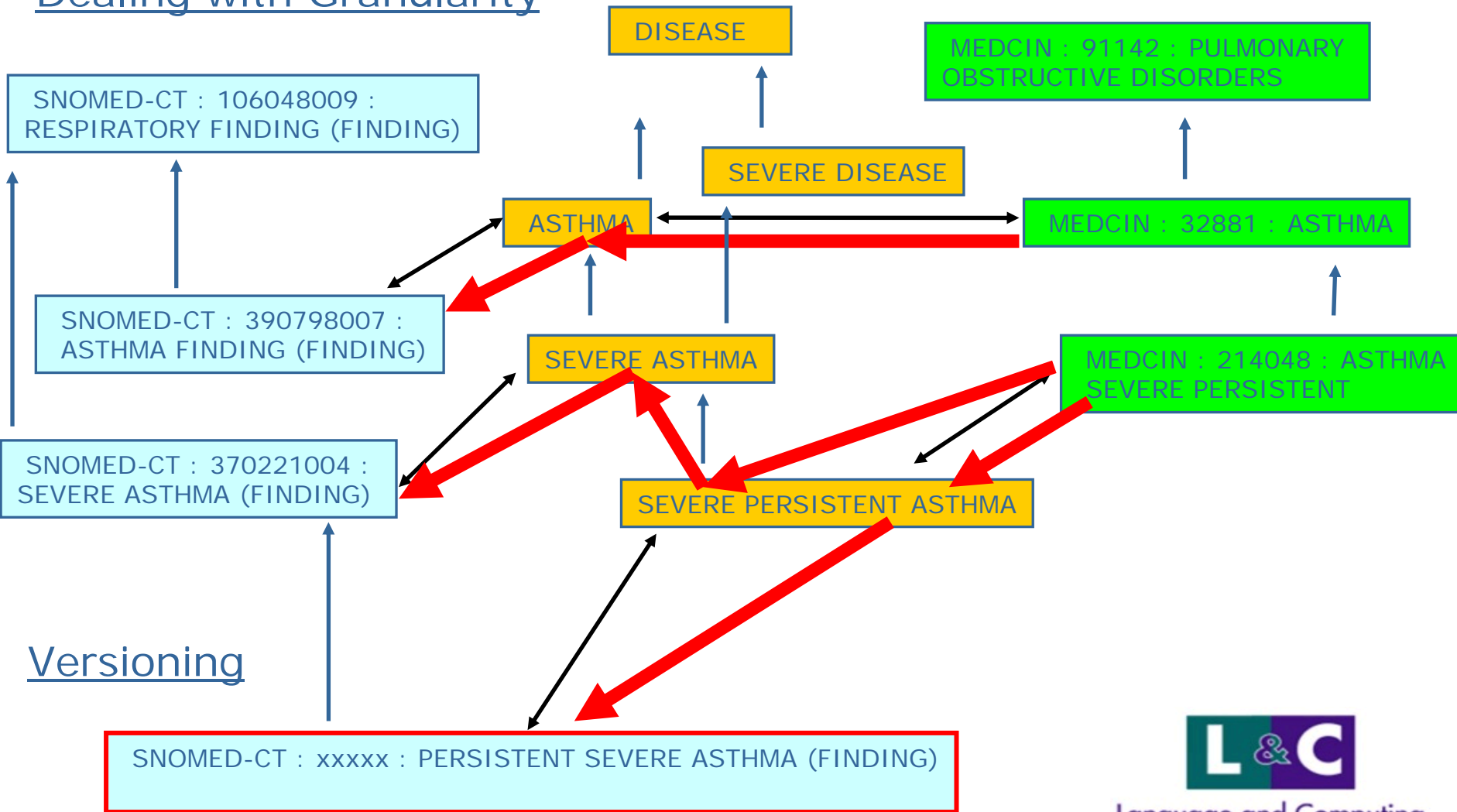




# Ontology for Alignment

## The Meta and the Domain

### Dealing with Granularity

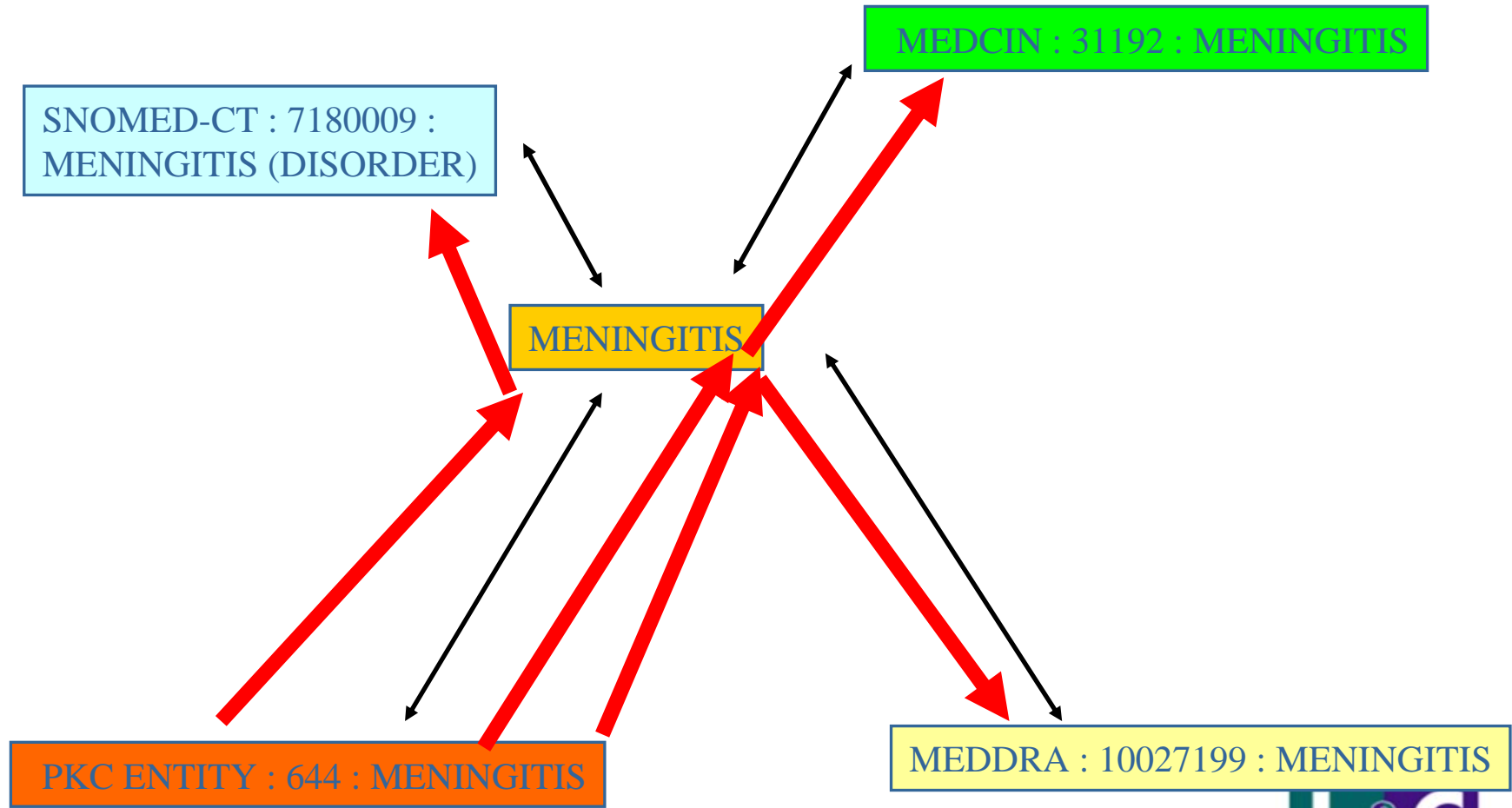


### Versioning

# Ontology for Alignment

## The Meta and the Domain

### Reusability of Mappings



# Mapping Terminologies Lexical Layer Issues

LinkFactory Workbench 2.4 (Mariana Casella Dos Santos on rmi://192.168.0.244:7000/LFserver)

File Layout Beans Bookmarks Tools Windows Help

The screenshot displays four ConceptTree panes in a grid layout, illustrating mapping issues between different terminologies. The panes are:

- Top-Left:** STANDARD CODING & CLASSIFICATION CATEGORY
  - MEDICIN ENTITY
    - MEDICIN : 6000 : PHYSICAL EXAMINATION
      - MEDICIN : 6425 : EXAMINATION OF THE EYES
        - MEDICIN : 11815 : RETINA** (highlighted in blue)
        - MEDICIN : 6665 : CHORIORETINITIS

- Top-Right:** IOMED-CT
- SNOMED-CT : 119263007 : EYE PART (BODY STRUCTURE)
  - SNOMED-CT : 302151004 : EYEBALL SEGMENT STRUCTURE (BODY STRUCTURE)
    - SNOMED-CT : 77205005 : POSTERIOR EYEBALL SEGMENT STRUCTURE (BODY STRUCTURE)
      - SNOMED-CT : 303295006 : CHOROIDAL AND RETINAL STRUCTURES (BODY STRUCTURE)
        - SNOMED-CT : 5665001 : RETINAL STRUCTURE (BODY STRUCTURE)
          - SNOMED-CT : 181171005 : ENTIRE RETINA (BODY STRUCTURE)** (highlighted in blue)
          - SNOMED-CT : 113323002 : STRUCTURE OF RETINA OF BOTH EYES
          - SNOMED-CT : 17139002 : ROD OF RETINA (CELL)
          - SNOMED-CT : 280611009 : ZONE OF RETINA (BODY STRUCTURE)
          - SNOMED-CT : 280657006 : LAYER OF RETINA (BODY STRUCTURE)
          - SNOMED-CT : 280691005 : SUBDIVISION OF RETINA (BODY STRUCTURE)
          - SNOMED-CT : 28758002 : STRUCTURE OF BLOOD VESSEL OF RETINA
          - SNOMED-CT : 34407003 : ROD CELLS OF INNER NUCLEAR LAYER
          - SNOMED-CT : 368847001 : RETINAL FOLD (MORPHOLOGIC ABNORMALITY)
          - SNOMED-CT : 5597008 : STRUCTURE OF RETINA OF RIGHT EYE (BODY STRUCTURE)
- Bottom-Left:** SNOMED-CT
- SNOMED-CT : 371409005 : DISORDER OF EYE REGION (DISORDER)
  - SNOMED-CT : 371405004 : DISORDER OF EYE PROPER (DISORDER)
    - SNOMED-CT : 128536001 : DISORDER OF POSTERIOR SEGMENT OF EYE (DISORDER)
      - SNOMED-CT : 128468007 : DISORDER OF CHOROID OF EYE (DISORDER)
        - SNOMED-CT : 16553002 : CHOROIDITIS (DISORDER)
          - SNOMED-CT : 46627006 : CHORIORETINITIS (DISORDER)** (highlighted in blue)
          - SNOMED-CT : 1979001 : FOCAL CHOROIDITIS (DISORDER)
          - SNOMED-CT : 231993001 : PNEUMOCYSTIS CARINII CHOROIDITIS (DISORDER)
          - SNOMED-CT : 27339005 : DISSEMINATED CHOROIDITIS (DISORDER)
          - SNOMED-CT : 95683004 : GRANULOMATOUS CHOROIDITIS (DISORDER)
- Bottom-Right:** IOMED-CT
- SNOMED-CT : 118690002 : PROCEDURE ON HEAD (PROCEDURE)
  - SNOMED-CT : 371548008 : PROCEDURE ON EYE REGION (PROCEDURE)
    - SNOMED-CT : 118897002 : PROCEDURE ON EYE (PROCEDURE)
      - SNOMED-CT : 118910000 : PROCEDURE ON POSTERIOR SEGMENT OF EYE (PROCEDURE)
        - SNOMED-CT : 118906003 : PROCEDURE ON RETINA (PROCEDURE)** (highlighted in blue)
        - SNOMED-CT : 274798009 : EXAMINATION OF RETINA (PROCEDURE)** (highlighted in blue and circled in red)
        - SNOMED-CT : 112965005 : DIAGNOSTIC PROCEDURE ON RETINA (PROCEDURE)
        - SNOMED-CT : 20067007 : OCULAR FUNDUS PHOTOGRAPHY (PROCEDURE)
        - SNOMED-CT : 20452001 : REATTACHMENT OF CHOROID AND RETINA (PROCEDURE)
        - SNOMED-CT : 21965005 : REMOVAL OF RETINAL IMPLANT (PROCEDURE)
        - SNOMED-CT : 231760002 : CRYOTHERAPY TO RETINA (PROCEDURE)
        - SNOMED-CT : 231782009 : REMOVAL OF GAS/AIR FROM VITREOUS (PROCEDURE)
        - SNOMED-CT : 303833005 : RETINAL ANGIOGRAPHY (PROCEDURE)
        - SNOMED-CT : 45817009 : REMOVAL OF FOREIGN BODY FROM RETINA (PROCEDURE)
        - SNOMED-CT : 53696006 : OPERATION ON RETINA (PROCEDURE)

Annotations:

- A blue arrow points from **MEDICIN : 11815 : RETINA** to **SNOMED-CT : 181171005 : ENTIRE RETINA (BODY STRUCTURE)**.
- A red arrow points from **MEDICIN : 11815 : RETINA** to **SNOMED-CT : 274798009 : EXAMINATION OF RETINA (PROCEDURE)**.
- A red 'X' is drawn over the blue arrow.
- A red circle highlights **SNOMED-CT : 274798009 : EXAMINATION OF RETINA (PROCEDURE)**.

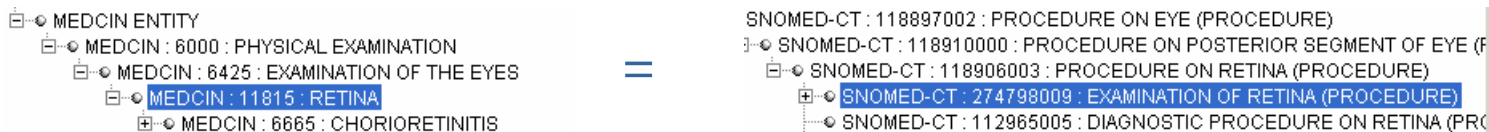
# Lexical Variations Match with Upper Level Constraints

- 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.



# Lexical Variations Match with Upper Level Constraints

## Pattern of equivalence found:



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

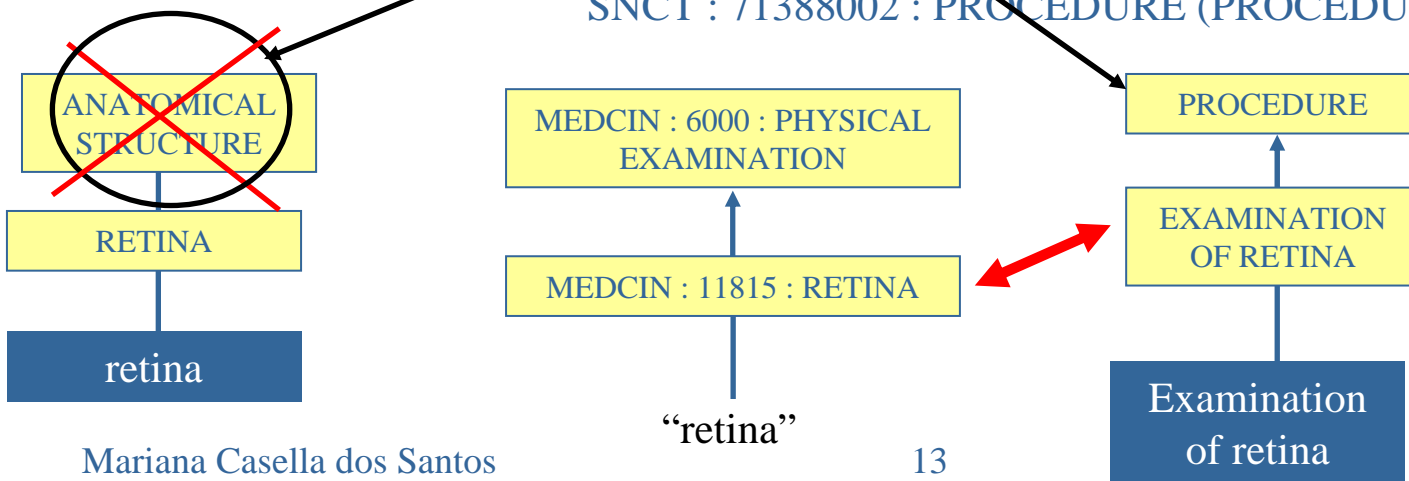
### Setting constraint for defined expression:

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

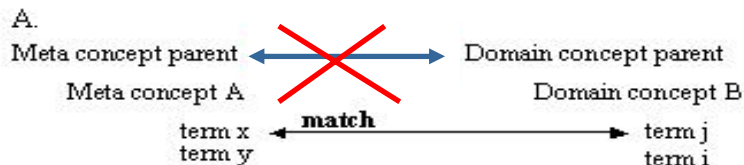
Concept B IS-A PROCEDURE

(\*MHS Core originating from and linked to SNCT : 71388002 : PROCEDURE (PROCEDURE))



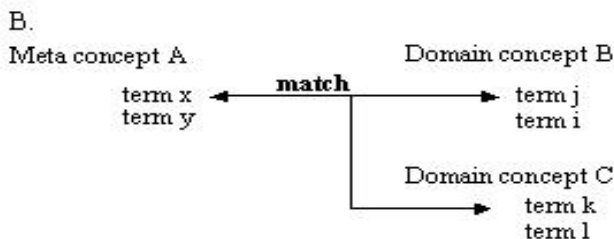
# Secure Mapping Algorithm

## Invalid exact string matches



Constraint Violation

\*Occurs for untruth variations or real homonyms.



‘Homonymous’ Term Variations

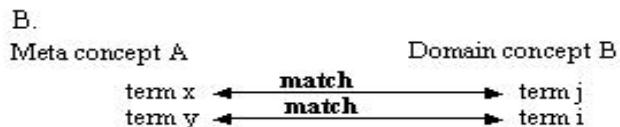
\*Not disambiguated by constraint.

Occurs for cases of redundancy, concepts w/ NOS/‘other’, broad constraints.

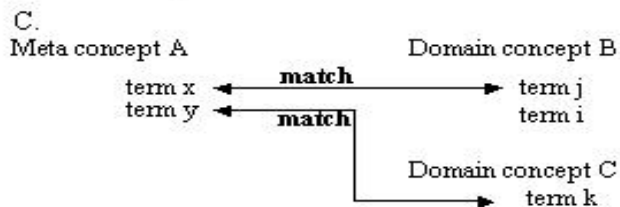
## Valid exact string matches



1 Synonymous Match



>1 Synonymous Match



Meta concept linked to

>1 MHS Domain concept

\*Occurs for ‘untruth’ synonyms.



# Medical Terminologies/Ontologies

## Different Perspectives

LinkFactory Workbench 2.4 (Mariana Casella Dos Santos on rmi://192.168.0.244:7000/LFserver)

File Layout Beans Bookmarks Tools Windows Help

ConceptTree

ENTITY  
CODING & CLASSIFICATION CATEGORY  
STANDARD CODING & CLASSIFICATION CATEGORY

- MEDCIN ENTITY
  - MEDCIN : 102905 : TESTS
  - MEDCIN : 2952 : SYMPTOMS
  - MEDCIN : 39448 : DIAGNOSES, SYNDROMES AND CONDITIONS
  - MEDCIN : 40000 : THERAPY
  - MEDCIN : 5141 : HISTORY
  - MEDCIN : 6000 : 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 : 6001 : VITAL SIGNS
    - MEDCIN : 6142 : 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 : 208834 : ASSESSMENT OF INDWELLING CATHETER
      - MEDCIN : 209291 : ASSESSMENT OF INJECTION SITE
      - MEDCIN : 209344 : 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

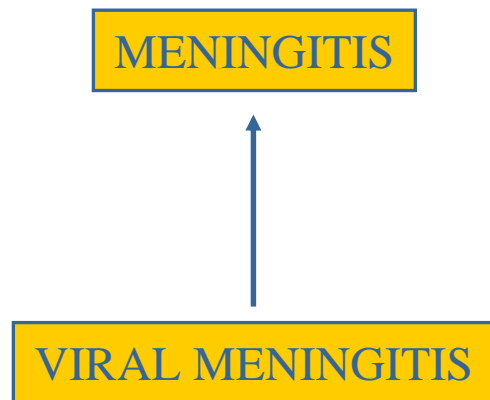
ConceptTree

- SNOMED-CT : 91334004 : CONDYLOMA LATUM (DISORDER)
- SNOMED-CT : 93458008 : FOREIGN BODY IN SKIN (DISORDER)
- SNOMED-CT : 95319004 : SKIN NODULE (DISORDER)
- SNOMED-CT : 95322002 : EDEMATOUS SKIN (DISORDER)
- SNOMED-CT : 95339000 : DISTICHIASIS (DISORDER)
- SNOMED-CT : 95347000 : SKIN NECROSIS (DISORDER)
- SNOMED-CT : 95504004 : ECTOPIC CILIA OF EYELID (DISORDER)
- SNOMED-CT : 95546008 : PERIANAL SKIN TAGS (DISORDER)
- SNOMED-CT : 64572001 : DISEASE (DISORDER)
- SNOMED-CT : 123946008 : DISEASE BY BODY SITE (DISORDER)
- SNOMED-CT : 80659006 : DISEASE OF SKIN AND/OR SUBCUTANEOUS TISSUE (DISORDER)
- SNOMED-CT : 95320005 : DISEASE OF SKIN (DISORDER)
- SNOMED-CT : 95324001 : SKIN LESION (DISORDER)
- SNOMED-CT : 95321009 : FISSURE IN SKIN (DISORDER)
- SNOMED-CT : 367523002 : PERLICHE (DISORDER)
- SNOMED-CT : 10278007 : FACTITIOUS PURPURA (DISORDER)
- SNOMED-CT : 105966009 : TUMOR-LIKE LESION OF SKIN (DISORDER)
- SNOMED-CT : 109550008 : CONGENITAL COMMISSURAL PURPURA (DISORDER)
- SNOMED-CT : 109957002 : DRUG-INDUCED PURPURA (DISORDER)
- SNOMED-CT : 111021003 : NEOVASCULARIZATION OF SKIN (DISORDER)
- SNOMED-CT : 111025007 : HYPERTROPHY OF DERMAL PAPILLAE (DISORDER)
- SNOMED-CT : 111032003 : MACULAR CUTANEOUS AMYLOIDOSIS (DISORDER)
- SNOMED-CT : 123637000 : SEBACEOUS GLAND DUCT DILATATION (DISORDER)
- SNOMED-CT : 123638005 : SWEAT GLAND DUCT DILATATION (DISORDER)
- SNOMED-CT : 123639002 : SWEAT GLAND PORE CLOSURE (DISORDER)
- SNOMED-CT : 123706005 : HAIR DYSTROPHY (DISORDER)
- SNOMED-CT : 126488004 : NEOPLASM OF SKIN (DISORDER)
- SNOMED-CT : 1271009 : KNUCKLE PADS, DEAFNESS AND MENTAL RETARDATION (DISORDER)
- SNOMED-CT : 1515008 : GORHAM'S DISEASE (DISORDER)
- SNOMED-CT : 1521007 : BLISTER OF BUTTOCK WITHOUT INFECTION (DISORDER)
- SNOMED-CT : 1525003 : BLISTER OF FOOT WITHOUT INFECTION (DISORDER)
- SNOMED-CT : 182782007 : DERMATITIS (DISORDER)
- SNOMED-CT : 18503002 : PERINATAL CUTANEOUS HEMORRHOIC CRUSTS (DISORDER)
- SNOMED-CT : 18749008 : CONGENITAL KOILONYCHIA (DISORDER)

# IS-A Link

## Defined Perspective for MHS Ontology

- Axiom: If concept A is-child-of (IS-A) concept B then every instance of A is a instance of B.

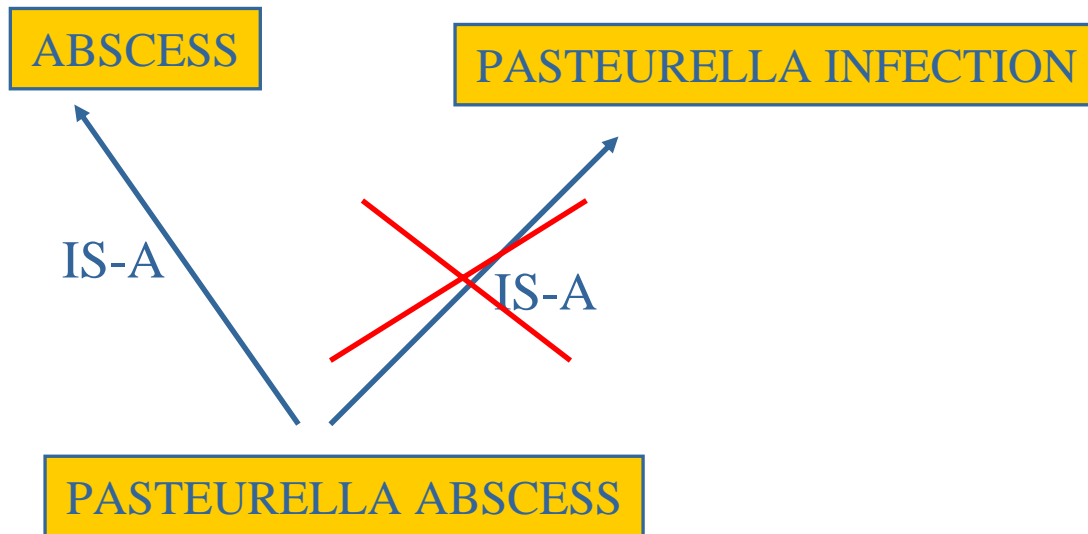


Every instance of viral meningitis is an instance of meningitis.

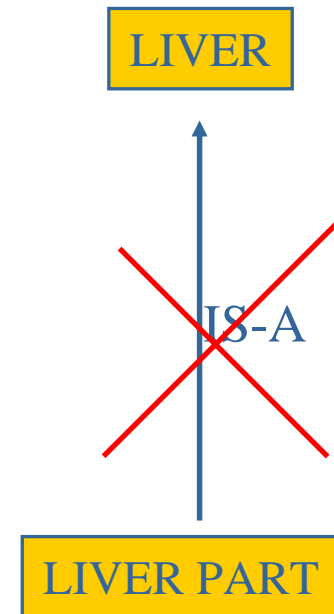


# IS-A Link Common Uses in Medical Terminologies

## Causality



## Parthood



# Transposing Hierarchies from Mapped Terminologies

- ☐ MEDICIN CATEGORY
  - ☐ MEDICIN ENTITY
    - ☐ MEDICIN : 6000 : PHYSICAL EXAMINATION
      - ☐ MEDICIN : 7358 : EXAMINATION OF THE ABDOMEN
        - ☐ MEDICIN : 7436 : ABDOMEN HERNIA
          - ☐ MEDICIN : 10297 : ABDOMEN HERNIA INGUINAL
            - ☐ MEDICIN : 10298 : INGUINAL HERNIA RIGHT
            - ☐ MEDICIN : 10299 : INGUINAL HERNIA LEFT
            - ☐ MEDICIN : 10300 : INGUINAL HERNIA BILATERAL
            - ☐ MEDICIN : 263357 : INGUINAL HERNIA TENDER
            - ☐ MEDICIN : 62509 : INGUINAL HERNIA NOT REDUCIBLE

**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.

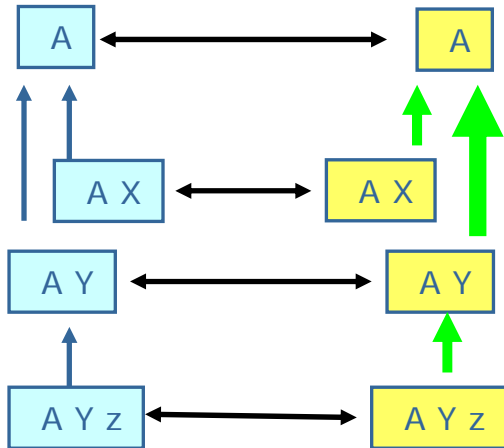
- Analysis of the terminology for the identification of certain 'cutting points' – concepts that are at the top of a small hierarchy with ontologically valid IS-A relations.
- Analysis of the terminology for the identification of hierarchical levels where IS-A relations are always valid.
- Transposition Procedure utilizing 'cutting points' and hierarchical levels identified.



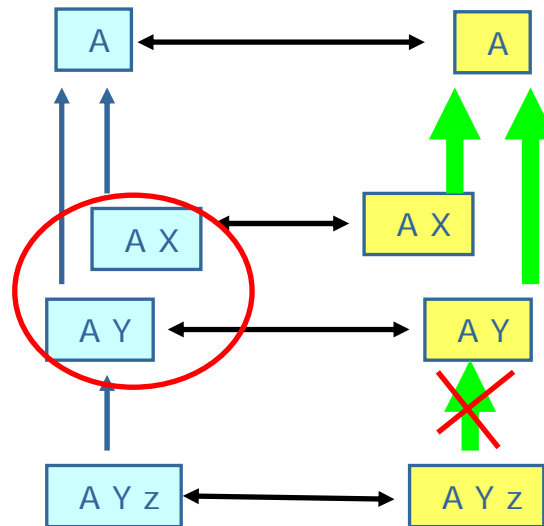
# Transposing Hierarchies from Mapped Terminologies

## 'Cutting Point' Methods:

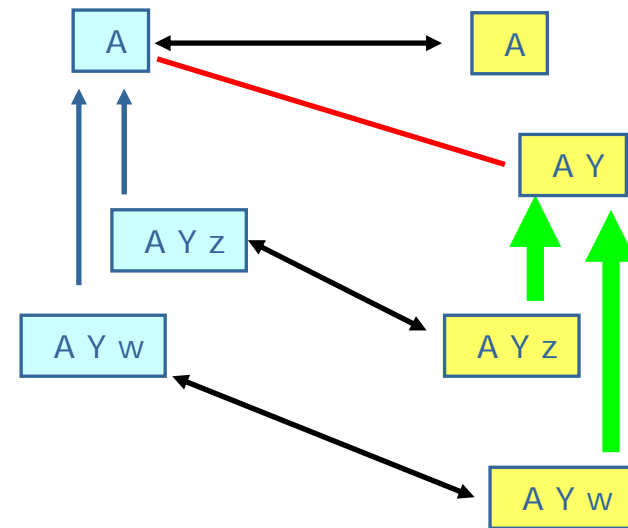
### Transpose to existing map n levels



### Transpose to existing map 1 level



### Transpose through levels of granularity



# Transposing Mapped Terminologies Hierarchies

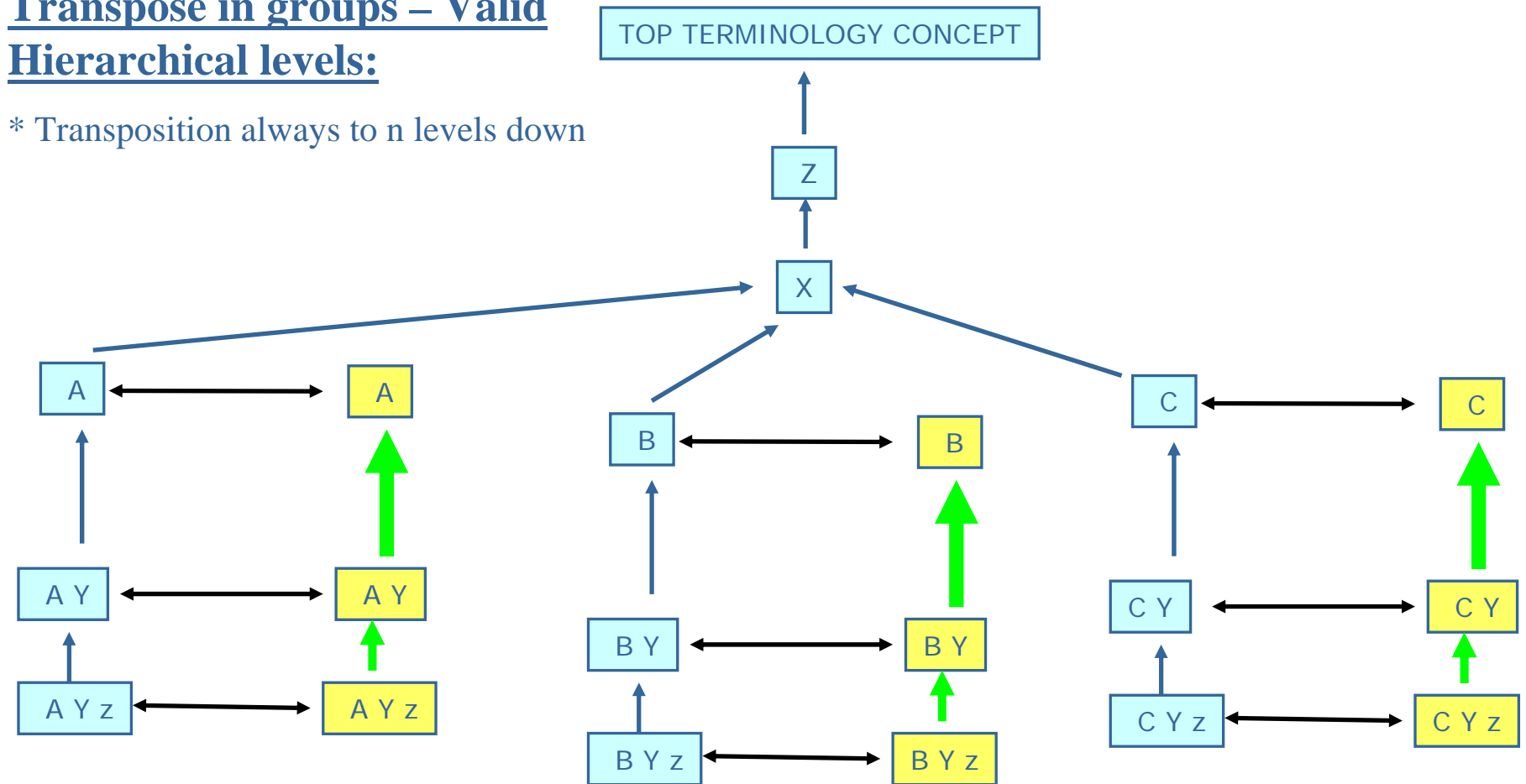
The screenshot displays the LinkFactory Workbench 2.4 interface, showing two side-by-side ConceptTree views. The left view, labeled 'Medcin', shows a hierarchy starting with 'MEDICIN CATEGORY' and 'MEDICIN ENTITY'. A red arrow points to 'MEDICIN : 6000 : PHYSICAL EXAMINATION' (Level 0), and another red arrow points to 'MEDICIN : 178575 : NECK SWELLING' (Level 3). A red box highlights a sub-hierarchy under 'NECK SWELLING', including terms like 'NECK SWELLING RIGHT SIDE', 'NECK SWELLING LEFT SIDE', and 'NECK SWELLING ANTERIOR'. The right view, labeled 'Domain', shows a hierarchy starting with 'CLINICAL FINDING' and 'FINDING BY SITE'. A red arrow points to 'NECK SWELLING' under 'FINDING OF NECK REGION'. A red box highlights a sub-hierarchy under 'NECK SWELLING', including terms like 'NECK SWELLING ANTERIOR', 'NECK SWELLING ANTERIOR INFRAHYOID', and 'NECK SWELLING ANTERIOR PRETRACHEAL'. A large red arrow labeled 'transpose' points from the 'Medcin' view to the 'Domain' view. The bottom of the interface shows search bars for 'Term' and 'KN' in both views.



# Transposing Mapped Terminologies Hierarchies

## Transpose in groups – Valid Hierarchical levels:

\* Transposition always to n levels down



# Transposing Mapped Terminologies Hierarchies

The image displays three windows from a software application, illustrating the mapping of terminology hierarchies. The first window on the left shows a list of MEDCIN categories under 'EDCIN CATEGORY'. The middle window shows a hierarchy of 'ABNORMAL COLOR' and 'DISCOLORATION OF SKIN', with 'ERYTHEMA' highlighted in blue and circled in red. The right window shows a hierarchy for 'SNCT : 247441003 : ERYTHEMA (FINDING)', with 'ERYTHEMA' highlighted in blue. A red arrow points from the circled 'ERYTHEMA' in the middle window to the 'ERYTHEMA' in the right window. A red box highlights the right window's title bar and content. A blue box with the text 'Mapping relationship' is overlaid on the right window.



# 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.



# Mapping Extraction 'Many to One'

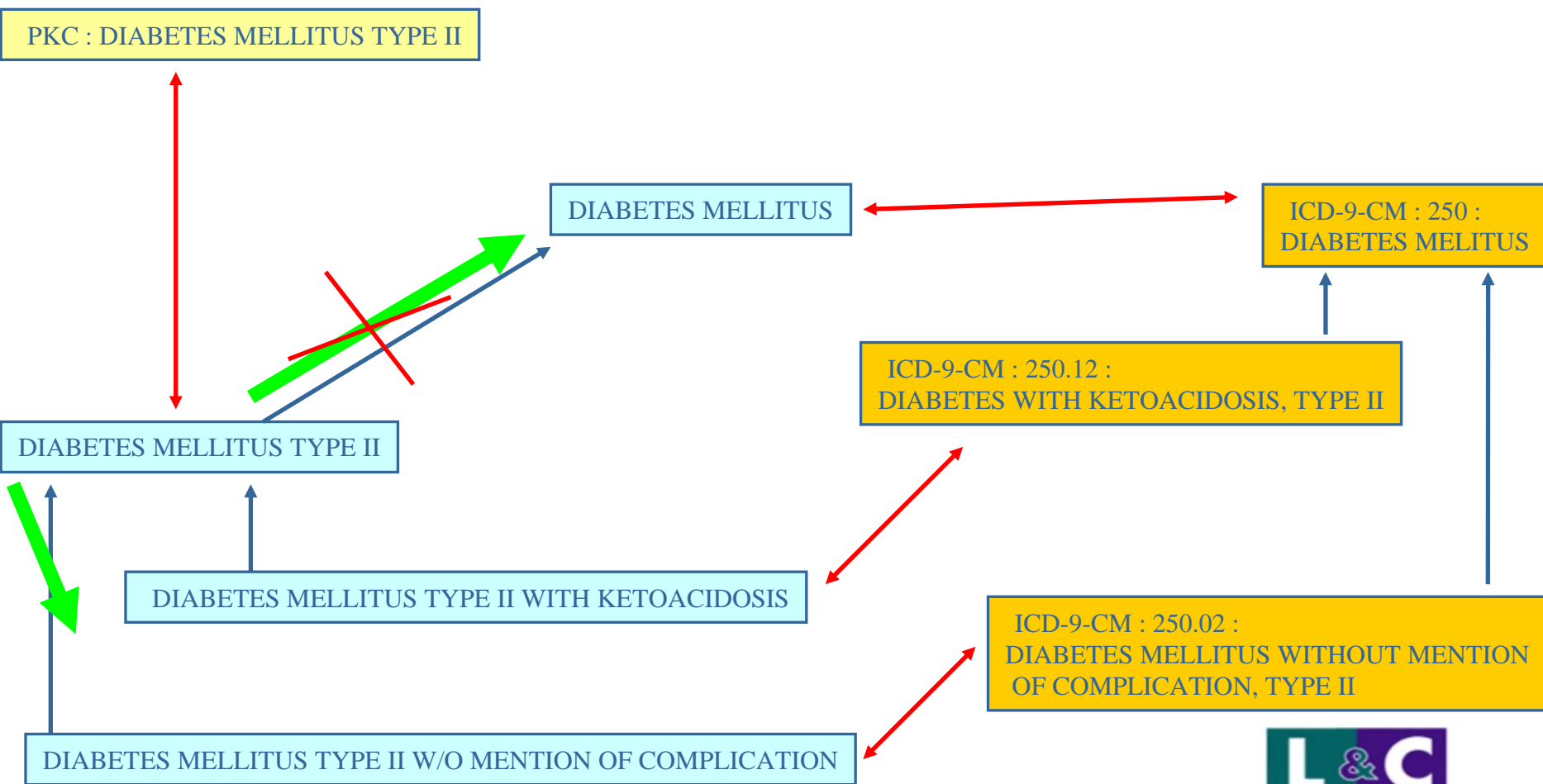
The screenshot displays three panels of a hierarchical medical ontology. The left panel, titled 'EDCIN CATEGORY', shows a tree structure under 'MEDICIN ENTITY' with 'MEDICIN : 167698 : ANKLE ERYTHEMA' circled in red. A red arrow points from this node to the middle panel. The middle panel, titled 'ABNORMAL COLOR', shows 'DISCOLORATION OF SKIN' expanded to 'ERYTHEMA', which is highlighted with a green arrow. A red arrow points from 'ERYTHEMA' to the right panel. The right panel, titled 'SNCT : 247441003 : ERYTHEMA (FINDING)', shows a list of criteria and definitions for 'ERYTHEMA'. A red arrow points from the top of this panel back to the middle panel. The interface includes a 'Save Definitions' button at the bottom right.

\*As MHS domain modeling evolves post-coordinated mappings extraction becomes also possible.





# Mapping Extraction 'One to Many'



# Conclusions

Mapping Processes Key Words  
Reusability and Flexibility





Clinical Vocabulary Mapping Methods Institute  
Saturday, October 15, 2005