Ubiquitous Computing Laboratory
Kyung Hee University, Korea

PhD Dissertation Presentation

Semantic Reconciliation Model for Shareable and Interoperable Medical Knowledge Acquisition

Mr. Taqdir Ali

Department of Computer Science and Engineering

Kyung Hee University, Republic of Korea

Email: taqdir.ali@oslab.khu.ac.kr

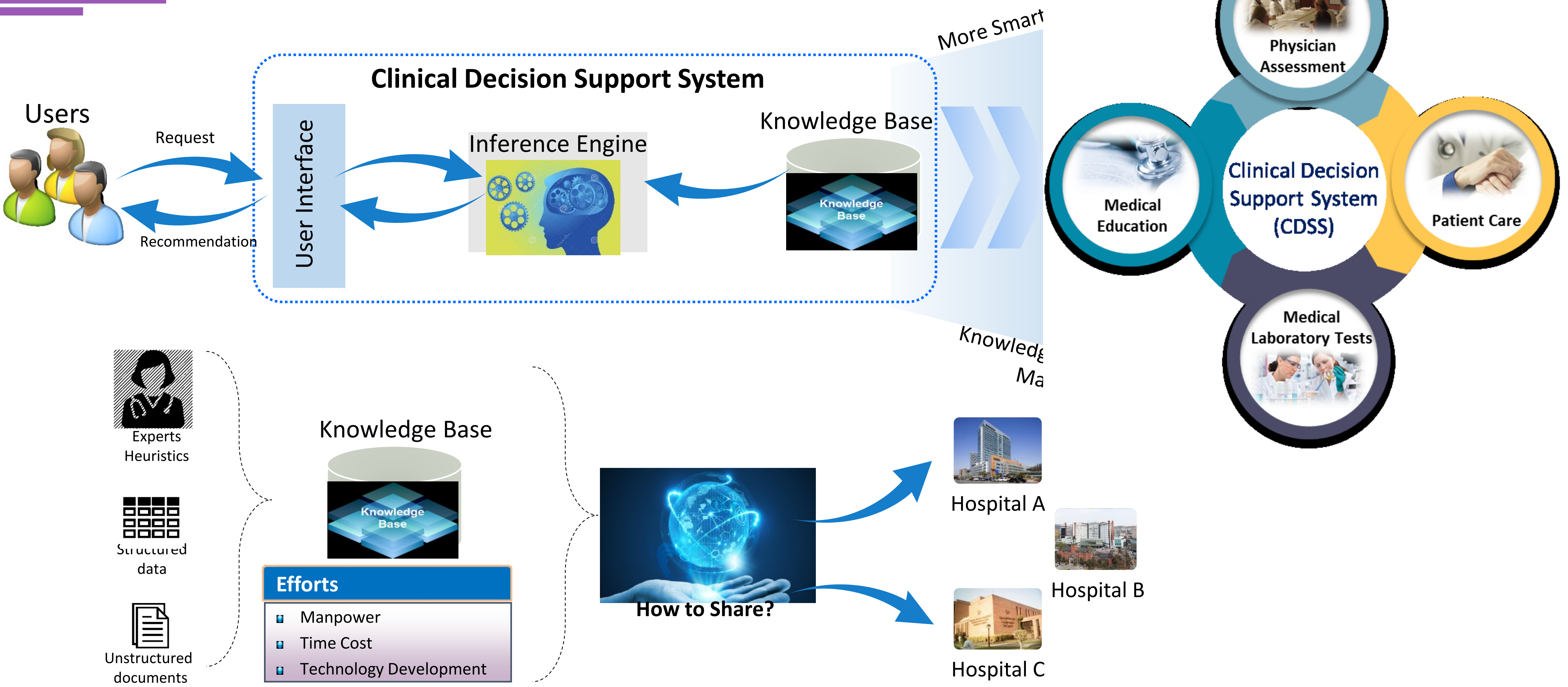
Date: November 26th, 2018

Advisor: Prof. Sungyoung Lee, PhD

Presentation Agenda

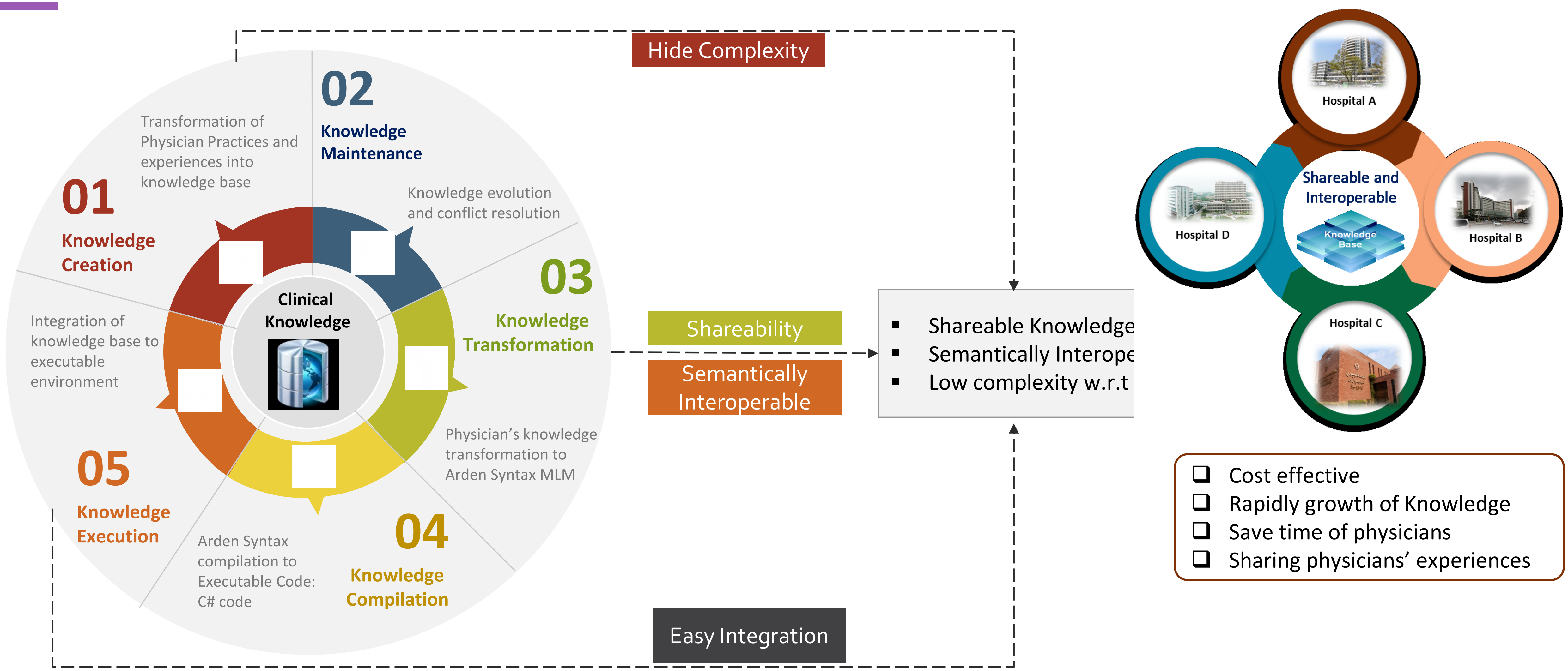
- Introduction
  - Background and Motivation
  - Problem Statement
- Research Scope
- Related Work
- Proposed Methodology
- Solutions
- Experiments and Results
- Uniqueness and Contributions
- Conclusion and Future Research

Background



*Clinical Decision Support System

Knowledge Acquisition Features



- ❑ Cost effective
- ❑ Rapidly growth of Knowledge
- ❑ Save time of physicians
- ❑ Sharing physicians' experiences

Motivation

How to make it Shareable?

Sharing Knowledge

- ❑ If Disease = Head and Neck cancer and Age >= 40
- ❑ Then Treatment Plan = Surgery

How to make it Interoperable?

Hospital A

Hospital B

EMR-1 Schema

Concepts	Schema Concept
Disease	Illness
Head and Neck cancer	Head and Neck cancer
Treatment Plan	Treatment
Surgery	Surgical Treatment

How to Share?

EMR-2 Schema

Concepts	Schema Concept
Disease	Disorder
Head and Neck cancer	HNC
Treatment Plan	Treatment Plan
Surgery	Surgical Procedure

Motivation- Tradeoff among shareability, interoperability and User-friendliness

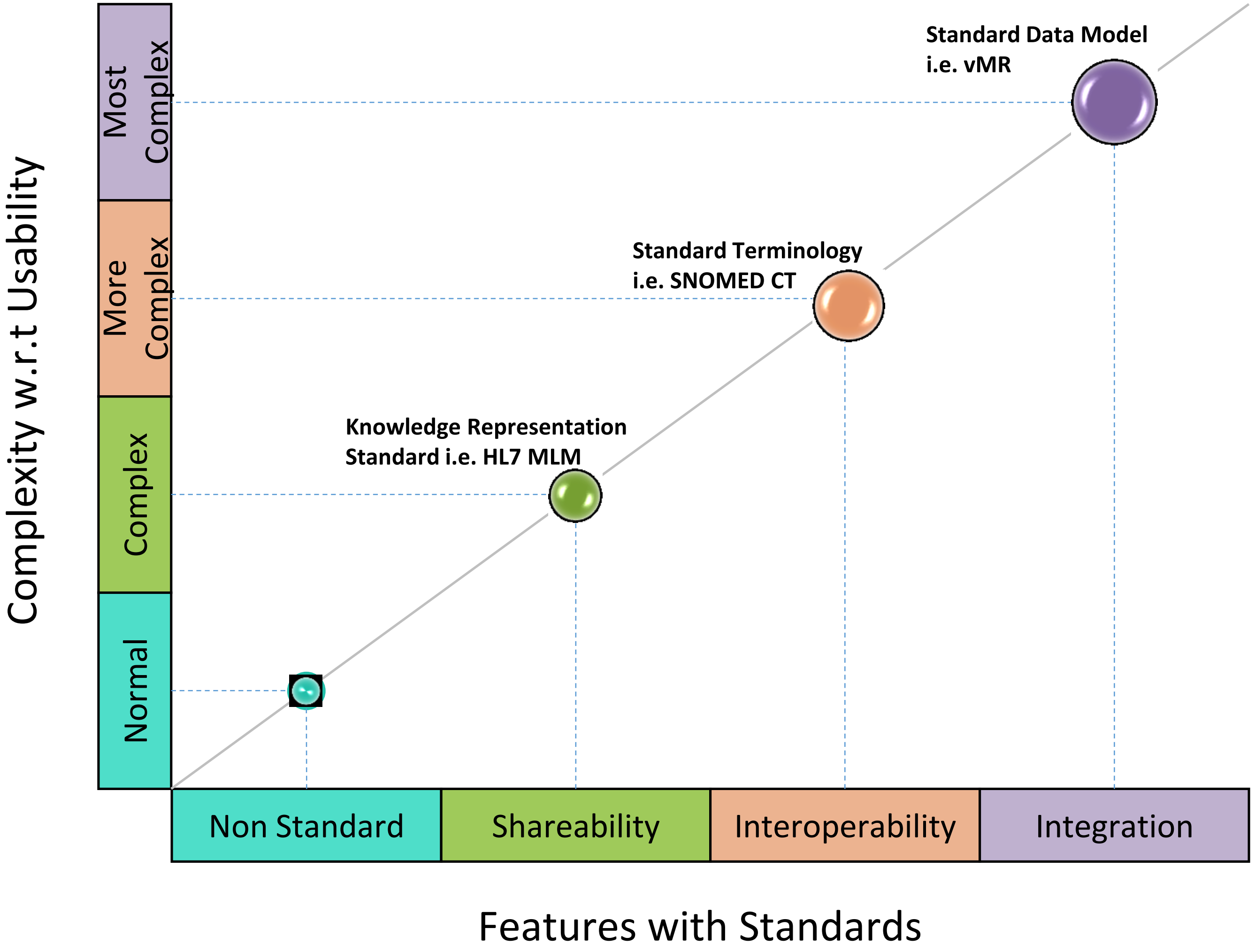
Shareable and reusable Knowledge

is cost and time effective for clinical communities

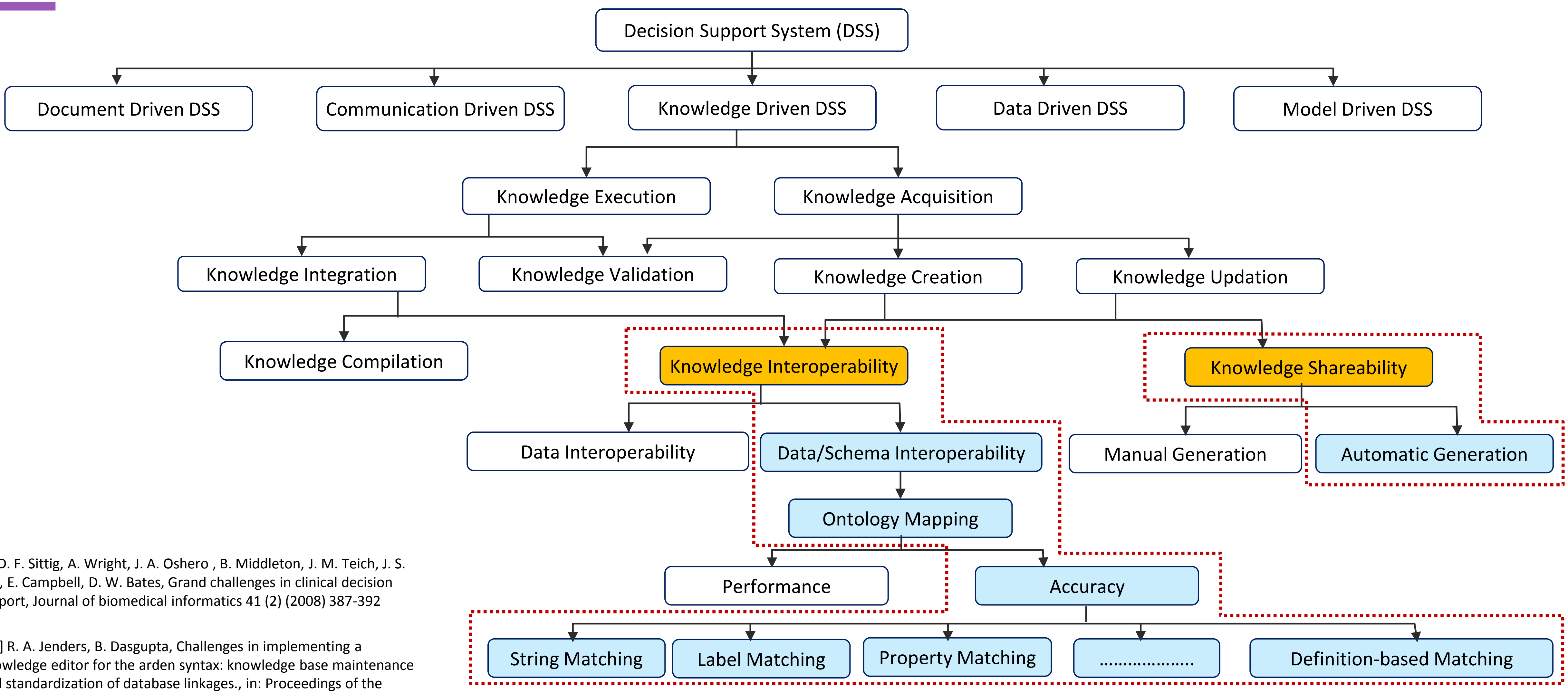
Interoperable knowledge

HL7 vMR and SNOMED CT terminologies makes the integration easy with legacy systems

User friendly rule authoring environment enhances the physician's capability to acquire and maintain knowledge base



Research Scope



[9] D. F. Sittig, A. Wright, J. A. Oshero , B. Middleton, J. M. Teich, J. S. Ash, E. Campbell, D. W. Bates, Grand challenges in clinical decision support, Journal of biomedical informatics 41 (2) (2008) 387-392

[10] R. A. Jenders, B. Dasgupta, Challenges in implementing a knowledge editor for the arden syntax: knowledge base maintenance and standardization of database linkages., in: Proceedings of the AMIA Symposium, American Medical Informatics Association, 2002, p. 355

Related Work (Multi-model mapping – Knowledge Interoperability)

GOMMA [13, 14]

Methodology:
- ❑ Provides component-based GOMMA infrastructure
- ❑ Components are divided into layers of repository, functional components, and tools
- ❑ Hybrid match approaches combine metadata- and annotation-based approaches.

Limitations:
- ❑ Lack of effective mapping representation
- ❑ Lack of definition based mapping
- ❑ Lack of Semantics mapping

LogMapLight [15, 16]

Methodology:
- ❑ Ontology-matching tool address scalability issue
- ❑ Provides better accuracy as compare to other systems
- ❑ Using built-in reasoning and diagnosis capabilities.
- ❑ Implements algorithms for 'on the fly' un-satisfiability detection and repair

Limitations:
- ❑ Lack of definition based mapping
- ❑ It has lack of flexible mapping representation
- ❑ Lack of Semantics Mapping

AgreementMakerLight [17, 18]

Methodology:
- ❑ Provides a combination of flexible and extensible framework with user interface
- ❑ Focused on computational efficiency to handle large ontologies
- ❑ Lexicon based matching for names, labels, and synonyms.

Limitations:
- ❑ No flexible mapping representation
- ❑ Low mapping accuracy
- ❑ Lack of Semantics Mapping
- ❑ Lack of definition based mapping

Falcon-AO++ [19, 20]

Methodology:
- ❑ Provides fundamental technologies for finding, aligning, and learning ontologies using divide-and-conquer approach
- ❑ Using linguistic, structural, partitioned block, similarity combination strategies for matching and aligning.
- ❑ Have highest accuracy in OAEI campaigns.

Limitations:
- ❑ Lack of flexible mapping representation
- ❑ Lack of Semantics Mapping
- ❑ Lack of definition based mapping

Related Work (Multi-model mapping – Knowledge Interoperability)

● Implemented ○ Not Implemented

Systems	Mapping Schemes	Semantics Mapping	Flexibility in mapping Rep.	Definition-based mapping	Accuracy of mapping
GOMMA [13] [14]	●	○	○	○	●
LogMapLight [15] [16]	●	○	●	○	○
AgreementMakerLight [17] [18]	●	●	○	○	●
Falcon-AO++ [19] [20]	●	○	●	○	○
Proposed System	●	●	●	●	●

Related Work (Knowledge Creation – Knowledge Shareability)

UMLS-Based [1]

Methodology:
- Rule Editor for MLM creation based on UMLS terminologies
- Multi-phases selection of concepts in rule creation
- Achieve values level of interoperability.

Limitations:
- Complex interfaces for rule creation
- Lack of interoperability to integrate with legacy systems

XML-based KAT [2]

Methodology:
- Knowledge represented in XML-based knowledge repository
- Embedding data and knowledge into XML files simultaneously
- XML files of knowledge and data are considered as shareable

Limitations:
- Complex structure of knowledge and data storage in XML files
- XML is only preferable of data exchange
- Lack of interoperability

Rule Editor – Arden Syntax [3]

Methodology:
- Provides well-organized tool to create MLM
- Three phase process based on library, maintenance, and knowledge
- Improved work for bibliographic and standardized database linkages

Limitations:
- Depends on local controlled vocabulary
- Lack of easy integration and interoperability
- Complex interfaces for rule creation

ArdenSuite [5, 12]

Methodology:
- Commercialized tool integrated AS development and test environment
- Working on resolving curly brace using GELLO and vMR.
- Provides different slots to write MLM.

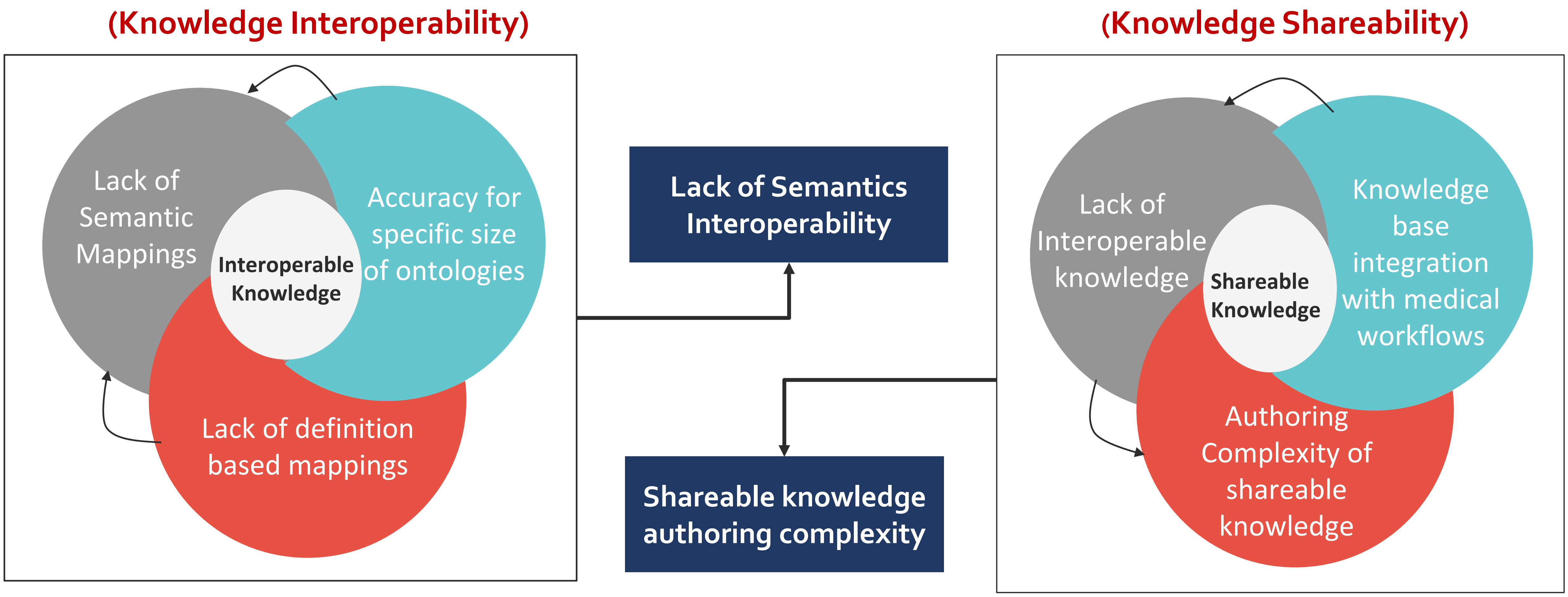
Limitations:
- Lack of interoperability and integration
- Physicians are responsible to remember structure and syntax of MLM
- Lack of standard terminology for easy integration

Related Work (Knowledge Creation – Knowledge Shareability)

● Implemented ○ Not Implemented

Approaches	Shareable KB (Arden Syntax)	Interoperability (vMR)	Standard Terminology	User Friendly	Scope of System	Remarks
KAT (UMLS-Based) [1]	●	○	●	○	Specific Domain	Construction of Domain Ontology By Clinicians
KAT (XML- Based) [2]	○	○	○	○	Extendable	Executable information in CDA
Rule Editor-Arden Syntax [3]	●	●	○	○	Specific Domain	Clinician understandability to vMR
KAT (Anesthesia) [4]	○	○	○	●	Specific to Anesthesia	XML file for Parameters and values
ArdenSuite [5,12]	●	Depends on Physician's skills	○	○	Extendable	Physician's understandability to vMR
I-KAT	●	●	●	●	Extendable	Standard rules based on standard HL7 vMR and SNOMED CT terminologies

Limitation of Existing Work



Problem Statement and Solution

Answer:

- **Knowledge is growing with new research, innovations, daily practice, new disease, and new treatment methodologies.**
- **Knowledge is heterogeneous and non-shareable**
- **HL7 community designed MLM as standard knowledge representation.**
- **But MLM structure and syntax is complex to follow by physician.**
- **MLM structure, syntax, terminologies, and data models overburden the physicians.**

Novel Research and Innovations

Daily Real Practices

New Diseases

New Treatment Methods

Knowledge Base

- ❑ Frequent Knowledge Evolution

How to make it Shareable?
How to make it Interoperable?

- ❑ Heterogeneous Knowledge
- ❑ Non-shareable Knowledge

Sharing Knowledge

- ❑ If Disease = Head and Neck cancer and Age >= 40
- ❑ Then Treatment Plan = Surgery

Physician 2: Knowledge

Concepts	Schema Concept
Disease	Illness
Head and Neck cancer	Head and Neck cancer
Treatment Plan	Treatment
Surgery	Surgical Treatment

How to Share?

Physician 1: Knowledge

Concepts	Schema Concept
Disease	Disorder
Head and Neck cancer	HNC
Treatment Plan	Treatment Plan
Surgery	Surgical Procedure

Problem Statement and Solution

Problem

Current knowledge acquisition methods produce shareable knowledge [1][2][3] with high authoring complexity [4][5] and minimal interoperability [6][7] support.

Thesis Goal

Providing flexible/robust mapping model to enabling easy-to-use knowledge authoring environment for accurate, shareable and interoperable knowledge.

Solution

- To design and develop mapping model to ensure accurate knowledge creation.
- To develop methods for automated creation of a shareable and Interoperable knowledge base

Way of Solution

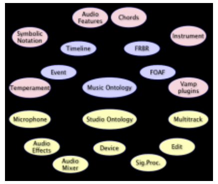
- How to map the concepts of standard data model and terminology based on definition?
- How to generate shareable knowledge with high level abstraction to hide authoring complexity?

Problem Statement and Solution

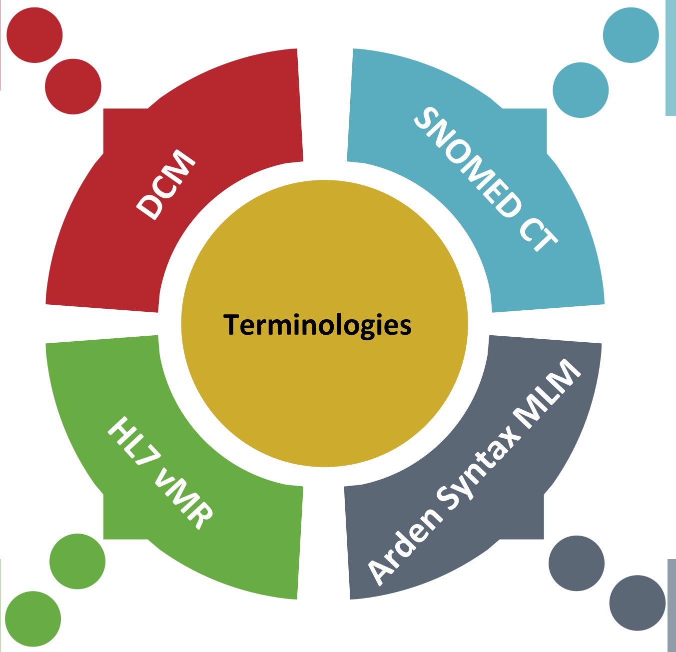
	Knowledge Heterogeneity	Knowledge Non-shareability	Knowledge Authoring Complexity
Problem	Knowledge Heterogeneity	Knowledge Non-shareability	Knowledge Authoring Complexity
Solution	Knowledge Interoperability	Knowledge Shareability	User-friendly Authoring Environment
Methodology	It is achieved using Multi-Modal Mapping among standard and non-standard terminologies, data models, and localized ontologies	The shareability is achieved through standard Medical Logic Module (MLM) representation.	It is achieved through auto generation of MLM by hiding structural and syntax complexity
Results (Quantitative)	**Precision, Recall, F-Measure:** The multi-model Mappings are measured in form of Precision, Recall, and F-Measure ▪ A) Pre-0.95, Rec-0.92, FM-0.93 ▪ B) Pre-0.89, Rec-0.97, FM-0.93	**Success Rate:** Errors are measured in terms of syntax and semantics. ▪ In term of syntax, our proposed approach is error-free, while semantic errors reduces our Success Rate to 90.63%, which is better than Existing system's Success Rate 46.87%	**Performance: Efficiency** User-friendliness is measured through physicians' performance w.r.t Average number of tasks completed successfully per minute. ▪ Our system's Average Efficiency 56.625 completion rate/time

Used Elements

- **Domain Clinical Model (DCM)** is clinical terminologies that are using in existing **HMIS** systems to represent some **specific domain**
- It is designed and developed from **HMIS concepts** using Clinical Information Modelling Process **(CIMP)**
- This vocabulary is **user-friendly** and easily **understandable** to physicians



DCM



- **SNOMED Clinical Terms** provide **standardized** way to represent **clinical terminologies**
- It is clinically **validated**, semantically **rich**, controlled vocabulary, uses to achieve **shareability** and **interoperability**
- SNOMED CT supports the development of **comprehensive** high-quality **clinical content**

SNOMED INTERNATIONAL

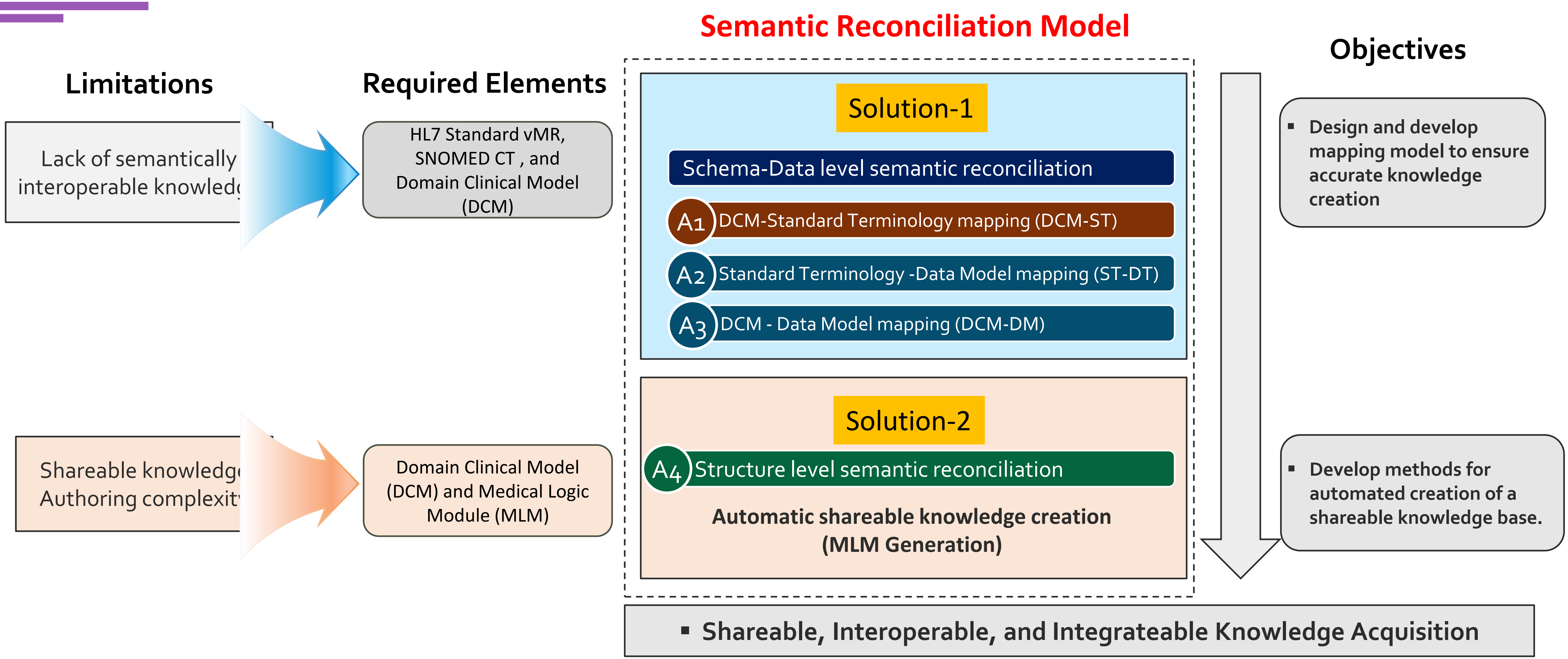
HL7 **vMR**

- **Virtual Medical Record (vMR)** is **data model** for representing medical data relevant to CDSS
- It entails providing clinicians and patients with **clinical knowledge** and patient-related **information**
- It is **simple** and **intuitive** representation of data that is **easy and safe** for CDSS **integration** with HMIS systems

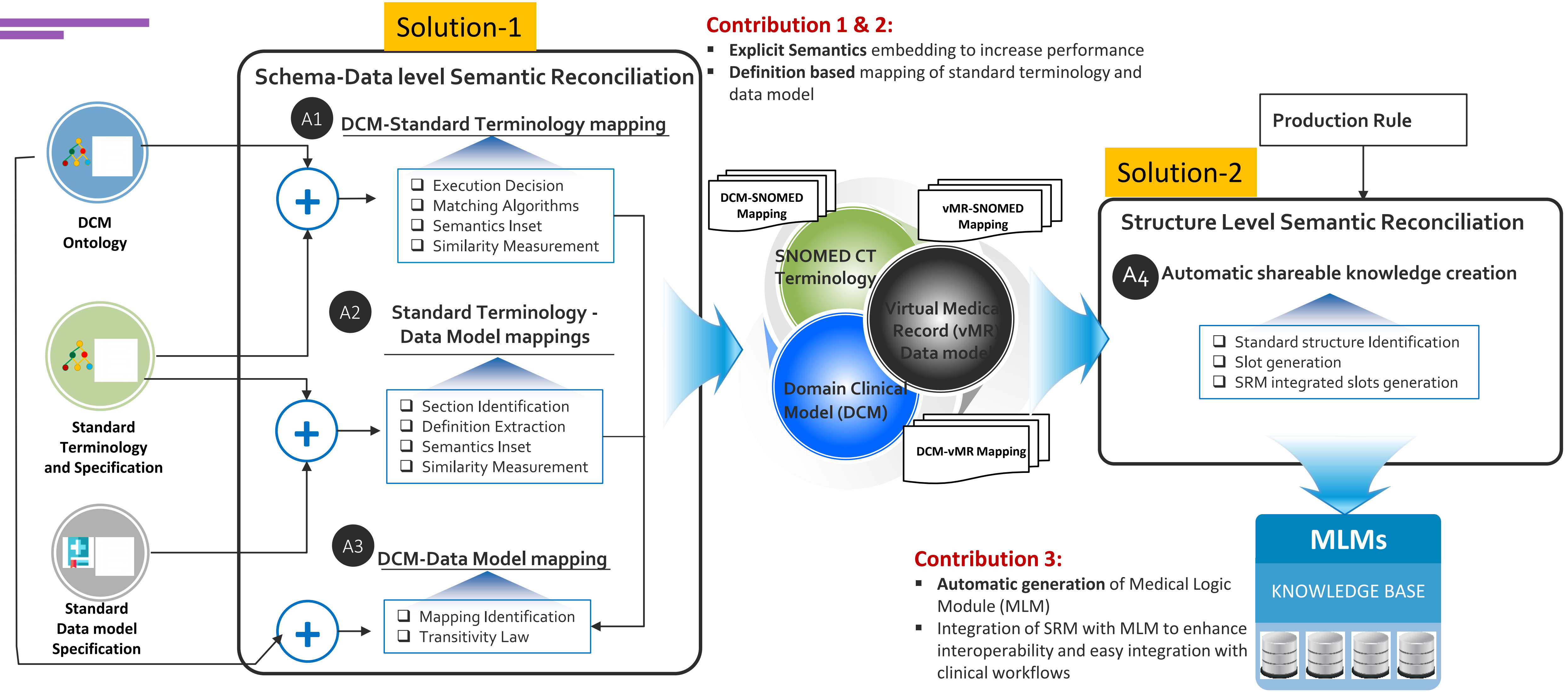
HL7 **MLM**

- **Medical Logic Module (MLM)** is independent module that encodes and represents medical knowledge.
- **Arden Syntax** is a language for writing **MLMs**.
- It is **standard representation** of knowledge to **share** medical knowledge among medical experts and community.

Proposed Idea



Proposed Methodology: Conceptual Model



Contribution 1 & 2:

- **Explicit Semantics** embedding to increase performance
- **Definition based** mapping of standard terminology and data model

Contribution 3:

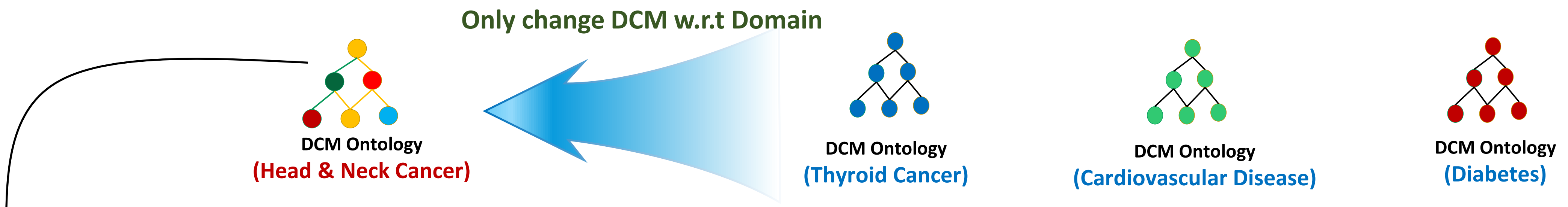
- **Automatic generation** of Medical Logic Module (MLM)
- Integration of SRM with MLM to enhance interoperability and easy integration with clinical workflows

System's Domain Flexibility

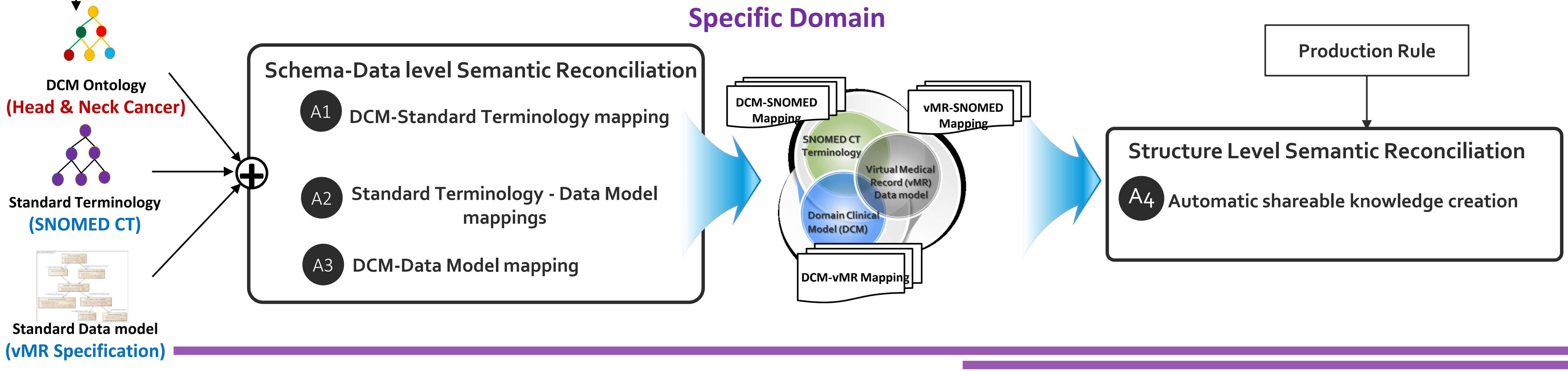
Answer:

- Knowledge shareability and interoperability are problems faced by every medical domain.
- Achieving these objectives lead to complexity in knowledge authoring.
- A case study of **Head & Neck cancer** domain is executed to realized the proposed system.

Domain Generalization



Specific Domain



Map of Thesis Idea

INPUT → Solution-1 → Solution-2 → OUTPUT

INPUT

Ontologies

1. Standard Ontology
2. Standard data model
3. Localized Ontology DCM

1. Schema-Data Level Semantic Reconciliation – (Interoperability)

Solution 1-1 DCM-Standard Terminology Mapping

Solution 1-2 Standard Terminology - Data Model Mapping

Solution 1-3 DCM – Data Model Mapping

Related Work

1. GOMMA [13, 14]
2. LogMapLight [15, 16]
3. AgreementMakerLight [17, 18]
4. Falcon-AO++ [19, 20]

1. Base-line Algorithms
   1. Cosine Similarity
   2. Jaccard Similarity
   3. Overlap Similarity

Problems

- ❑ Lack of semantics mapping
- ❑ Low Accuracy in standard and non-standard ontologies

- ❑ Lack of definition-based mapping

Solutions

- ❑ **Algorithm 1.** DCM-Standard Terminology Mapping (DCM-ST Mapping)
- ❑ **Algorithm 2.** Standard Terminology – Data Model Mapping (ST-DM) Mapping)
- ❑ **Algorithm 3.** DCM – Data Model Mapping (DCM-DM) Mapping)

Mapping Files

2. MLM Generation (Shareability)

Solution 2 Structure Level Semantic Reconciliation

1. KAT (UMLS-Based) [1]
2. KAT (XML-Based) [2]
3. Rule Editor – Arden Syntax [3]
4. KAT (Anesthesia) [4]
5. ArdenSuite - Medexter [5, 12]

- ❑ Knowledge base integration is difficult
- ❑ Lack of medical standards, no Interoperability
- ❑ High authoring complexity due to complex structure and syntax

- ❑ **Algorithm 4**. Shareable and Interoperable Knowledge (MLM Generation)
  - ❑ **Algorithm 4-1** – Data Slot Generation
  - ❑ **Algorithm 4-2** – Logic Slot Generation
  - ❑ **Algorithm 4-3** – Action Slot Generation

Knowledge (MLM)

OUTPUT

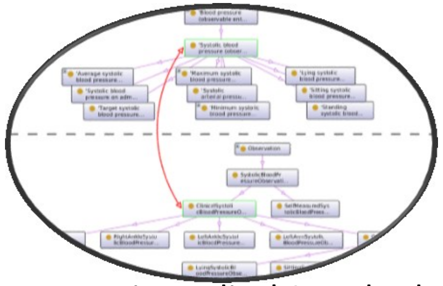
Shareable & Interoperable Knowledge (MLM)

Size of Used Ontologies

- In ontologies matching, size is considered as number of classes, attributes, relationships, individuals. Therefore, following are the size of used ontologies.

1 SNOMED CT

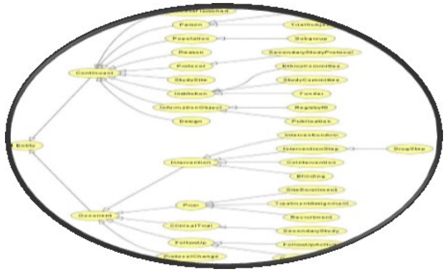
- Classes = 347,358
- Properties = 200
- Maximum Depth = 28
- Maximum Number of Children = 3,355
- Classes with no definition = 343,630



Large Biomedical Standard Ontology (SNOMED CT)

2 Domain Clinical Model (DCM)

- Classes = 214
- Properties = 7
- Maximum Depth = 9
- Maximum Number of Children = 13
- Definition of concepts = All provided



Non - Standard domain clinical model (DCM)

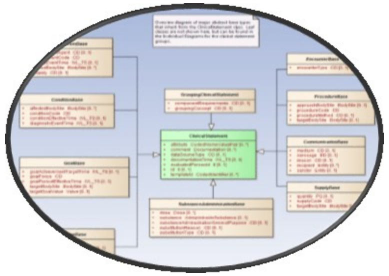
3 Virtual Medical Record (vMR)

Total

- Classes = 94
- Properties = 335
- Maximum Depth = 5

Used for CDSS

- Classes = 69
- Properties = 171
- Maximum Depth = 5



Standard Data model (vMR)

Solution 1: Schema-Data Level Semantic Reconciliation [Background]

Plain Rule:
If Disease = Head and Neck cancer and Age >= 40
Then Treatment Plan = Surgery

HMIS-1

Only work with single clinical system HMIS-1

Plain Rule:
If Disorder = HNC and Age >= 40 Then Treatment Plan = Surgical Procedure

HMIS-2

Only work with single clinical system HMIS-2

Plain Rule:
If Illness = Head and Neck cancer and Age >= 40
Then Treatment = Surgical Treatment

Rules' Facts	Representation in vMR classes and SNOMED CT codes
Disease = "Head and Neck cancer"	ObservationResult.observationFocus = "64572001" and ObservationResult.ObservationValue = "255056009"
Age >= 40	EvaluatedPerson.age >= 40
Treatment Plan >= Surgery	ProcedureEvent.procedureCode = "413737006" and ProcedureEvent.procedureMethod = "387713003"

Interoperable, integrateable

If ObservationResult.observationFocus = "64572001" and ObservationResult.ObservationValue = "255056009" and EvaluatedPerson.age >= 40
Then ProcedureEvent.procedureCode = "413737006" and ProcedureEvent.procedureMethod = "387713003"

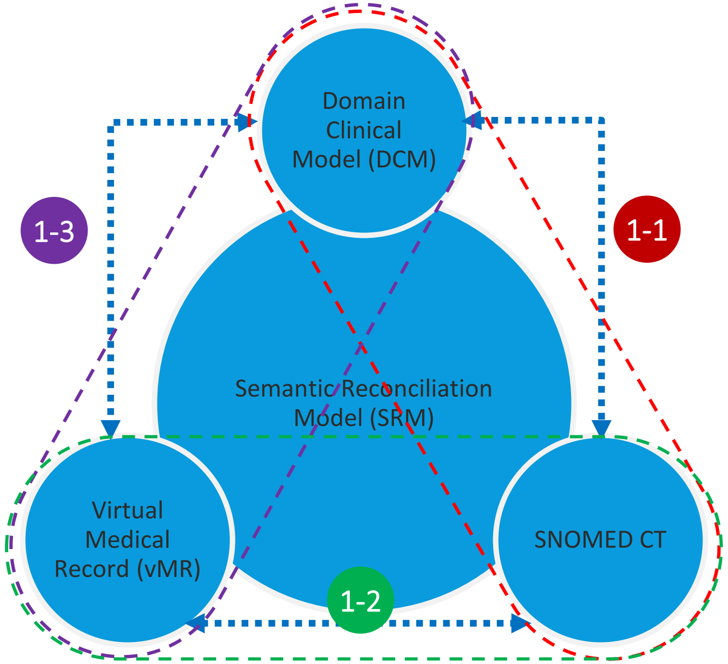
Solution 1: Schema-Data Level Semantic Reconciliation [Motivation]

Plain Rule:
If Disease = Head and Neck cancer and Age >= 40
Then Treatment Plan = Surgery

Comprises → **DCM Concepts**

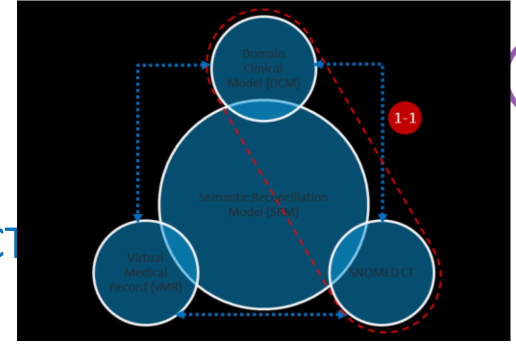
vMR Concepts + **SNOMED CT Concepts**

Transformed rule into clinical standardization:
If ObservationResult.observationFocus = "64572001" and ObservationResult.ObservationValue = "255056009" and EvaluatedPerson.age >= 40
Then ProcedureEvent.procedureCode = "413737006" and ProcedureEvent.procedureMethod = "387713003"



- DCM to SNOMED Mapping:
  - It is mapping between clinical concepts of HMIS systems and SNOMED CT.
  - It enhances the shareability, flexibility, and user friendliness
- DCM to vMR Mapping:
  - It provides mappings between clinical concepts of HMIS systems and vMR schema classes.
  - It increase user friendliness and Interoperability
- SNOMED to vMR:
  - It is mapping between the concepts of SNOMED CT and vMR schema classes.
  - It is useable for shareability, and Interoperability

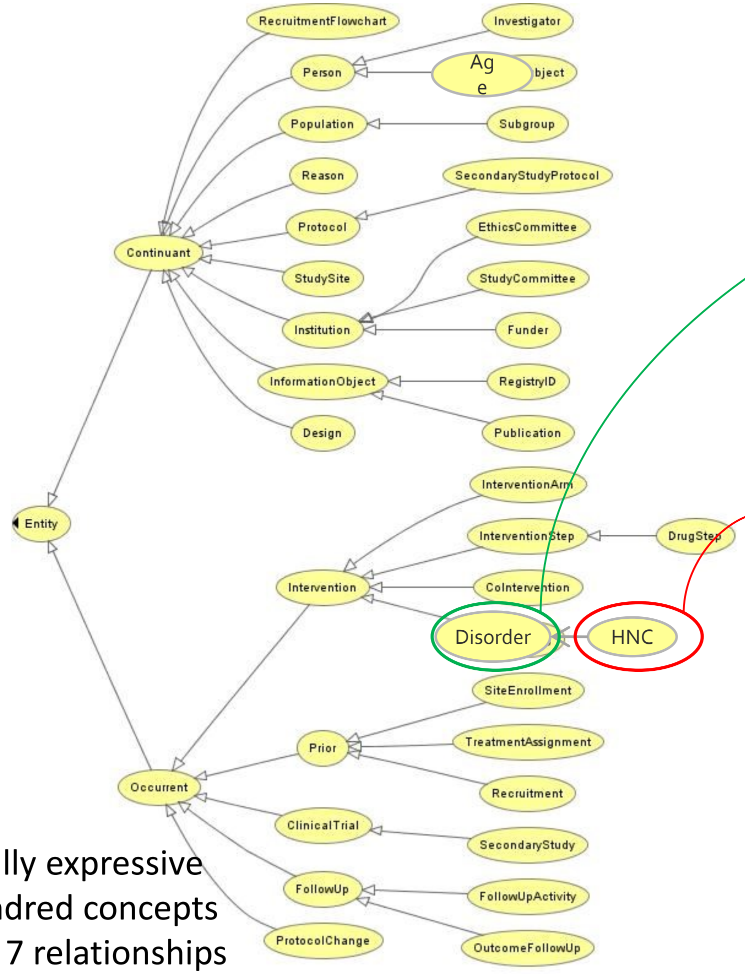
Solution 1-1): DCM-Standard Terminology Mapping [Motivation]

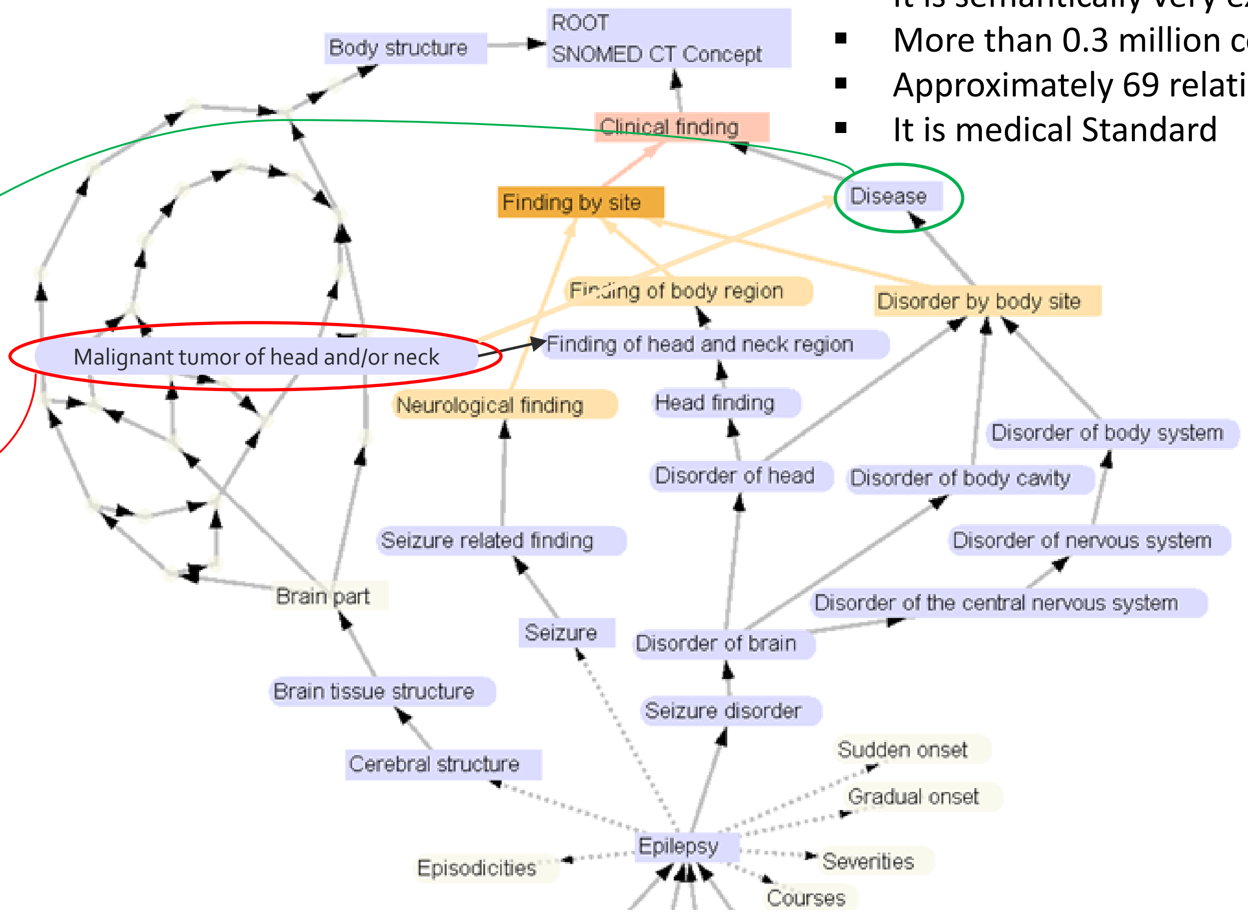


- It is semantically very expressive
- More than 0.3 million concepts
- Approximately 69 relationships
- It is medical Standard

Domain Clinical Model (DCM)

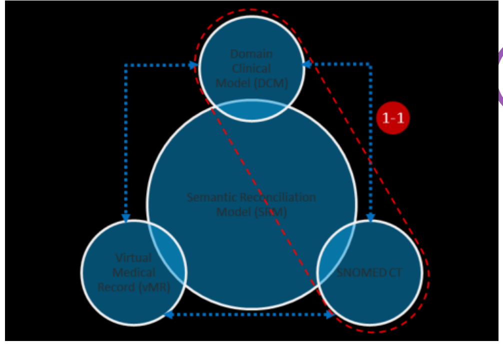
Standard Terminology SNOMED CT





- Less semantically expressive
- More than hundred concepts
- Approximately 7 relationships
- It is localized ontology

Solution 1-1): DCM-Standard Terminology Mapping [Abstract View]



(As – is)

Standard Ontology 1

Standard Ontology 2

1 **Execute matching Algorithms**
String Matching, Label Matching, Child Matching, Property Matching, Synonym Matching

2 **Generate Mappings**
Based on similarity score generates mapping

3 **Verify Mappings**
Using ontology reasoner to verify conflict and duplicate

Mapping Repository

(To – be)

Standard Ontology 1

Non Standard Ontology 2

1 **Execute matching Algorithms**
String Matching, Label Matching, Child Matching, Property Matching, Synonym Matching

Concept Mapped?

Yes

No

2 **Generate Mappings**
Based on similarity score generates mapping

4

3 **Verify Mappings**
Using ontology reasoner to verify conflict and duplicate

5

Mapping Repository

2 **Explicit Semantic Inset**
Concept Preprocessing, Acronyms Inset, Synonym Inset, Hypernym Inset, Hypernym Inset, Meronym Inset

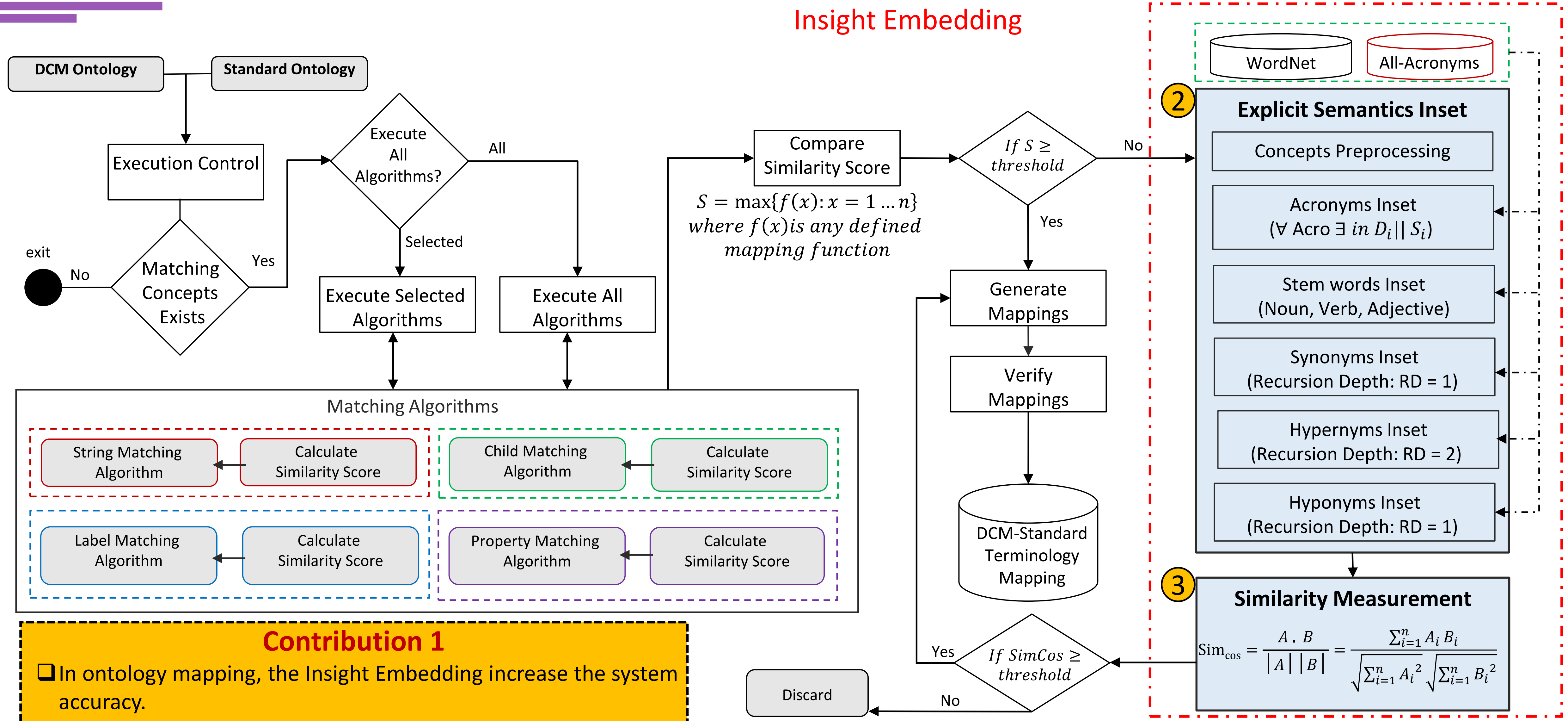
3 **Similarity Measurement**
Measure similarity based on the highest accuracy

Existing Approach Limitations

- ❑ Only focus on the internal semantics of ontology.
- ❑ Lack of definition-based mappings
- ❑ These limitations cause to the low accuracy.

$S = \max\{f(x): x = 1 \ldots n\}$
where f(x) is any defined mapping function

Solution 1-1): DCM-Standard Terminology Mapping [Detail Workflow]



Contribution 1

- In ontology mapping, the Insight Embedding increase the system accuracy.

Algorithm 1: DCM-Standard Terminology Mapping

$C_i \in O_s$

Insight Embedding

Matching Algorithms ③

$$f(x) \leftarrow Ci(O_s) \cap C_j(O_t) \; \forall \; S_k \therefore C_i \in O_s, C_j \in O_t$$

$$g(x) \leftarrow Ci(O_s).label \cap C_j(O_t).label \; \forall \; S_k \therefore C_i \in O_s, C_j \in O_t$$

$$h(x) \leftarrow Child_{i.c}(O_s) \cap Child_{j.c}(O_t) \; \forall \; S_k \therefore Childi_{.c} \in O_s, Child_{j.c} \in O_t$$

$$p(x) \leftarrow Property_{i.c}(O_s) \cap Property_{j.c}(O_t) \; \forall \; S_k \therefore Property_{i.c} \in O_s, Property_{j.c} \in O_t$$

$$M_{apped}C_{oncepts} = \mathbb{R}_{i=1}^{C_n}.\mathbb{R}_{j=1}^{C_n} f(x)$$

$$where \; S_k \in \{set \; of \; string \; matching \; algorithms\}$$

Explicit Semantic Inset ②

$C_i \in O_s$

$$SV(Ci) = [Stem((Ci), Acro(Ci), \; \sum_{k=0}^{n} Syn(C_i), \; \sum_{k=0}^{n} Hyper(C_i), \; \sum_{k=0}^{n} Hypo(C_i), \sum_{k=0}^{n} Mero(C_i),$$

$C_j \in O_t$

$$SV(Cj) = [Stem((Ci), Acro(Cj), \; \sum_{k=0}^{n} Syn(C_j), \; \sum_{k=0}^{n} Hyper(C_j), \; \sum_{k=0}^{n} Hypo(C_j), \sum_{k=0}^{n} Mero(C_j),$$

where $SV(C_i)$ is space vector of source concept and
$SV(C_j)$ is space vector of target concept.
$\sum_{k=0}^{n}$ is dynamic collection of synonyms, hypernyms, hyponyms, and myronyms

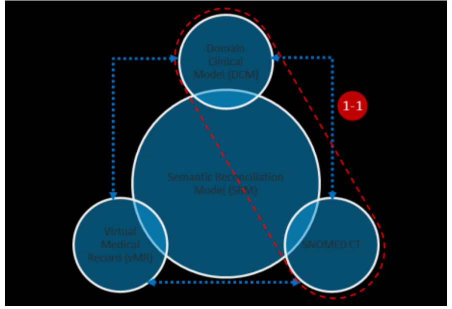
Algorithm 1: DCM and Standard Terminology mapping algorithm

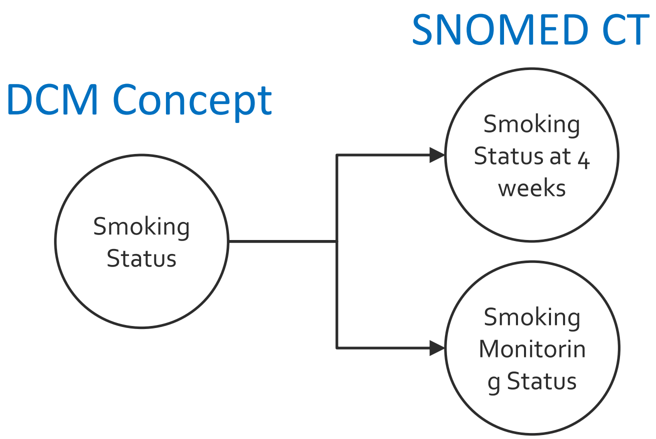
Input : O_s Source Ontology
O_t Target Ontology
Result: $M_{apped}C_{oncepts}$

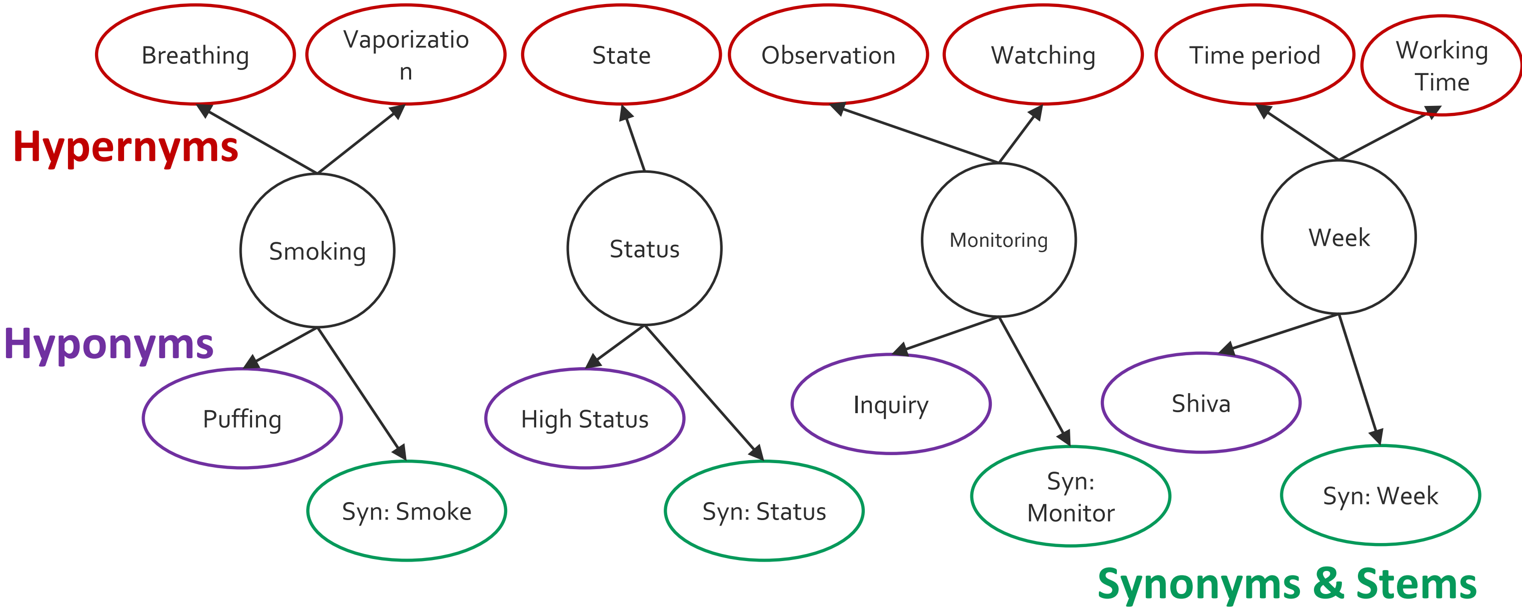
Initialization: InitialSimilarityScore = 0.0;
Initialization: FinalSimilarityScore = 0.0;
foreach C_i *of* O_s **do**
  foreach C_j *of* O_t **do**
    $stringSimilarity$ = stringMatching(C_i, C_j);
    $childSimilarity$ = childgMatching(C_i, C_j);
    $labelSimilarity$ = labelMatching(C_i, C_j);
    $propertySimilarity$ = propertyMatching(C_i, C_j);
    InitialSimilarityScore = $max(stringSimilarity, childSimilarity, labelSimilarity, propertySimilarity)$;
    if *InitialSimilarityScore is less than* $ThresholdValue$ **then**
      $SV(C_i)$ = preprocessConcepts(C_i);
      $SV(C_i) = SV(C_i)$ + InsertExplicitSemantics(C_i);
      $SV(C_j)$ = preprocessConcepts(C_j);
      $SV(C_j) = SV(C_j)$ + InsertExplicitSemantics(C_j);
      $Cosine_{sim} = \frac{\sum_{i=1}^{n} SV(C_i).SV(C_j)}{\sqrt{\sum_{i=1}^{n} SV(C_i)^2}\sqrt{\sum_{i=1}^{n} SV(C_j)^2}}$;
      FinalSimilarityScore = $Cosine_{sim}$;
      if *FinalSimilarityScore is greater than or equal to* $ThresholdValue$ **then**
        $M_{apped}C_{oncepts} \leftarrow C_i(O_s) \bigcap C_j(O_t)$;
        Break;
      end
    end
    else
      $M_{apped}C_{oncepts} \leftarrow C_i(O_s) \bigcap C_j(O_t)$;
      Break;
    end
  end
end

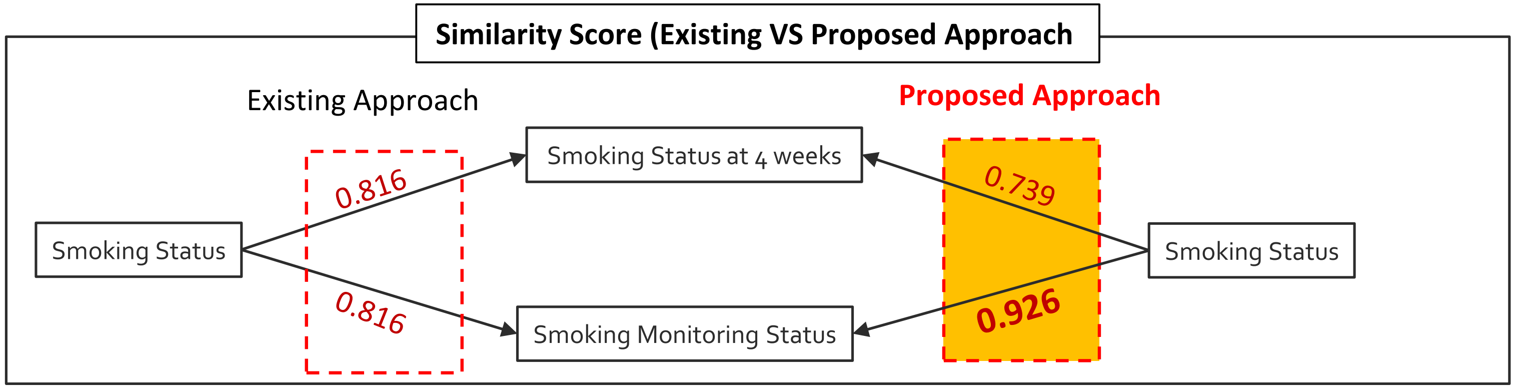
Taqdir Ali, and Sungyoung Lee, "Reconciliation of SNOMED CT and domain clinical model for interoperable medical knowledge creation", 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017), July 11-15, 2017

Solution 1-1): Outcome of Algorithm 1

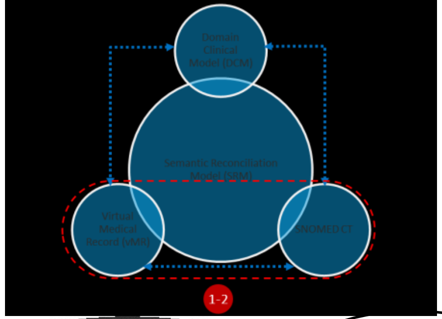






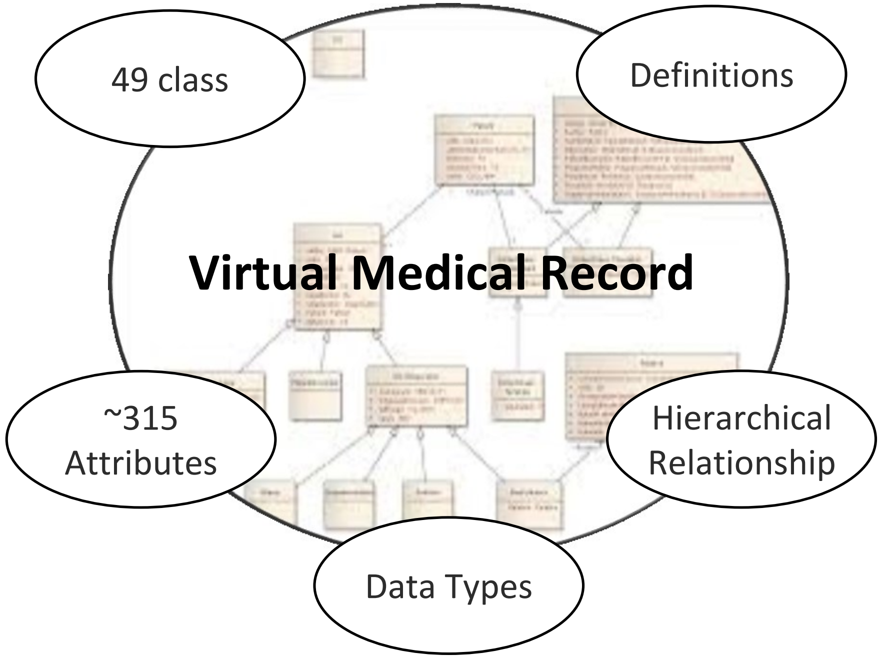


Solution 1-2): Standard Terminology and Data Model Mapping [Motivation]



ObservationFocus

Clinical Findings



<u>String Matching</u>

ObservationFocus **!=** Clinical Findings

<u>Child Matching</u>

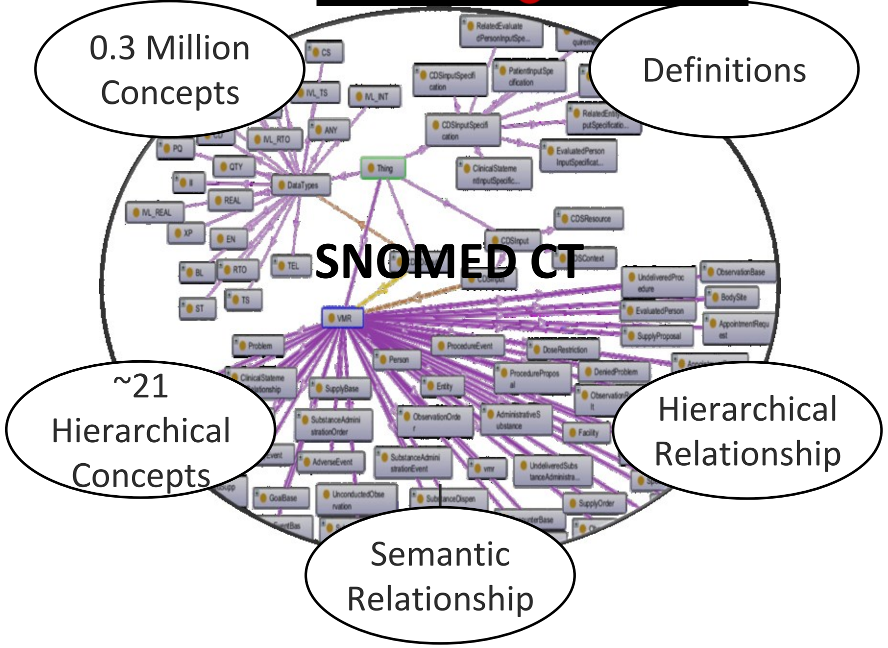
ObservationFocus **!=** Clinical Findings

<u>Label Matching</u>

ObservationFocus **!=** Clinical Findings

<u>Property Matching</u>

ObservationFocus **!=** Clinical Findings

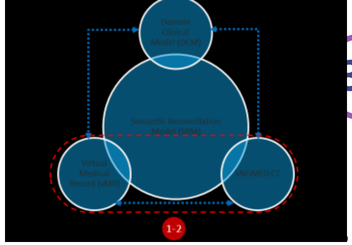


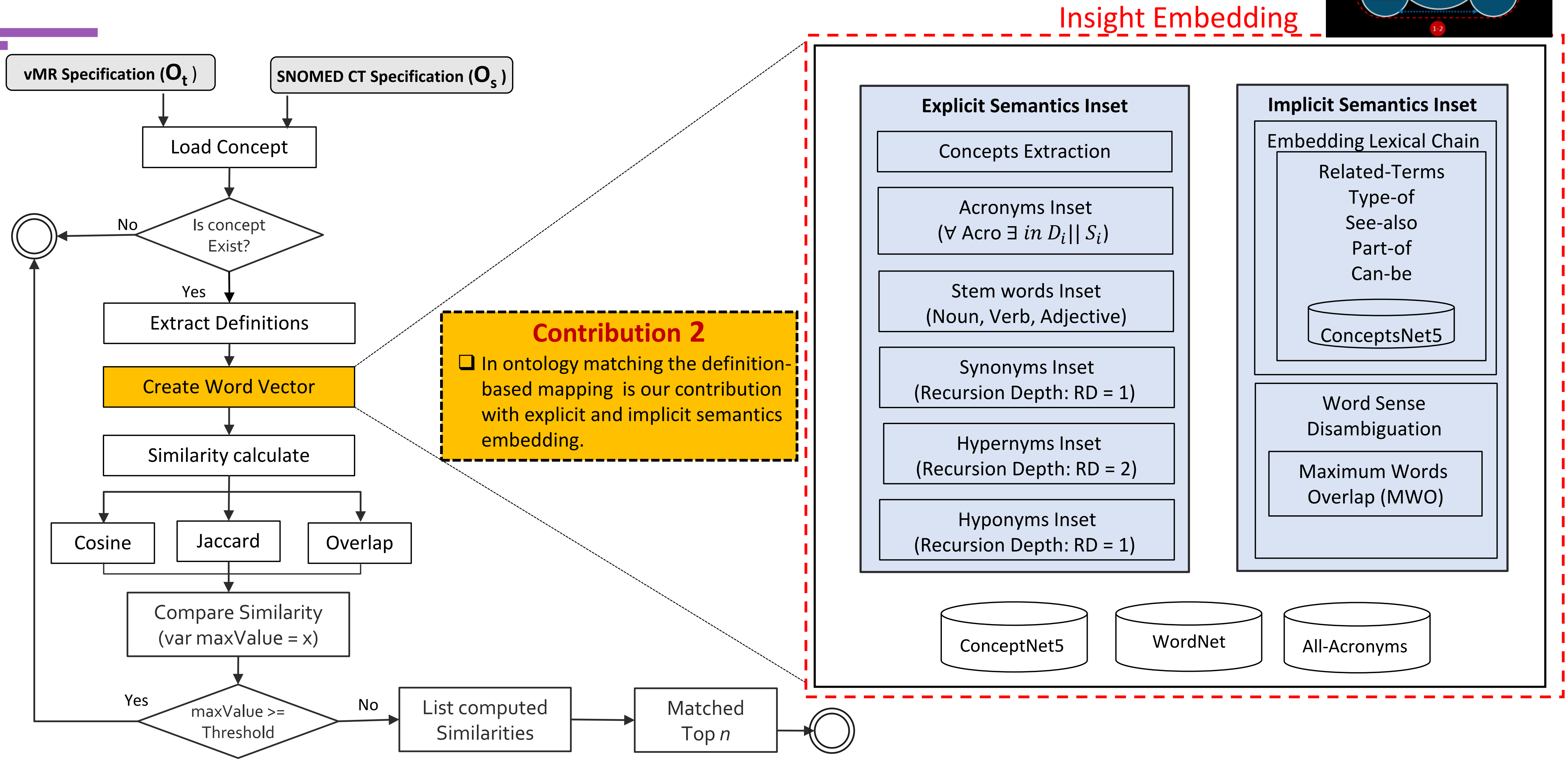
ObservationFocus
The code that identifies the focus of the observation with as much specificity as available, or as required by a template. E.g., serum potassium level, hemoglobin A1c level, smoking status.

≅

Clinical Findings
Represents the result of a clinical observation, assessment or judgment and includes normal and abnormal clinical states (e.g. |asthma|, |headache|, |normal breath sounds|). The |clinical finding| hierarchy includes concept used to represent diagnoses.

Solution 1-2): Standard Terminology and Data Model Mapping [Detail Workflow]





Algorithm 2: Standard Terminology and Data Model Mapping

Explicit Semantic Insets

1

$C_i \in O_s$

$ESV(Ci) = [Stem((Ci), Acro(Ci), \sum_{k=0}^{n} Syn(C_i), \sum_{k=0}^{n} Hyper(C_i), \sum_{k=0}^{n} Hypo(C_i), \sum_{k=0}^{n} Mero(C_i),$

$C_j \in O_t$

$ESV(Cj) = [Stem((Ci), Acro(Cj), \sum_{k=0}^{n} Syn(C_j), \sum_{k=0}^{n} Hyper(C_j), \sum_{k=0}^{n} Hypo(C_j), \sum_{k=0}^{n} Mero(C_j),$

where $ESV(C_i)$ *is explicit space vector of source concept and*
$ESV(C_j)$ *is explicit space vector of target concept.*
$\sum_{k=0}^{n}$ *is dynamic collection of synonyms, hypernyms, hyponyms, and myronyms*

Implicit Semantic Insets (Lexical Chain)

2

$C_i \in Os$

$ISV(Ci) = lc[\sum_{k=0}^{n} Related - Term(Ci), \sum_{k=0}^{n} Type - of(Ci), \sum_{k=0}^{n} Part - of(Ci), \sum_{k=0}^{n} See - also(Ci)]$

$C_j \in O_t$

$ISV(Cj) = lc[\sum_{k=0}^{n} Related - Term(C_j), \sum_{k=0}^{n} Type - of(C_j), \sum_{k=0}^{n} Part - of(C_j), \sum_{k=0}^{n} See - also(C_j)]$

where $ISV(C_i)$ *is implicit space vector of source concept and*
$ISV(C_j)$ *is implicit space vector of target concept.*
$\sum_{k=0}^{n}$ *is dynamic collection of implicit properties*

Vectors Integration

$SV(C_i) = ESV(C_i) + ISV(C_i)$ $\qquad SV(C_j) = ESV(C_j) + ISV(C_j)$

Similarity Measurement

3

$$\text{Sim}_{\text{cos}} = \frac{\sum_{i=1}^{n} SV(C_i).SV(Cj)}{\sqrt{\sum_{i=1}^{n} SV(Ci)^2}\sqrt{\sum_{i=1}^{n} SV(Cj)^2}} \quad \text{Sim}_{\text{jaccard}} = \frac{|SV(Ci) \cap SV(Cj)|}{|SV(Ci) \cup SV(Cj)|} \quad \text{Sim}_{\text{overlap}} = \frac{|SV(Ci) \cap SV(Cj)|}{\min(|SV(Ci)|, |SV(Cj)|)}$$

Algorithm 2: Standard Terminology and data model mapping algorithm

Input : O_s Source Ontology as data model specification
O_t Target Ontology as standard terminology specification
$O_s = \{C_1, C_2, C_3, .., C_i, ..., C_n\}$
$O_t = \{\}C_1, C_2, C_3, .., C_j..., C_n\}$

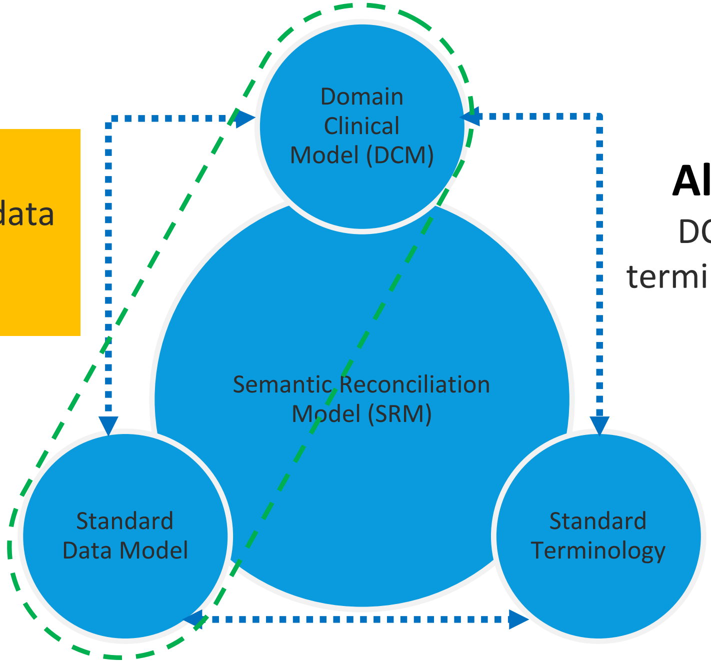
Result: $M_{apped}C_{oncepts}$

```
Initialization: FinalSimilarityScore = 0.0;
foreach C_i of O_s do
    Definition(C_i) = ExtractDefinition(C_i);
    SV(C_i) = preprocessDefinition(C_i);
    SV(C_i) = SV(C_i) + InsertExplicitSemantics(C_i);   1
    SV(C_i) = SV(C_i) + InsertImplicitSemantics(C_i);   2
    foreach C_j of O_t do
        Definition(C_j) = ExtractDefinition(C_j);
        SV(C_j) = preprocessDefinition(C_j);
        SV(C_j) = SV(C_j) + InsertExplicitSemantics(C_j);   1
        SV(C_j) = SV(C_j) + InsertImplicitSemantics(C_j);   2
    end
    Cosine_sim = sum_{i=1}^{n} SV(C_i).SV(C_j) / (sqrt(sum_{i=1}^{n} SV(C_i)^2) sqrt(sum_{i=1}^{n} SV(C_j)^2));   3
    Jaccard_sim = |SV(C_i) ∩ SV(C_j)| / |SV(C_i) ∪ SV(C_j)|;
    Overlap_sim = |SV(C_i) ∩ SV(C_j)| / min(|SV(C_i)|, |SV(C_j)|);
    FinalSimilarityScore = max(Cosine_sim, Jaccard_sim, Overlap_sim);
    if FinalSimilarityScore isgreaterthanorequaltoThresholdValue then
        M_appedC_oncepts <- C_i(O_s) ∩ C_j(O_t);
        Break;
    end
end
```

Algorithm 3: DCM and Data Model Mapping

Algorithm 3
DCM and Standard data model mapping algorithm

DCM-vMR Mapping



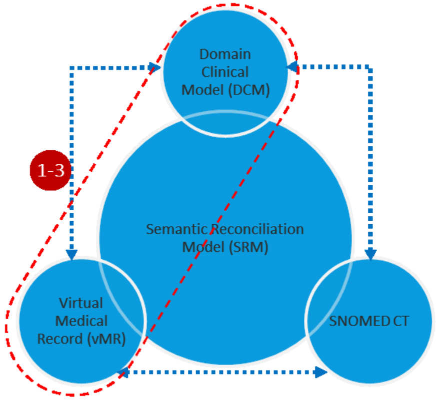
Algorithm 1
DCM-Standard terminology mapping algorithm

DCM-SNOMED Mapping

Algorithm 2
Standard terminology and data model mapping algorithm

vMR-SNOMED Mapping

- Why **reconciliation inferencing**?
  - We already mapped DCM-SNOMED and SNOMED-vMR.
  - Both of these mappings can give results for DCM-vMR automatically by **transitivity law**.



Algorithm 3: DCM and standard data model mapping algorithm

Begin
inputs: **DCM Ontology**, $\boldsymbol{DCM - SNOMED\ mapping}, \boldsymbol{vMR - SNOMED}$
$\mathrm{A} = \{D_1 D_2, \dots, D_\mathrm{n}\}$ // n objects of DCM
$\mathrm{B} = \{DS_1 DS_2, \dots, DS_\mathrm{n}\}$ // n objects of DCM-SNOMED Mapping
$\mathrm{C} = \{\mathrm{v}S_1 \mathrm{v}S_2, \dots, \mathrm{v}S_\mathrm{n}\}$ // n objects of vMR-SNOMED Mapping
output: $Mapping\ File$

foreach **Dc in A**
loadConcept(Dc)
foreach DSc **in** B // scanning of DCM-SNOMED Mapping
if Dc ***is equal to*** DSc // (if A = B)
foreach vSc **in** C // Scanning of vMR-SNOMED Mapping
if vSc ***is equal to*** DSc // (if B = C)
generateMappingIDs(Dc**,** vSc**)** // (means A = C)
Mapping File.Add(Dc, vSc**)**
end if
end foreach
end if
end foreach
end foreach
Return *Mapping File*
End

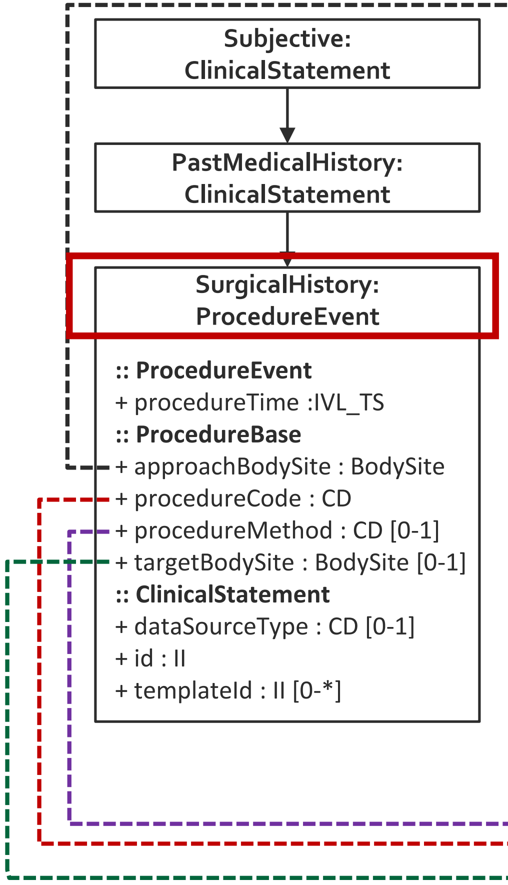
Outcome of Solution 1

Shows vMR classes/attributes

Shows localized concepts of DCM

Shows the SNOMED CT concepts and codes

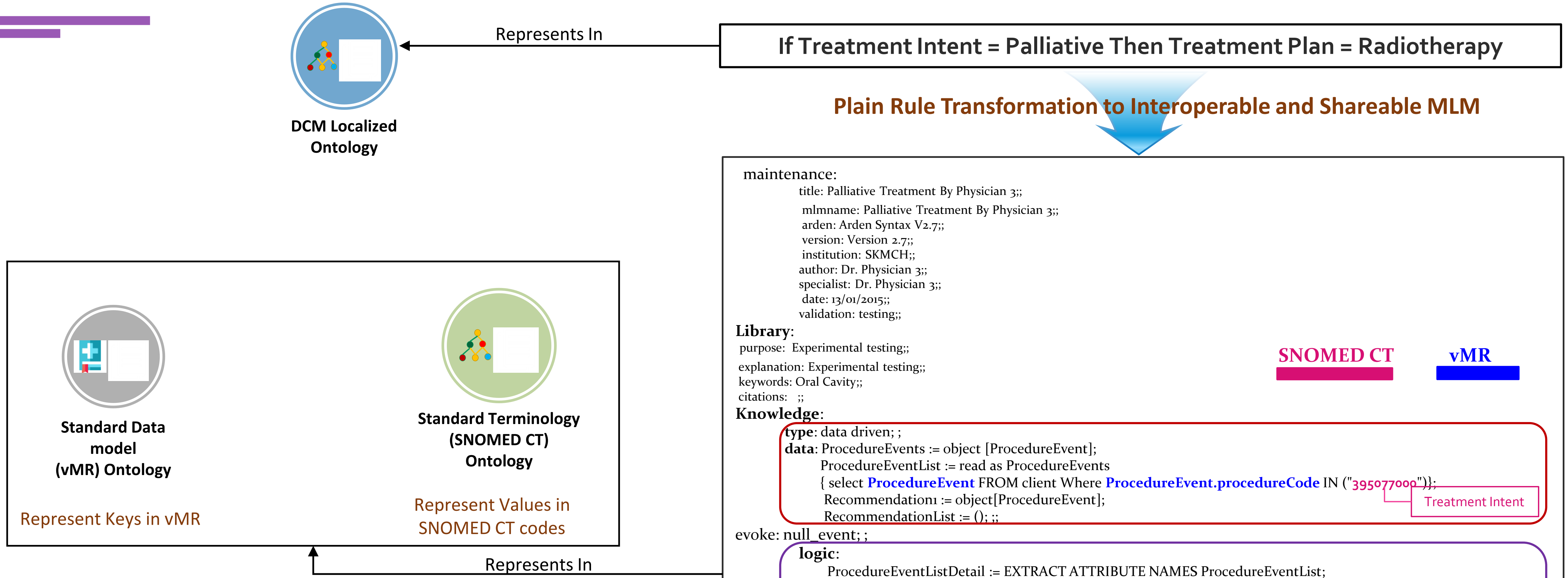
- **Partial Example of the outcome of Solution 1.**
  - It provides three mapping files.
    - DCM – SNOMED CT
    - vMR – SNOMED CT
    - DCM - vMR
  - It is offline process, it is needed once when its going to deploy in hospital.
  - On each new deployment only DCM mapping is required.

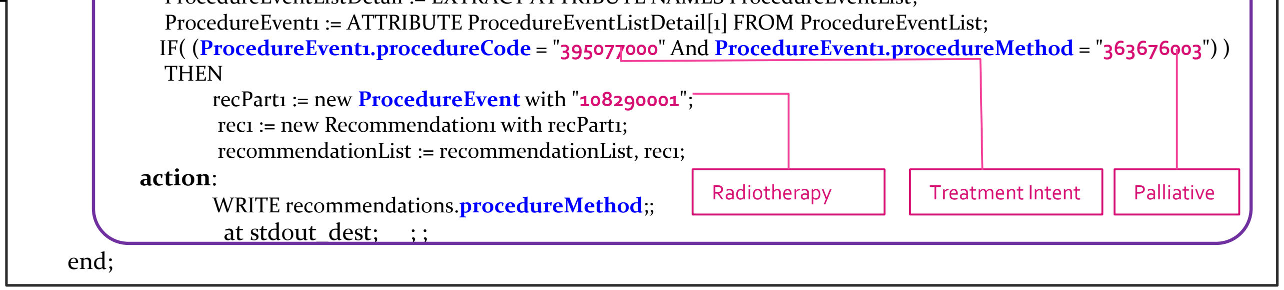


SurgicalHistory - ProcedureEvent

vMR Classes' Attributes	DCM concepts	SNOMED CT Description	SNOMED CT Code
approachBodySite	Vocal Cord	Vocal Cord Structure	46105003
	larynx	Laryngeal structure (body structure)	4596009
	Laryngeal part	Laryngeal part (body structure)	119196003
	Entire larynx	Entire larynx (body structure)	181212004
	Head and neck	Head and neck structure (body structure)	774007
	Muscle tissue	Muscle tissue (body structure)	91727004
	Pharynx	Pharyngeal structure (body structure)	54066008
	Skin	Skin structure (body structure)	39937001
	Lymph node	Structure of lymph node (body structure)	59441001
procedureCode	SurgicalHistory	Surgical procedure (procedure)	387713003
procedureMethod	Vocal cord stripping	Stripping of vocal cords (procedure)	9373007
	Laser surgery	Laser surgery (procedure)	35631009
	Cordectomy	Cordectomy of vocal cord (procedure)	47935000
	Laryngectomy	Laryngectomy (procedure)	72791001
	Partial laryngectomy	Partial laryngectomy (procedure)	80513001
	Total laryngectomy	Total laryngectomy (procedure)	66478002
	Myocutaneous flaps	Myocutaneous flap of head and neck (procedure)	65635008
	Pharyngectomy	Pharyngectomy (procedure)	51265002
	Free flaps	Reconstruction with free flap (procedure)	240985005
	Lymph node removal	Biopsy of lymph node (procedure)	21911005
	Tracheotomy/Tracheostomy	Incision of trachea (procedure)	48387007
targetBodySite	As above (approachBodySite)		

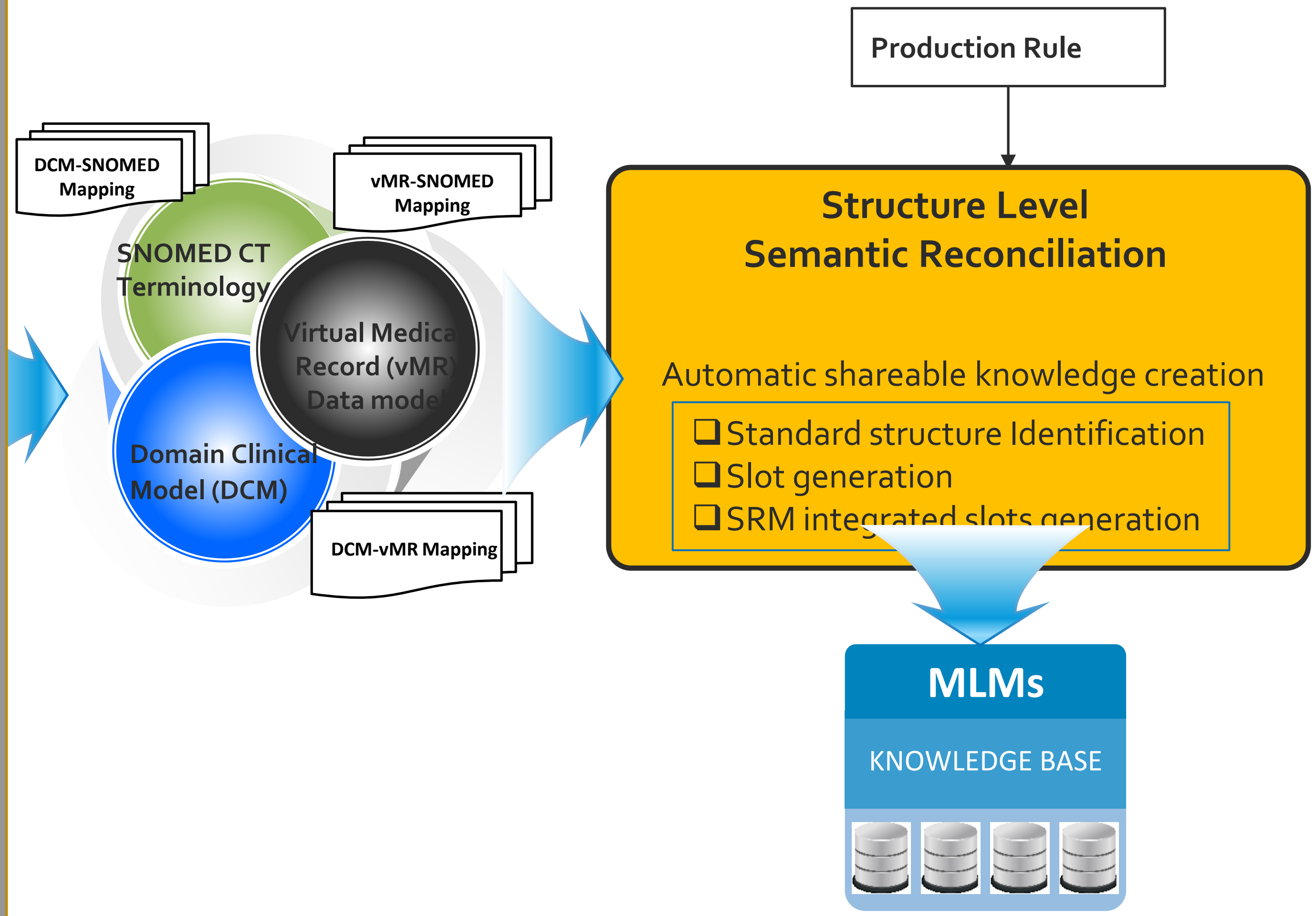
Medical Logic Module Generation





- **The Medical Logic Module (MLM) is generated from plane rules instead of ontologies.**
- **Standard Terminology (SNOMED CT) and standard data model (vMR) ontology concepts are using to create shareable and interoperable MLM.**

Solution 2: Structure Level Semantic Reconciliation (MLM Generation)



Solution 2: Structure Level Semantic Reconciliation (MLM Generation) [Motivation]

Standard Clinical knowledge representation Using HL7 Arden Syntax MLM

Structure

```
maintenance:
        title: Palliative Treatment By Physician 3;;
        mlmname: Palliative Treatment By Physician 3;;
        arden: Arden Syntax V2.7;;
        version: Version 2.7;;
        institution: SKMCH;;
        author: Dr. Physician 3;;
        specialist: Dr. Physician 3;;
        date: 13/01/2015;;
        validation: testing;;
Library:
 purpose: Experimental testing;;
 explanation: Experimental testing;;
 keywords: Oral Cavity;;
 citations: ;;
Knowledge:
    type: data driven; ;
    data:
        LET varTreatmentIntent = BE Read { Select TreatmentInten from ClientDB }
        ;;
evoke: null_event; ;
        logic:
            if ( varTreatmentIntent is equal to Palliative)
            {
               Conclude true;
            };;
        action:
                WRITE "The recommended treatment plan is Radiotherapy"
                at stdout_dest;    ; ;
   end;
```

Syntax

Solution

Complexity , Interoperability, and Integration

If Treatment Intent = Palliative Then Treatment Plan = Radiotherapy

SNOMED CT

vMR

```
maintenance:
        title: Palliative Treatment By Physician 3;;
        mlmname: Palliative Treatment By Physician 3;;
        arden: Arden Syntax V2.7;;
        version: Version 2.7;;
        institution: SKMCH;;
        author: Dr. Physician 3;;
        specialist: Dr. Physician 3;;
        date: 13/01/2015;;
        validation: testing;;
Library:
 purpose: Experimental testing;;
 explanation: Experimental testing;;
 keywords: Oral Cavity;;
 citations: ;;
Knowledge:
    type: data driven; ;
    data: ProcedureEvents := object [ProcedureEvent];
          ProcedureEventList := read as ProcedureEvents
          { select ProcedureEvent FROM client Where ProcedureEvent.procedureCode IN ("395077000")};
           Recommendation1 := object[ProcedureEvent];
           RecommendationList := (); ;;
evoke: null_event; ;
    logic:
       ProcedureEventListDetail := EXTRACT ATTRIBUTE NAMES ProcedureEventList;
       ProcedureEvent1 := ATTRIBUTE ProcedureEventListDetail[1] FROM ProcedureEventList;
      IF( (ProcedureEvent1.procedureCode = "395077000" And ProcedureEvent1.procedureMethod = "363676003") )
       THEN
            recPart1 := new ProcedureEvent with "108290001";
             rec1 := new Recommendation1 with recPart1;
             recommendationList := recommendationList, rec1;
     action:
            WRITE recommendations.procedureMethod;;
             at stdout_dest;    ; ;
   end;
```

Treatment Intent

Radiotherapy

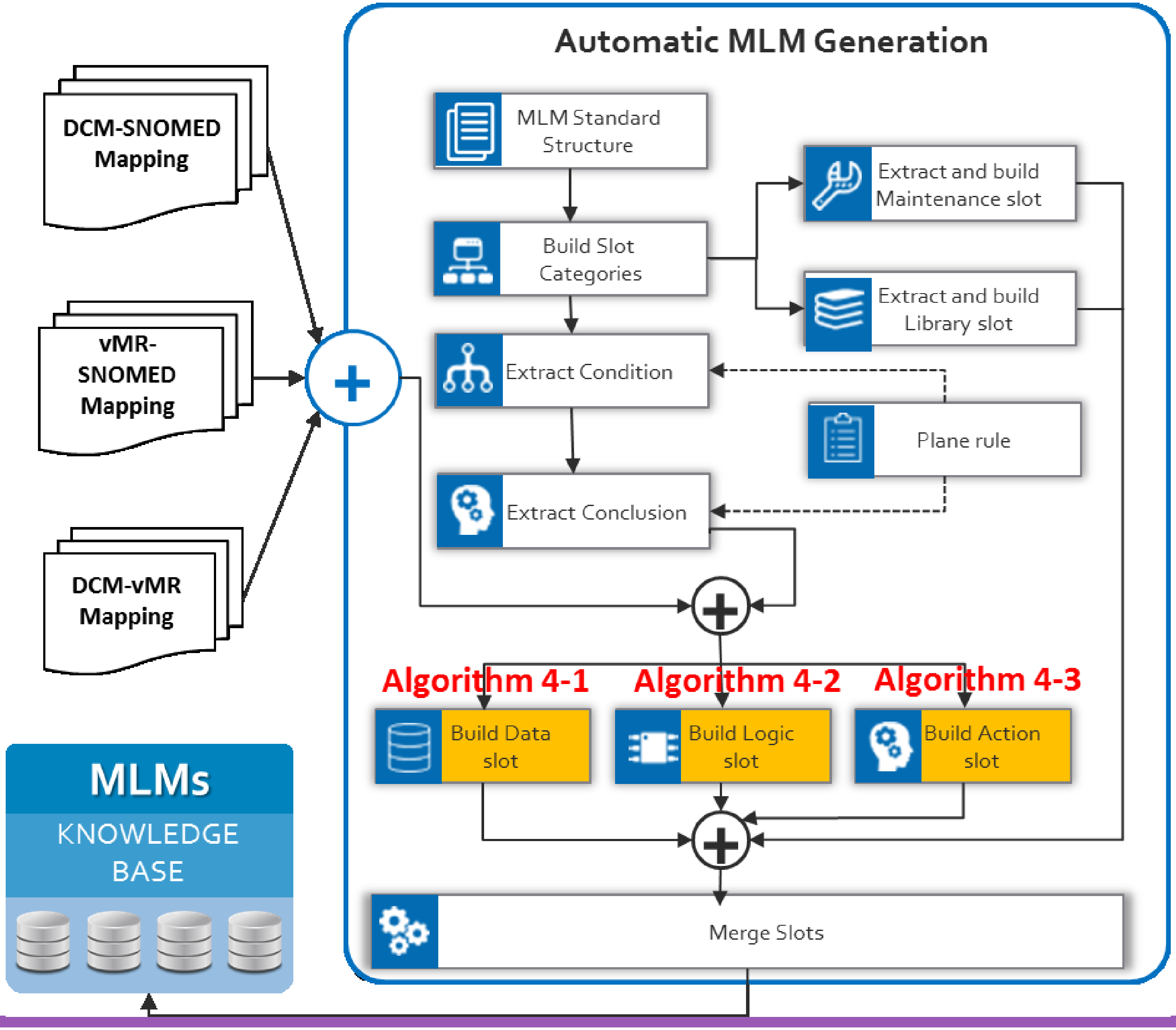
Treatment Intent

Palliative

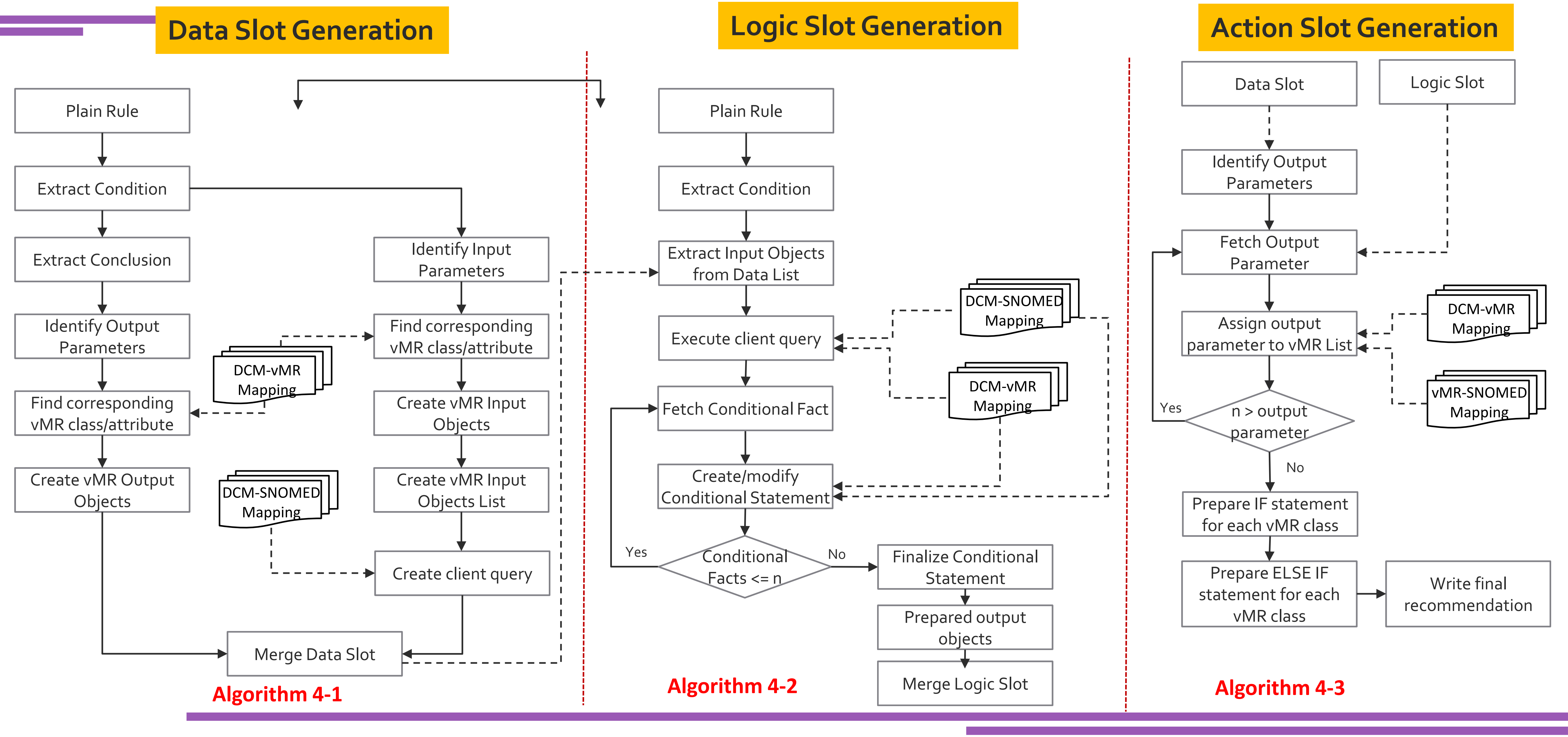
- **Solution: Structure level semantic reconciliation**
  - **Proposed solution hides this complexity by automatic generation of MLM.**

Solution 2: Structure Level Semantic Reconciliation (MLM Generation) [Workflow]

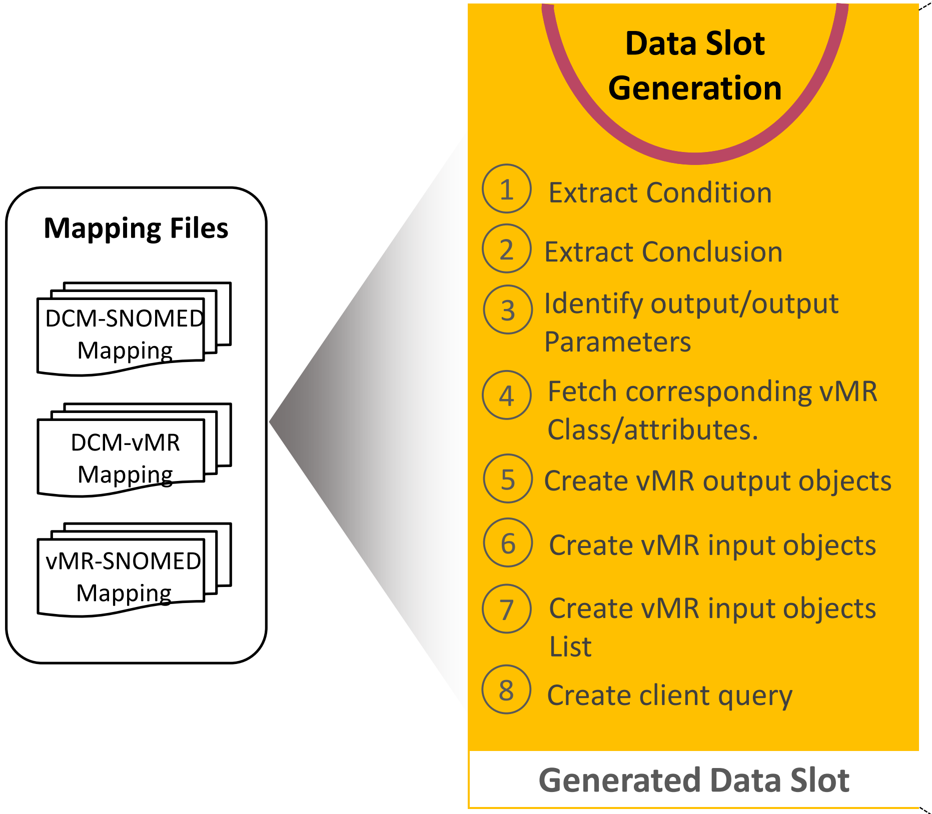
Contribution 3:
- **Automatic generation** of Medical Logic Module (MLM)
- Integration of SRM with MLM to enhance interoperability and easy integration with clinical workflows



Solution 2: Structure Level Semantic Reconciliation [Detail Workflow]



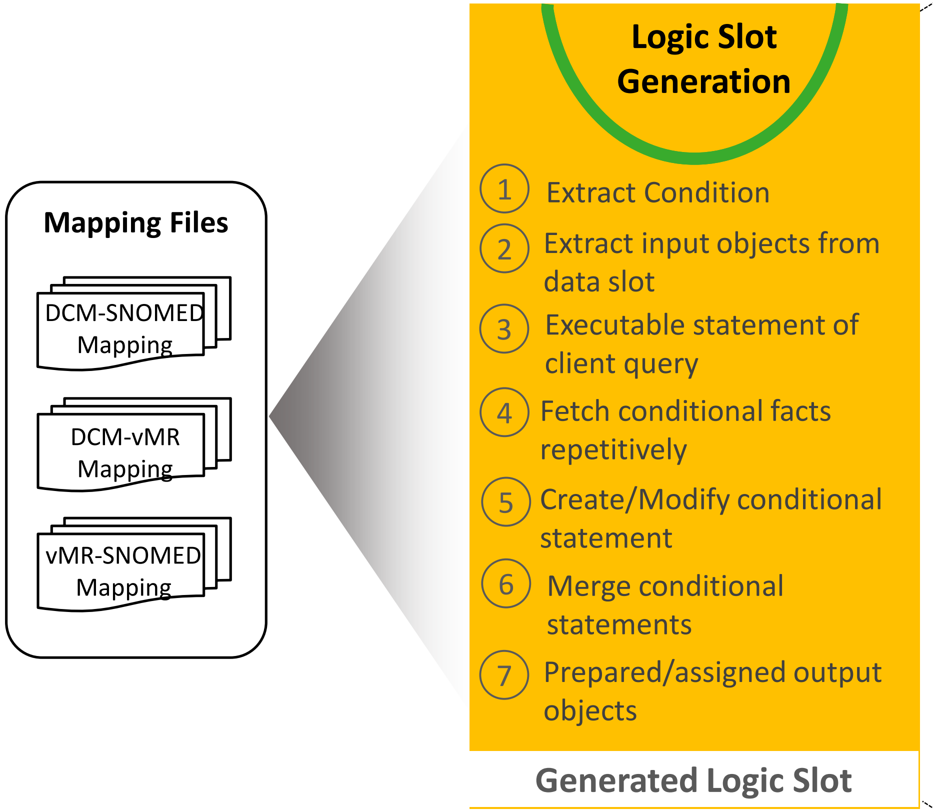
Algorithm 4-1: Structure Level Semantic Reconciliation (MLM Generation)



Algorithm 4-1: Semantic Reconciliation Model realization for Data Slot generation

```
Input  : Rule production rule
         mfDCMSNOMED mapping file of DCM-SNOMED
         mfvMRSNOMED mapping file of vMR-SNOMED
         mfDCMvMR mapping file of DCM-vMR
Result: DataSlot
Initialization:
inputObjectList = Empty;
outputObjectList = Empty;
listConditionalOperator[2] = {"and", "or"};
listConditionalFacts = Rule.condition.getAllFacts(listConditionalOperator);
listConclusionFacts = Rule.conclusion.getAllFacts(listConditionalOperator);
listComparisonOperator[] = {" = "," > "," < "," >= "," <= "};
foreach Fact f of listConditionalFacts do
    operands[2] = f.getOperands(listComparisonOperator);
    foreach ConceptC in mfDCMvMR do
        if C is mapped with operand[1] then
            vMRClassAttribute ← getCorrespondingvMRClass(C);
            if inputObjectList.contains(vMRClassAttribute = False) then
                inputObjectList.CreateObject(vMRClassAttribute.class);
                SNOMEDCode = mfDCMvMR.getSNOMEDCode(C);
                queryList.createQuery(vMRClassAttribute.class, SNOMEDCode)
                 Break;
            end
        end
    end
end
foreach Fact f of listConclusionFacts do
    operands[2] = f.getOperands(listComparisonOperator);
    foreach ConceptC in mfDCMvMR do
        if C is mapped with operand[1] then
            vMRClassAttribute ← getCorrespondingvMRClass(C);
            if outputObjectList.contains(vMRClassAttribute = False) then
                outputObjectList.CreateObject(vMRClassAttribute.class);
                Break;
            end
        end
    end
end
```

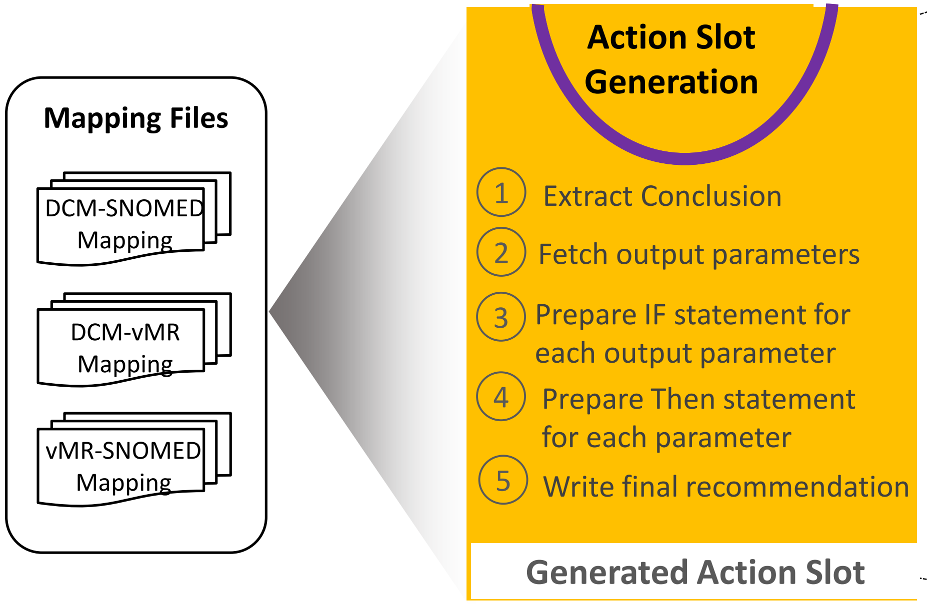
Algorithm 4-2: Structure Level Semantic Reconciliation (MLM Generation)



Algorithm 4-2: Semantic Reconciliation Model realization for Logic Slot generation

```
Input  : Rule production rule
         inputObjectsList Input object list generated by Data Slot
         outputObjectsList Output object list generated by Data Slot
         mfDCMSNOMED mapping file of DCM-SNOMED
         mfvMRSNOMED mapping file of vMR-SNOMED
         mfDCMvMR mapping file of DCM-vMR
Result: LogicSlot
Initialization:conditionStatement = "" and
conclusionStatement = "";
listConditionalOperator[2] = {"and", "or"};
listConditionalFacts = Rule.condition.getAllFacts(listConditionalOperator);
listConclusionFacts = Rule.conclusion.getAllFacts(listConditionalOperator);
listComparisonOperator[] = {" = "," > "," < "," >= "," <= "};
foreach Factf of listConditionalFacts do
    operands[2] = f.getOperands(listComparisonOperator);
    foreach ConceptC in operands do
        if inputObjectsList.find(ConceptC) = True then
            conditionStatement = conditionStatement +
             inputObjectsList[c].operands[0].getVMRClasses() + "." +
             inputObjectsList[c].getVMRAttribute();
            conditionStatement = conditionStatement +
             inputObjectsList[c].operands[0].getSNOMEDCTCode()
             conditionStatement = conditionStatement +
             operands[0].getFollowedOperator();
        end
    end
end
foreach Factf of listConclusionFacts do
    operands[2] = f.getOperands(listComparisonOperator);
    foreach ConceptC in operands do
        if outputObjectsList.find(ConceptC) = True then
            conclusionStatement = conclusionStatement +
             outputObjectsList[c].operands[0].getVMRClasses() + "." +
             outputObjectsList[c].getVMRAttribute();
            conclusionStatement = conclusionStatement +
             outputObjectsList[c].operands[0].getSNOMEDCTCode()
             conclusionStatement = conclusionStatement +
             operands[0].getFollowedOperator();
        end
    end
end
LogicSlot = "IF " + conditionStatement + " Then" + conclusionStatement;
```

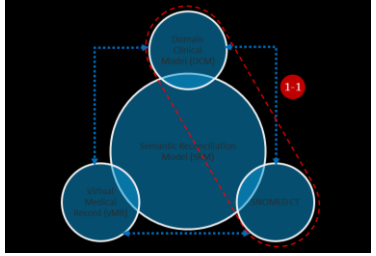
Algorithm 4-3: Structure Level Semantic Reconciliation (MLM Generation)



Algorithm 4-3: Semantic Reconciliation Model realization for Action Slot generation

```
Input  : Rule  production rule
         outputObjectsList  Output object list generated by Logic Slot
Result: ActionSlot
Initialization: recommendationStatement = "";
foreach Object obj of outputObjectsList do
    if obj.Is Not Null then
        recommendationStatement = "IF recommendation =" + obj.getVMRClass() +
          " then ";
        recommendationStatement = recommendationStatement + "Write " +
          obj.getVMRClass().getVMRAttribute() = obj.getValue;
        "Write " + recommendationStatement;
    end
end
if recommendationStatement is not NULL then
    ActionSlot = recommendationStatement;
end
```

Evaluation for Algorithm 1: DCM-Standard Terminology Mapping



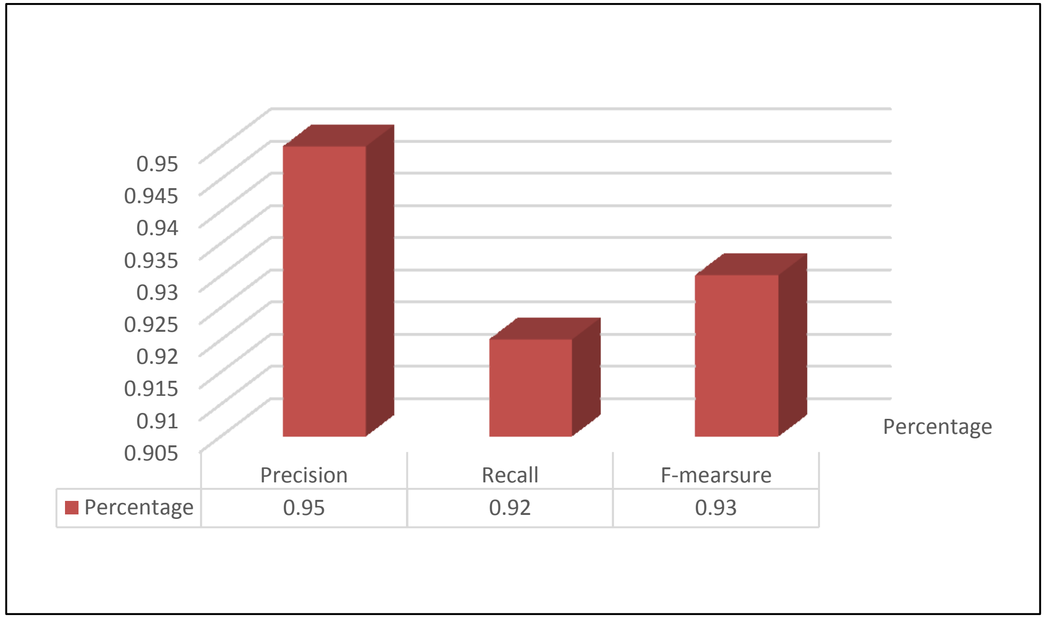
Experimental Setup

- **Experiments**
  - **Experiment 1:** Proposed Algorithm Performance for DCM-Standard Terminology Mapping
  - **Experiment 2:** Existing Algorithms Performance
    - AgreementMakerLight
    - GOMMA
    - LogMap Light
    - Proposed System
- **Datasets**
  - **DCM :** Domain Clinical Model (Collaborative Hospital)
    - Total number of concepts : 214
  - **SNOMED CT** : IHTSDO (Downloaded)
    - Total number of concepts : 0.3 million

$$\text{Precision} = \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Positive}}$$

$$\text{Recall} = \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Negative}}$$

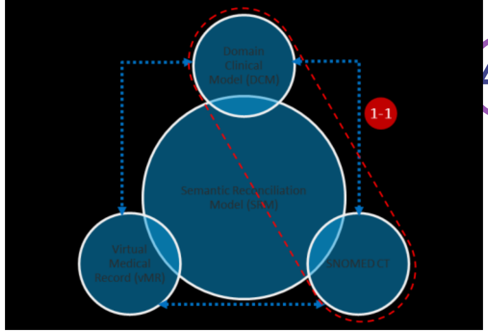
$$\text{F} - \text{measure} = 2 \times \frac{\textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$



Proposed System Results Statistics	
Total DCM Concepts	214
SNOMED CT Mapped Concepts	197
Wrong Mapped Concepts	9
Local Concepts	8

Taqdir Ali, and Sungyoung Lee, "Reconciliation of SNOMED CT and domain clinical model for interoperable medical knowledge creation", 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017), July 11-15, 2017

Evaluation for Algorithm 1: DCM-Standard Terminology Mapping



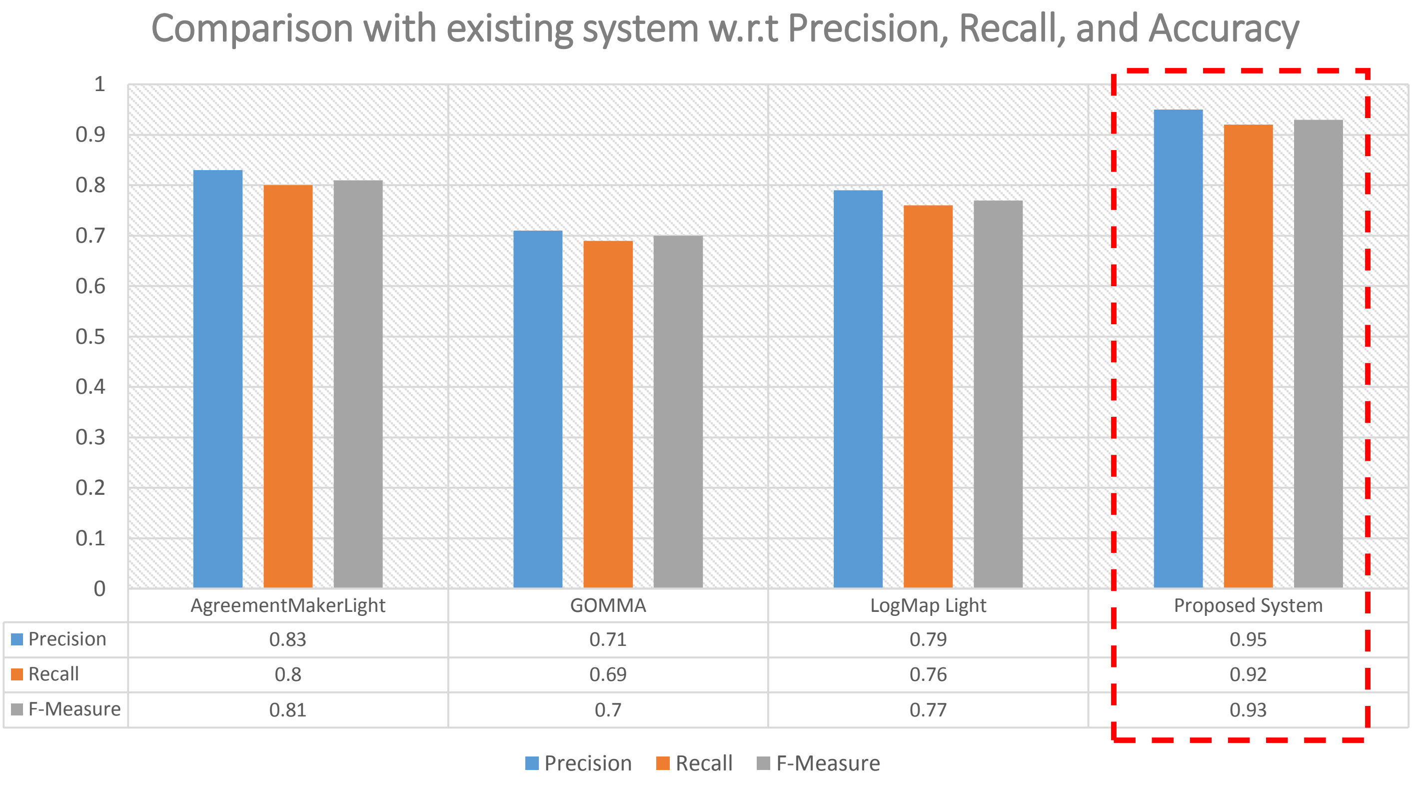
Relevant values of statistical measures finding

Systems	True Positive	False Positive	Local Concepts	False Negative
AgreementMakerLight	163	43	8	51
GOMMA	137	69	8	77
LogMap Light	153	53	8	61
Proposed System	197	9	8	17

Result Analysis

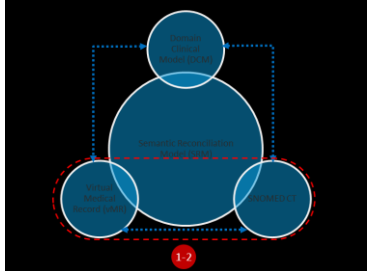
- SRM performed exceptionally well on the all features with 0.95 precision, 0.92 recall, and 0.93 F-measure on datasets
- The compared existing systems have good performance on both datasets as standard terminologies.
- Explicit semantics Inset enhanced performance

Statistical measures Precision, Recall, and F-measure



Taqdir Ali, and Sungyoung Lee, "Reconciliation of SNOMED CT and domain clinical model for interoperable medical knowledge creation", 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017), July 11-15, 2017

Evaluation for Algorithm 2: Standard Terminology and Data Model Mapping



Experimental Setup

- **Experiments**
  - **Experiment 1:** Proposed Algorithm Performance for Standard Terminology and data model mapping
  - **Experiment 2:** Existing ontology matching algorithms lack definition-based mapping, therefore, we evaluate our system with base-line (Jaccard) Similarity matching algorithm.
- **Datasets**
  - **Standard Data Model (vMR)** : Specification of standard data model with
    - Total number of concepts : 69
  - **SNOMED CT** : IHTSDO (Downloaded)
    - Total number of top level hierarchical concepts: 21

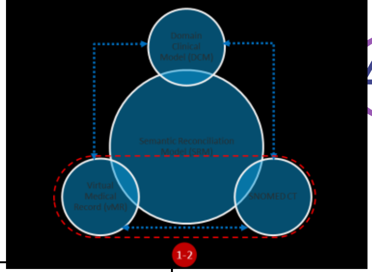
- **Tools and Technologies**
  - **Proposed Algorithm:** We implemented matching tool using java language.
    - Restful Service with Java
    - Knowledge base (SQL Server 2016)
    - Windows 10, RAM 4 GB
    - WordNet, ConceptNet5, AllAcronyms Repositories
  - **Jaccard Similarity Algorithm:** Existing base-line algorithm (Jaccard Similarity) is implement using existing tool RapidMinor
    - RapidMinor Studio Basic 6.5.002
    - Windows 10, RAM 4 GB
    - Python libraries for Preprocessing

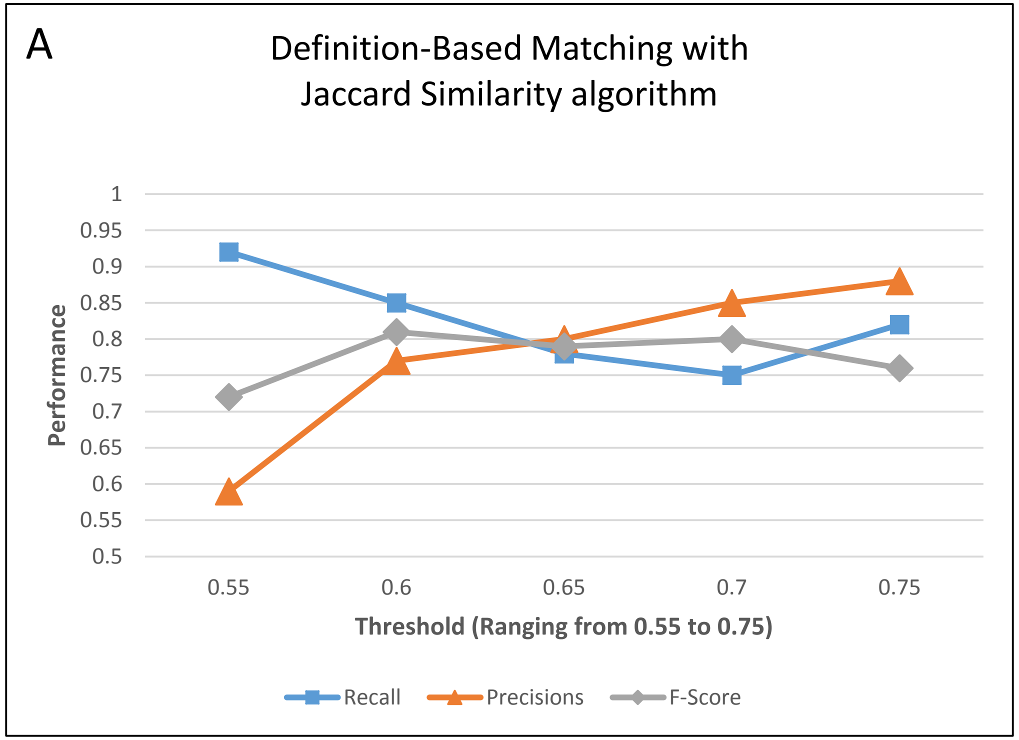
$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

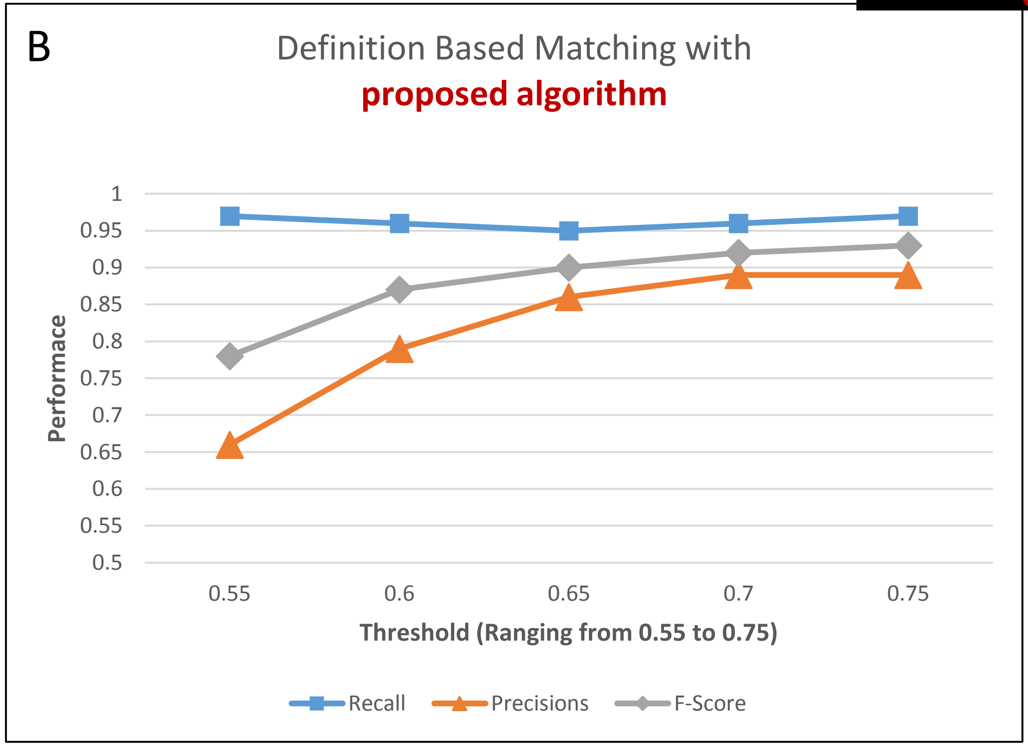
$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$\text{F} - \text{measure} = 2 \times \frac{Precision\ \times Recall}{Precision + Recall}$$

Evaluation for Algorithm 2: Standard Terminology and Data Model Mapping



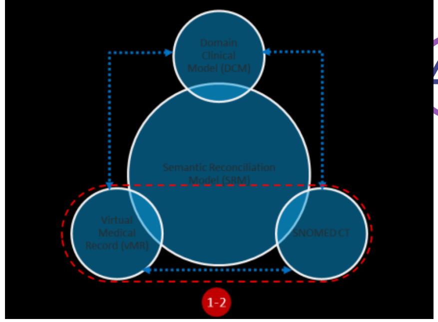




Results Analysis

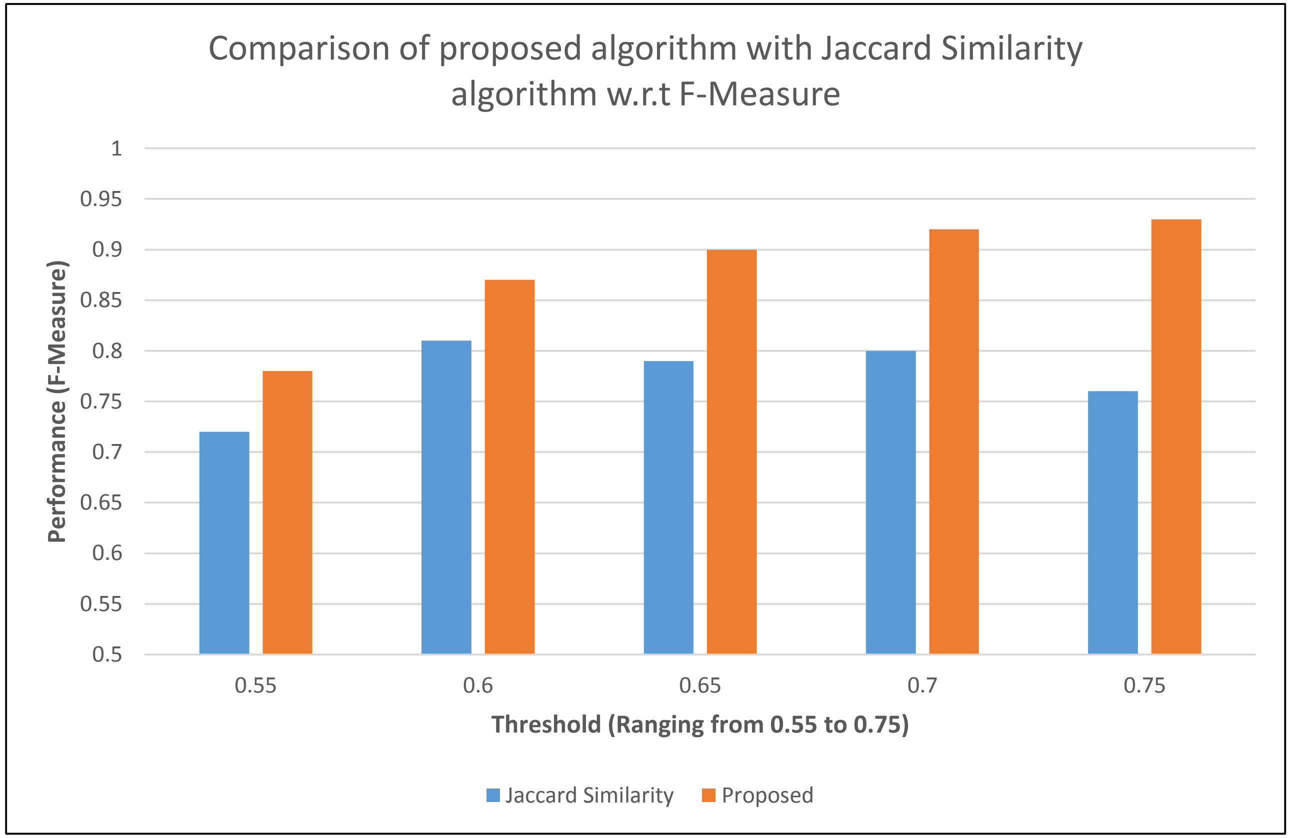
- A) The Jaccard similarity algorithm gives high recall (0.92) but the precision (0.59) was very low with threshold value 0.5. It gave very random results on different threshold values.
- B) The proposed system give better results than base-line (Jaccard) with each threshold. At threshold 0.75, the proposed algorithm gave good result with respect to Precision (0.89), Recall (0.97), and F-measure (0.93).
- We inserted Explicit and Implicit semantics into Space Vectors to increase the performance

Evaluation for Algorithm 2: Standard Terminology and Data Model Mapping



Results Analysis

- In summery, this graph shows the comparison of Proposed algorithm with Base-line algorithm with respect to F-Measure.
- The Proposed system showed better results F-Measure (0.93), while the Base-line algorithm has F-Measure (0.76) on the highest threshold value 0.75.
- On the lowest threshold value 0.55, the F-Measure of Proposed system is 0.78, while Base-line algorithm has 0.72 with very low precision.



Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

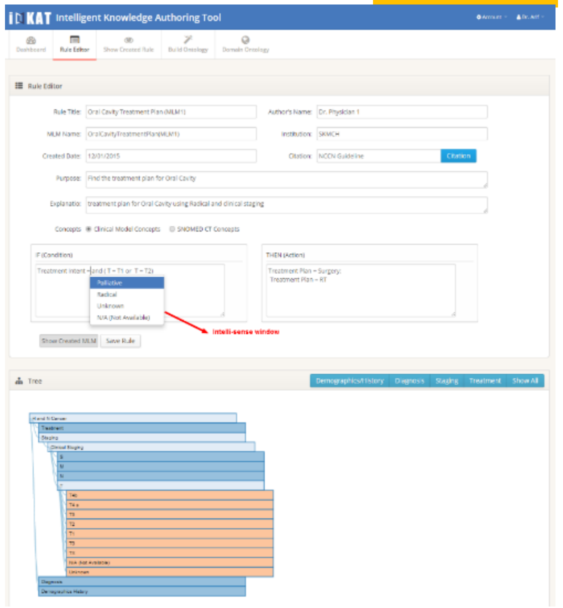
Evaluation for Solution 2 (Algorithm 4): Standard Terminology and Data Model Mapping

System-Centric Evaluation

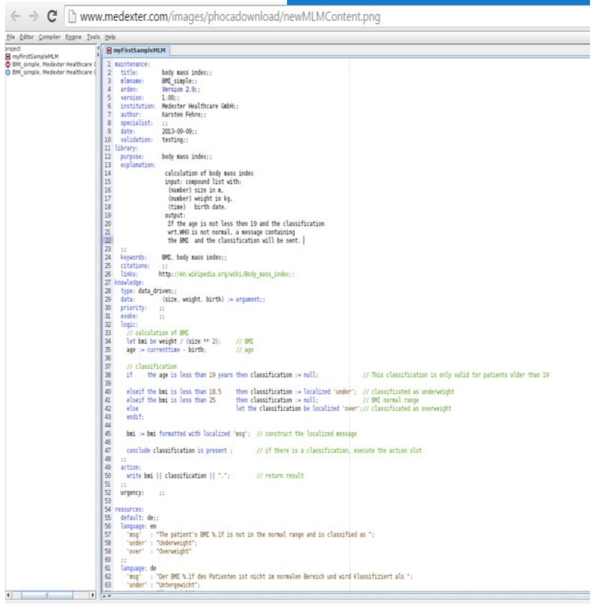
- System performance and quality measurement
- Dataset items collected from users during interaction with system [7]

COMPARISON TOOLS

I-KAT



ArdenSuite



User-Centric Evaluation

- System quality measurement
- Already existing dataset of preferences and opinions
- Items collected from users, not necessarily during interaction with system [7]

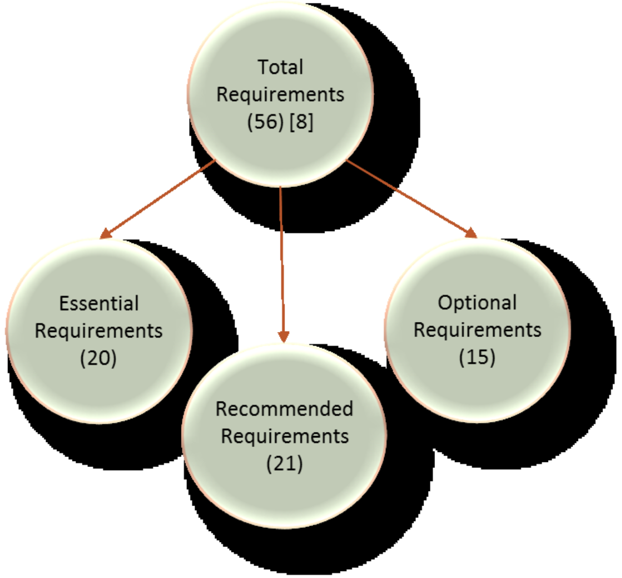
Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

Results of Solution 2 (Algorithm 4): System Centric

Initial Requirements	Phase 1 Reduction	Phase 2 Enhancement	Phase 3 Interpretation	Phase 4 Evaluation
	❑ Removed technology-oriented and specific to CMIT requirements	❑ Recommend new requirement of DCM tree and Intelli-sense window	❑ Interpret the CIMT related requirements for knowledge Authoring tool	❑ Compare and evaluate with existing system based on output of phase 3.

Requirements	Initial Requirements	Phase 1	Phase 2	Phase 3
Essential	**Total (20)** R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, R11, R12, R13, R14, R15, R16, R17, R18, R19, R20	**Total (16)** R1, R2, R3, R4, R5, R6, R8, R9, R10, R11, R13, R14, R16, R17, R18, R19	**Total (16)** R1, R2, R3, R4, R5, R6, R8, R9, R10, R11, R13, R14, R16, R17, R18, R19	**Total (16)** IR1, IR2, IR3, IR4, IR5, IR6, IR8, IR9, IR10, IR11, IR13, IR14, IR16, IR17, IR18, IR19
Recommended	**Total (21)** R23, R24, R25, R26, R27, R28, R29, R30, R31, R32, R33, R34, R35, R36, R37, R38, R39, R40, R41, R42, R43	**Total (16)** R23, R25, R26, R27, R29, R30, R31, R32, R33, R34, R37, R39, R40, R41, R42, R43	**Total (18)** R23, R25, R26, R27, R29, R30, R31, R32, R33, R34, R37, R39, R40, R41, R42, R43, ER57, ER58	**Total (18)** IR23, IR25, IR26, IR27, IR29, IR30, IR31, IR32, IR33, IR34, IR37, IR39, IR40, IR41, IR42, IR43, IR57, IR58
Optional	**Total (15)** R21, R22, R44, R45, R46, R47, R48, R49, R50, R51, R52, R53, R54, R55, R56	**Total (5)** R21, R45, R48, R51, R52	**Total (5)** R21, R45, R48, R51, R52	**Total (5)** IR21, IR45, IR48, IR51, IR52

Evaluations and Results
Evaluation of proposed system with exiting knowledge authoring tool [5] based on 40 requirements in the previous phase.



Final Requirement set for evaluation of knowledge Acquisition Tool

- Total (39)
  - Essential (16)
  - Recommended (18)
  - Optional (5)

Contribution 4

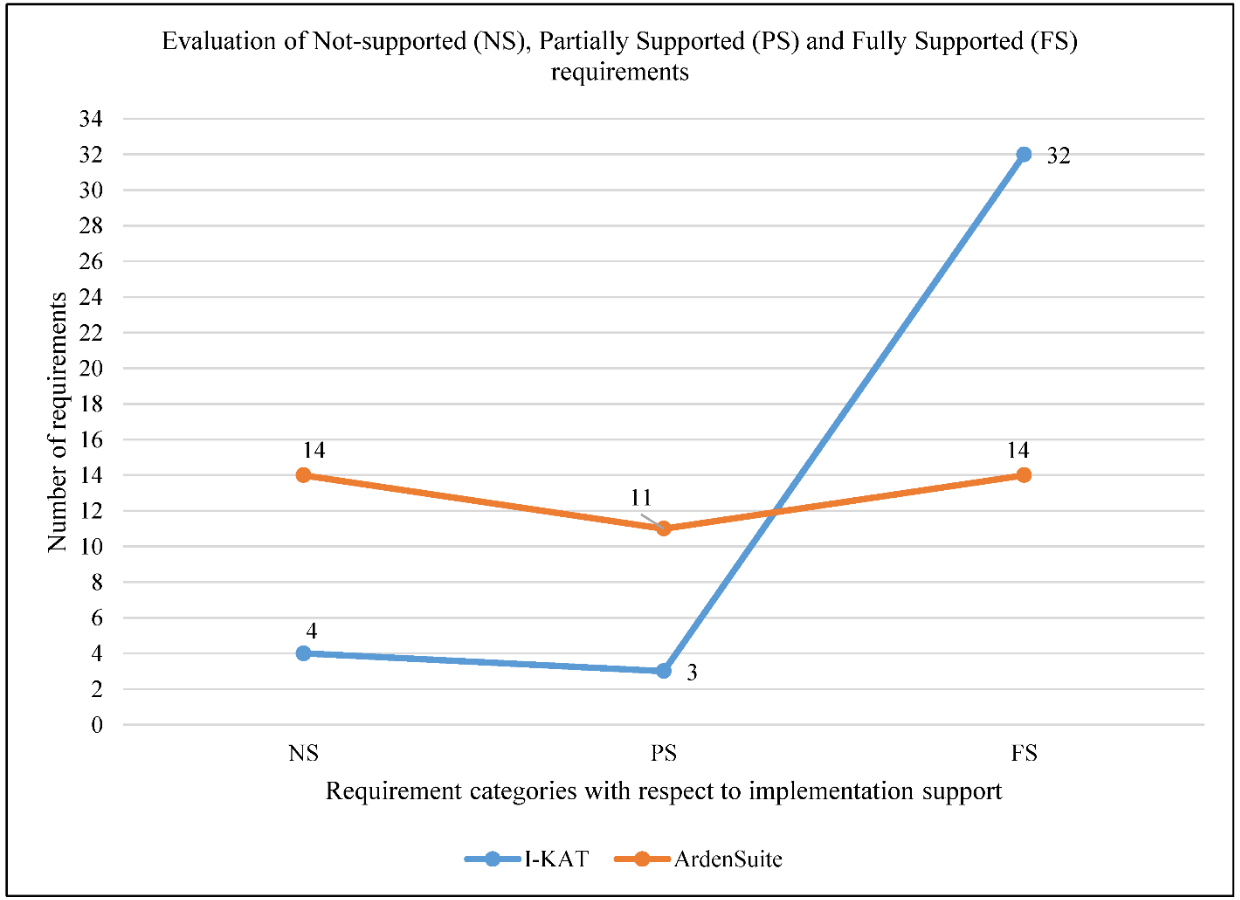
❑ The completeness evaluation of knowledge acquisition tool w.r.t. standard requirements.

Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

Results of Solution 2 (Algorithm 4): System Centric

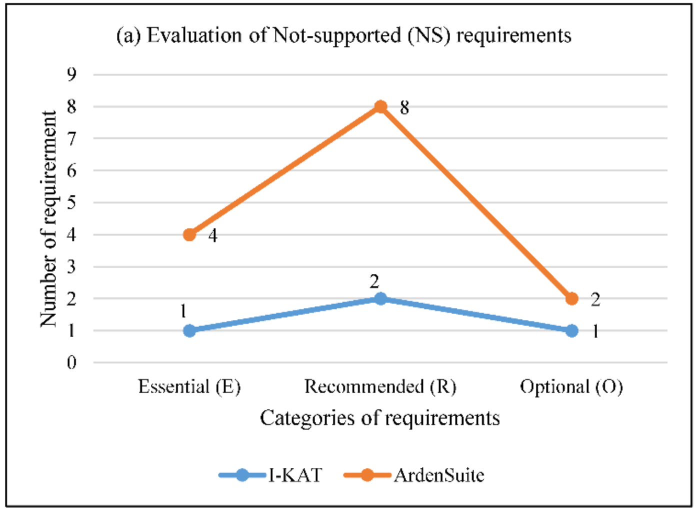
Result Analysis

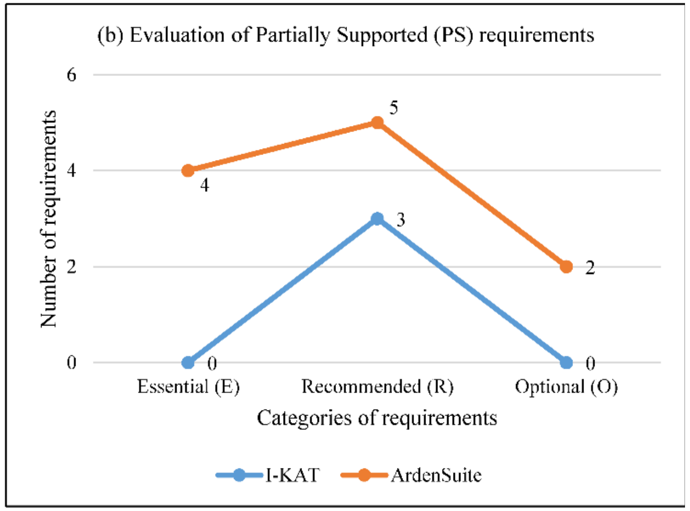
- I-KAT
  - has **full support** to 82.05% of requirements
  - Has **partial support** to 7.69% of requirements
  - Has **no support** to 10.25% of requirements
- ArdenSuite
  - has **full support** to **35.89%** of requirements
  - Has **partial support** to **28.20%** of requirements
  - Has **no support** to **35.89%** of requirements

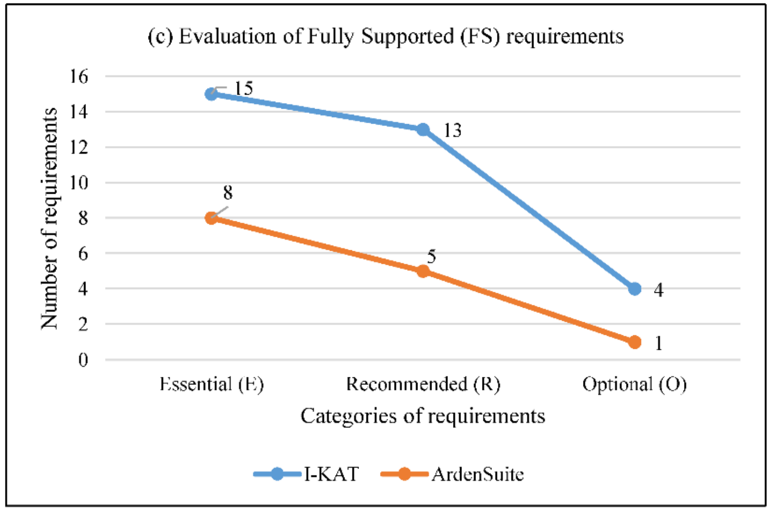


Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

Results of Solution 2 (Algorithm 4): System Centric







Interpretation:

a

I-KAT has minimum number of requirements which are not supported. Essential 1, Recommended 2, Optional 1

b

I-KAT has less partial support to Essential 0, Recommended 3, Optional 0

c

I-KAT has maximum full support to requirements. Essential 15, Recommended 13, Optional 4

Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

Results of Solution 2 (Algorithm 4): User Centric

Experiments and Objectives

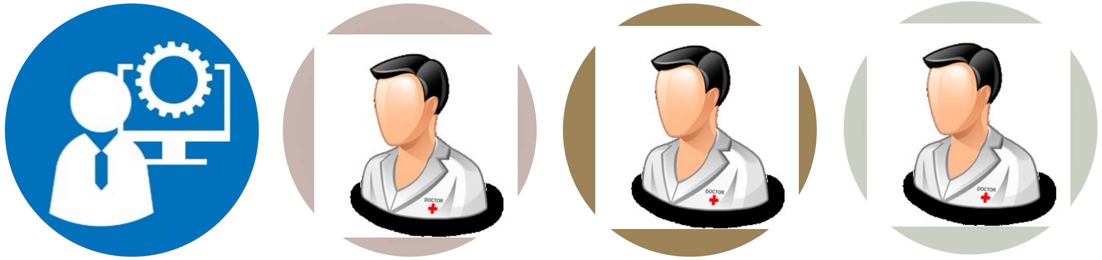
Experiments	Objectives
Experiment 1: User-friendliness	**Physicians' performance w.r.t. time**
Experiment 2: Shareability and Interoperability	**No. of errors in MLM creation**

Testing MLM: Total 8 MLMs are tested

- 2 MLM with simple structure
- 2 MLM with intermediate
- 4 MLM with complex structure

MLM ID	Associated Rule ID	Contents/Logic Complexity [No. of attributes, {No. of logical operators}]
MLM1	Rule1	[1, {And (0), Or (0)}]
MLM2	Rule2	[1, {And (0), Or (0)}]
MLM3	Rule3	[5, {And (3), Or (1)}]
MLM4	Rule4	[7, {And (5), Or (1) }]
MLM5	Rule5	[7, {And (5), Or (1)}]
MLM6	Rule6	[5, {And (3), Or (1)}]
MLM7	Rule7	[17, {And (7), Or (9)}]
MLM8	Rule8	[29, {And (7), Or (21)}]

Experimental Setup



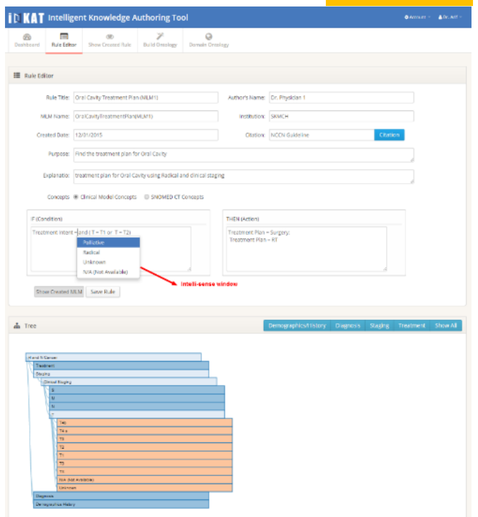
Experienced

Novice

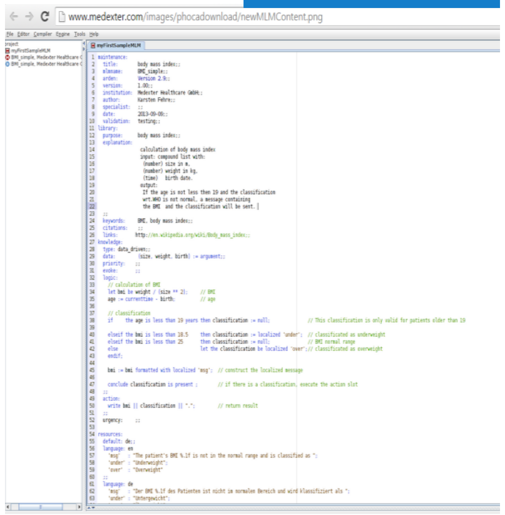
Intermediate

Experienced

I-KAT



ArdenSuite



Results of Solution 2 (Algorithm 4): User Centric

Experiment 1: User-friendliness

Physicians' performance w.r.t. Time

Result Analysis

- Create eight MLMs by each participant
- I-KAT has increased **34 times** faster than ArdenSuite in simplest MLM 1, while **5 times** faster in complex MLM 8.
- The overall average performance showed an improvement of **15 times**.
- We provided high level abstraction and hide the complexity.

Table 4: Ease-of-use evaluation with respect to time

MLM No	MLM Creation Time		User Involved
	Using ArdenSuite	Using I-KAT	
MLM1	18 min 20 Sec	22 sec	Physician 1
	21 min 15 Sec	46 sec	Physician 2
	Not Applicable	66 sec	Physician 3
	8 min 30 Sec	20 sec	Knowledge Engineer
MLM2	18 min 22 Sec	23 sec	Physician 1
	21 min 10 Sec	40 sec	Physician 2
	Not Applicable	69 sec	Physician 3
	8 min 34 Sec	23 sec	Knowledge Engineer
MLM3	32 min 20 Sec	2 min and 47 sec	Physician 1
	34 min 30 Sec	3 min and 5 sec	Physician 2
	Not Applicable	2 min and 40 sec	Physician 3
	19 min 15 Sec	2 min and 18 sec	Knowledge Engineer
MLM4	33 min 25 Sec	3 min and 49 sec	Physician 1
	35 min 39 Sec	3 min and 7 sec	Physician 2
	Not Applicable	3 min and 45 sec	Physician 3
	18 min 21 Sec	2 min and 19 sec	Knowledge Engineer
MLM5	33 min 25 Sec	3 min and 49 sec	Physician 1
	35 min 39 Sec	3 min and 7 sec	Physician 2
	Not Applicable	4 min and 47 sec	Physician 3
	21 min 21 Sec	3 min and 19 sec	Knowledge Engineer
MLM6	32 min 20 Sec	2 min and 47 sec	Physician 1
	34 min 30 Sec	3 min and 5 sec	Physician 2
	Not Applicable	2 min and 40 sec	Physician 3
	19 min 15 Sec	2 min and 18 sec	Knowledge Engineer
MLM7	34 min 45 Sec	4 min and 53 sec	Physician 1
	36 min 51 Sec	5 min and 51 sec	Physician 2
	Not Applicable	8 min and 27 sec	Physician 3
	21 min 34 Sec	4 min and 10 sec	Knowledge Engineer
MLM8	35 min 58 Sec	5 min and 23 sec	Physician 1
	37 min 51 Sec	6 min and 19 sec	Physician 2
	Not Applicable	9 min and 47 sec	Physician 3
	22 min 46 Sec	5 min and 13 sec	Knowledge Engineer

Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

Results of Solution 2 (Algorithm 4): User Centric

Experiment 2: Shareability and Interoperability → **No. of errors in MLM creation**

Result Analysis

- **Syntactic** (S) and **Logical** (L) Errors
- Participants made on **4 errors** on average in simplest MLM1 and **17 errors** in complex MLM 8 using ArdenSuite, its success rate is **46.88%** .
- Participant made only logical error on average **1 error** in all MLMs using I-KAT with overall Task Success Rate **90.625%** .
- We hide the structure and syntax of MLM from physicians

- **Syntactic** (S) Errors:
  - Mistakes in syntax of Arden Syntax Language of MLM.
- **Logical** (L) Errors:
  - Mistakes of wrong concept selection, wrong operator selection, and wrong values selection.

Table 5: MLM syntactic and semantic evaluation

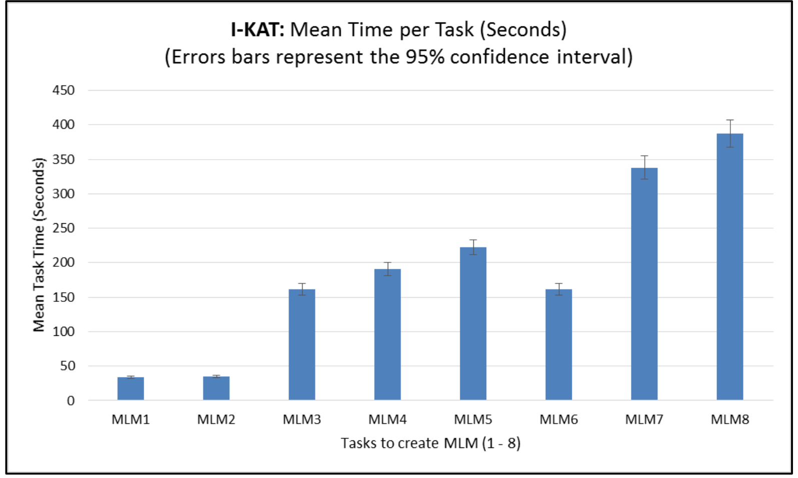
MLM No	MLM Errors Recorded (L: Logical errors, S: Syntax error) Using ArdenSuite	Using I-KAT	User Involved
MLM1	S:2, L:2	S:0, L:0	Physician 1
	S:3, L:5	S:0, L:0	Physician 2
	Not Applicable	S:0, L:0	Physician 3
	S:0, L:0	S:0, L:0	Knowledge Engineer
MLM2	S:2, L:1	S:0, L:0	Physician 1
	S:3, L:4	S:0, L:0	Physician 2
	Not Applicable	S:0, L:1	Physician 3
	S:0, L:0	S:0, L:0	Knowledge Engineer
MLM3	S:10, L:11	S:0, L:1	Physician 1
	S:5, L:18	S:0, L:1	Physician 2
	Not Applicable	S:0, L:2	Physician 3
	S:2, L:0	S:0, L:0	Knowledge Engineer
MLM4	S:9, L:13	S:0, L:1	Physician 1
	S:6, L:17	S:0, L:2	Physician 2
	Not Applicable	S:0, L:2	Physician 3
	S:3, L:0	S:0, L:1	Knowledge Engineer
MLM5	S:7, L:12	S:0, L:1	Physician 1
	S:6, L:16	S:0, L:1	Physician 2
	Not Applicable	S:0, L:3	Physician 3
	S:3, L:1	S:0, L:0	Knowledge Engineer
MLM6	S:8, L:9	S:0, L:1	Physician 1
	S:6, L:17	S:0, L:2	Physician 2
	Not Applicable	S:0, L:1	Physician 3
	S:1, L:0	S:0, L:0	Knowledge Engineer
MLM7	S:8, L:13	S:0, L:1	Physician 1
	S:8, L:17	S:0, L:2	Physician 2
	Not Applicable	S:0, L:4	Physician 3
	S:3, L:2	S:0, L:0	Knowledge Engineer
MLM8	S:9, L:15	S:0, L:2	Physician 1
	S:6, L:16	S:0, L:1	Physician 2
	Not Applicable	S:0, L:3	Physician 3
	S:2, L:3	S:0, L:0	Knowledge Engineer

Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72

Evaluation: Task-On-Time

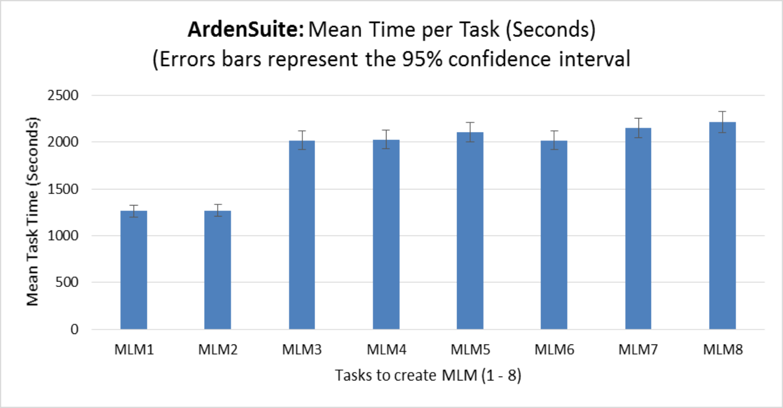
Time-On-Task:

- Called task completion time.
- Faster completion of task by user is the better experience.
- Mean time per MLM creation calculated in seconds.
- Find Mean, Median, Geometric mean, and confidence interval.
- The error bars represent the 95% (α = 0.05) confidence interval.
- **ArdenSuite:** The average time-on-task in simplest MLM is **1621.25 Sec**, and complex MLM is **2348.75 Sec**.
- **I-KAT:** The average time-on-task in simplest MLM is **38.5 Sec**, and complex MLM is **400.5 Sec**.



I-KAT: Time-On-Task data for 4 participants and 8 MLMs

	MLM1	MLM2	MLM3	MLM4	MLM5	MLM6	MLM7	MLM8
Physician 1	22	23	167	229	229	167	293	323
Physician 2	46	40	185	187	187	185	351	379
Physician 3	66	69	160	225	287	160	507	587
Knowledge Engineer	20	23	138	139	199	138	250	313
Mean	38.5	38.75	162.5	195	225.5	162.5	350.25	400.5
Median	34	31.5	163.5	206	214	163.5	322	351
Geometric mean	34	34.76	161.61	191.3	222.38	161.61	337.89	387.26
Standard Deviation	21.81	21.7	19.43	41.86	44.64	19.43	112.4	127.68
Confidence	21.37	21.27	19.04	41.02	43.75	19.04	110.15	125.12
Confidence Interval (+)	59.87	60.02	181.54	236.02	269.25	181.54	460.4	525.62
Confidence Interval (-)	17.13	17.48	143.46	153.98	181.75	143.46	240.1	275.38
Confidence Interval	42.74	42.54	38.08	82.04	87.5	38.08	220.3	250.24



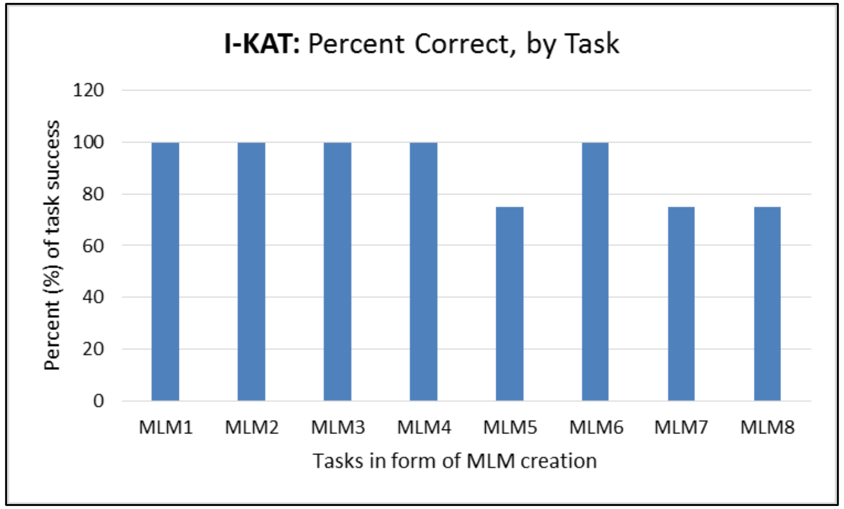
ArdenSuite: Time-On-Task data for 4 participants and 8 MLMs

	MLM1	MLM2	MLM3	MLM4	MLM5	MLM6	MLM7	MLM8
Physician 1	1100	1102	1940	2005	2005	1940	2085	2158
Physician 2	1275	1270	2070	2139	2139	2070	2211	2271
Physician 3	3600	3600	3600	3600	3600	3600	3600	3600
Knowledge Engineer	510	514	1155	1101	1281	1155	1294	1366
Mean	1621.25	1621.5	2191.25	2211.25	2256.25	2191.25	2297.5	2348.75
Median	1187.5	1186	2005	2072	2072	2005	2148	2214.5
Geometric mean	1266.76	1268.57	2021.46	2030.5	2108.84	2021.46	2152.7	2215.67
Standard Deviation	1359.16	1358.24	1022.45	1034.25	971.88	1022.45	958.5	926.26
Confidence	1331.95	1331.05	1001.98	1013.55	952.42	1001.98	939.31	907.72
Confidence Interval (+)	2953.2	2952.55	3193.23	3224.8	3208.67	3193.23	3236.81	3256.47
Confidence Interval (-)	289.3	290.45	1189.27	1197.7	1303.83	1189.27	1358.19	1441.03
Confidence Interval	2663.9	2662.1	2003.96	2027.1	1904.84	2003.96	1878.62	1815.44

Evaluation: Task Success Rate

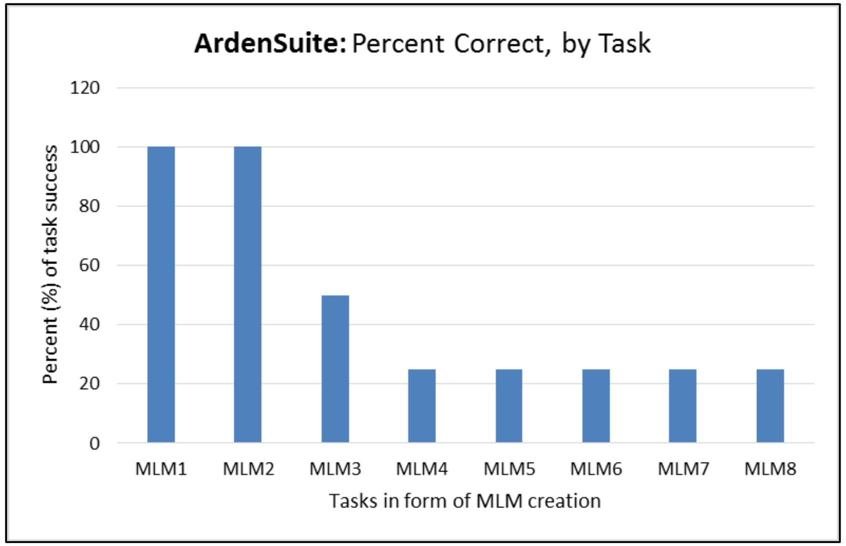
Task Success Rate:

- Binary success is the simplest way to measure task success.
- It depends on Pass Or Fail binary decision.
- Typically, 1's are using for success, and 0's are using for failure.
- Calculate task success by looking at the average success rate for each task across the participants Or each participant across task.
- The value at the right-bottom of the table is overall task success of the system
- ArdenSuite has **46.88%** task success rate, while proposed I-KAT has **90.62%** task success.



I-KAT: Task Success for 4 participants and 8 MLMs

	MLM1	MLM2	MLM3	MLM4	MLM5	MLM6	MLM7	MLM8	Average
Physician 1	1	1	1	1	1	1	1	1	**100%**
Physician 2	1	1	1	1	1	1	1	1	**100%**
Physician 3	1	1	1	1	0	1	0	0	**62.5%**
Knowledge Engineer	1	1	1	1	1	1	1	1	**100%**
Average	**100%**	**100%**	**100%**	**100%**	**75%**	**100%**	**75%**	**75%**	**90.62%**



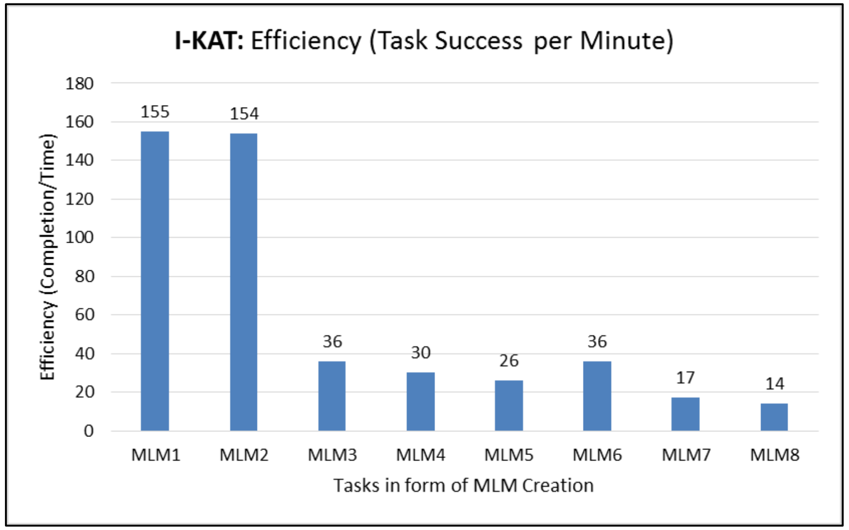
ArdenSuite: Task Success for 4 participants and 8 MLMs

	MLM1	MLM2	MLM3	MLM4	MLM5	MLM6	MLM7	MLM8	Average
Physician 1	1	1	1	0	0	0	0	0	**37.5%**
Physician 2	1	1	0	0	0	0	0	0	**25%**
Physician 3	1	1	0	0	0	0	0	0	**25%**
Knowledge Engineer	1	1	1	1	1	1	1	1	**100%**
Average	**100%**	**100%**	**50%**	**25%**	**25%**	**25%**	**25%**	**25%**	**46.87%**

Evaluation: Efficiency

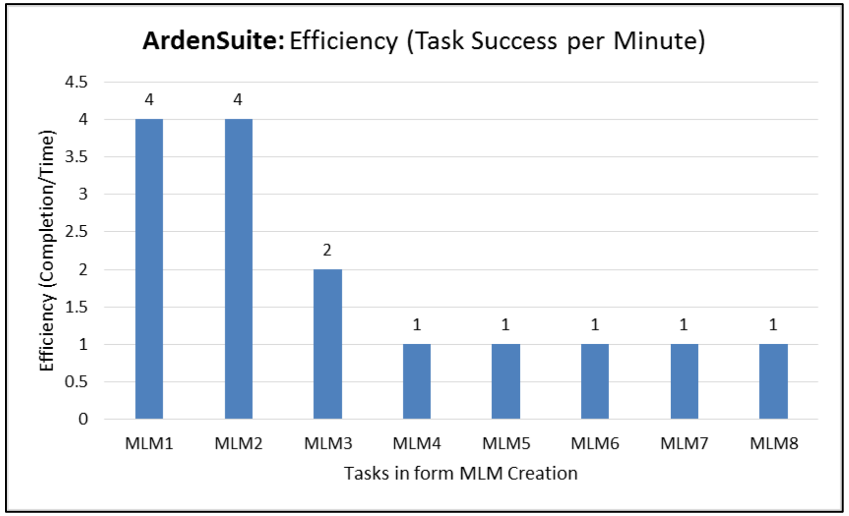
Efficiency:

- Efficiency is the combination of two metrics "Task Success" and "Time-On-Task".
- Usability Test Reports **(ISO/IEC 25062:2006)** specifies that the "core measure of efficiency" is the ratio of the task completion rate to the mean time per task.
- We already calculated "Time-On-Task" and "Task Success", previously.
- The calculated efficiency of existing system is **1.87% completion rate/time**.
- The calculated efficiency of our proposed system is **56.62% completion rate/time**.



I-KAT: Efficiency Measurement

	Task Success Rate	Time-On-Task (sec)	Time-On-Task(min)	Efficiency %
MLM1	100	38.5	0.64	**155**
MLM2	100	38.75	0.64	**154**
MLM3	100	162.5	2.70	**36**
MLM4	100	195	3.25	**30**
MLM5	75	225.5	3.75	**19**
MLM6	100	162.5	2.70	**36**
MLM7	75	350.25	5.83	**12**
MLM8	75	400.5	6.67	**11**
Efficiency (I-KAT) =				**56.62**



ArdenSuite: Efficiency Measurement

	Task Success Rate	Time-On-Task (sec)	Time-On-Task(min)	Efficiency %
MLM1	100	1621.25	27.02	**4**
MLM2	100	1621.5	27.02	**4**
MLM3	50	2191.25	36.52	**2**
MLM4	25	2211.25	36.85	**1**
MLM5	25	2256.25	37.60	**1**
MLM6	25	2191.25	36.52	**1**
MLM7	25	2297.5	38.29	**1**
MLM8	25	2348.75	39.14	**1**
Efficiency (ArdenSuite) =				**1.87**

Uniqueness and Contributions

Research Contribution

Knowledge Interoperability

- Semantic Reconciliation Model (SRM) to enhance the knowledge Interoperability
  - Explicit Semantic embedding to enhance ontology matching performance
    - Precision-0.95, Recall-0.92, F-Measure-0.93
  - Enhance data and schema level interoperability by definition based mappings
    - Precision-0.89, Recall-0.97, F-Measure-0.93

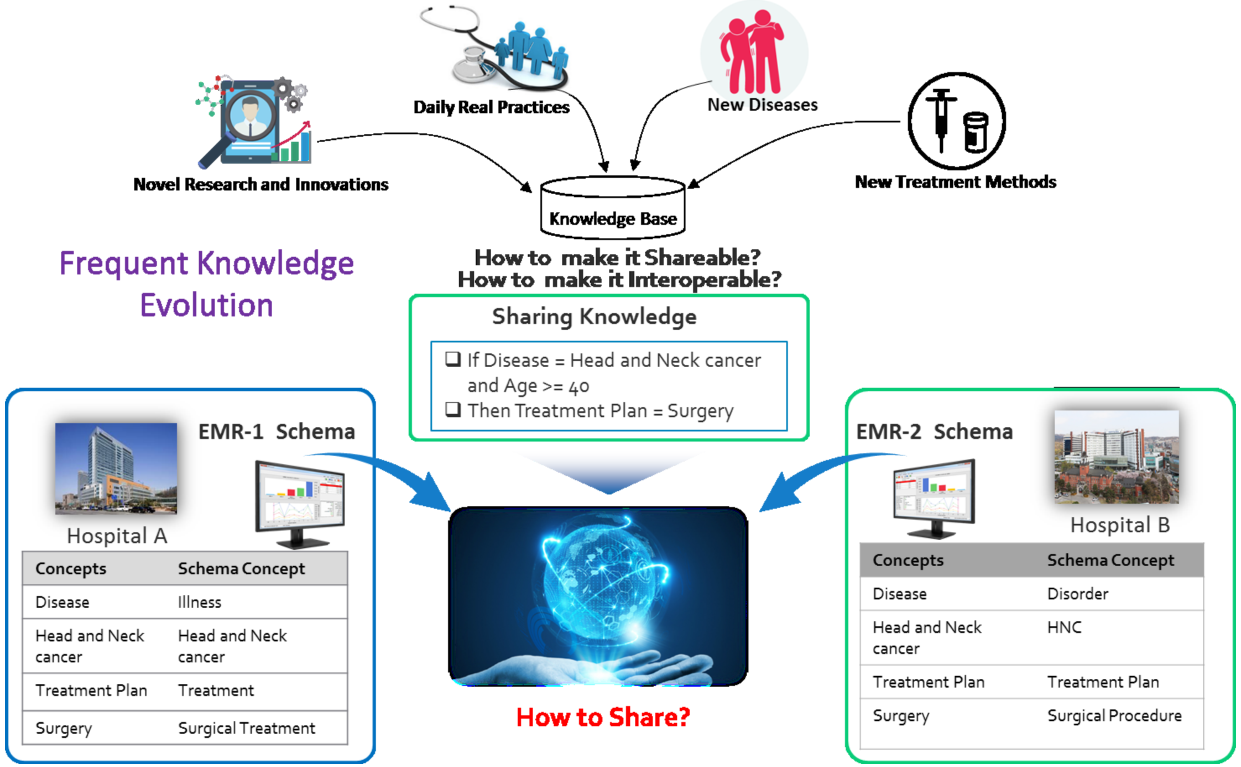
Knowledge Shareability

- Automatic MLM generation with the help of structure interoperability using medical standards
  - Task Success Rate **90.625%**

System evaluation Method

- We introduced new system evaluation methodology based on standard requirements to fulfill during design and development of knowledge acquisition system

Medical Domain Contribution



- **Clinical Knowledge shareability and interoperability is achieved**
  - **Intelligent Knowledge Authoring Tool (I-KAT) to create shareable and interoperable knowledge**

Uniqueness and Contributions

Knowledge Interoperability

- Semantic Reconciliation Model to enhance the knowledge shareability and Interoperability
- Explicit Semantic embedding to enhance ontology matching performance
- Enhance data and schema level interoperability by definition based mappings

Knowledge Shareability

- Automatic MLM generation with the help of structure interoperability using medical standards

System evaluation Method

- We introduced new system evaluation methodology based on standard requirements to fulfill during design and development of knowledge acquisition system

Conclusions and Future Works

Semantic Reconciliation Model (SRM) is proposed

- To make the standard and non standard knowledge base as interoperable.
- To acquire shareable knowledge using standard knowledge representation.
- Transform the knowledge into executable format.

Standards Data Models, Terminologies, and Knowledge Representation

- In SRM, we used Standard Data Model vMR and Standard Terminology (SNOMED CT) to achieve Interoperability.
- We used standard knowledge representation HL7 MLM for shareability and Production rules for non-standard knowledge interoperability.

Future Works

- In future, we will validate the interoperability aspect with other knowledge representations such as GLIF
- Uncertainty control will handle in Medical Logic Module using Fuzzy Logic

Selected References

- [1] A. Soumeya, et al. "A UMLS based knowledge acquisition tool for rule-based clinical decision support system development." Journal of the American Medical Informatics Association , vol 8, no. 4, pp. 351-360, 2012.
- [2] H. Nathan C, et al. " KAT: A Flexible XML-based Knowledge Authoring Environment", Journal of the American Medical Informatics Association, vol 12, no. 4, pp. 418-430, 2005.
- [3] J. Robert A, et al. “Challenges in Implementing a Knowledge Editor for the Arden Syntax: Knowledge Base Maintenance and Standardization of Database Linkages”, Proceedings of the AMIA Symposium. American Medical Informatics Association, pp. 355–359, 2002.
- [4] D. Dustin, et al. "A Knowledge Authoring Tool for Clinical Decision Support" Journal of Clinical Monitoring and Computing, vol 22, no. 3, pp. 189-198,2008.
- [5] http://www.medexter.com/ (Last visited on 03 April, 2018)
- [6] R. Rachel, et al. “A Clinical Rule Editor in an Electronic Medical Record setting: Development, Design, and Implementation”, Proceedings of the AMIA Annual Symposium. American Medical Informatics Association, pp. 537–541, 2009.
- [7] Cremonesi, Paolo, Franca Garzotto, and Roberto Turrin. "User-centric vs. system-centric evaluation of recommender systems." Human-Computer Interaction–INTERACT 2013. Springer Berlin Heidelberg, 2013. 334-351.
- [8] A. Moreno-Conde, F. Jodar-Sanchez, D. Kalra, Requirements for clinical information modelling tools, International journal of medical informatics 84 (7) (2015) 524-536
- [9] D. F. Sittig, A. Wright, J. A. Oshero , B. Middleton, J. M. Teich, J. S. Ash, E. Campbell, D. W. Bates, Grand challenges in clinical decision support, Journal of biomedical informatics 41 (2) (2008) 387-392
- [10] R. A. Jenders, B. Dasgupta, Challenges in implementing a knowledge editor for the arden syntax: knowledge base maintenance and standardization of database linkages., in: Proceedings of the AMIA Symposium, American Medical Informatics Association, 2002, p. 355
- [11] Berner, Eta S. "Clinical decision support systems: state of the art." AHRQ publication 90069 (2009).
- [12] M. Samwald, K. Fehre, J. De Bruin, K.-P. Adlassnig, The arden syntax standard for clinical decision support: Experiences and directions, Journal of biomedical informatics 45 (4) (2012) 711-718

Selected References

- [13] Kirsten, T., Gross, A., Hartung, M., & Rahm, E. GOMMA: a component-based infrastructure for managing and analyzing life science ontologies and their evolution. J.Biomedical Semantics, 2011; 2, 6
- [14] Hartung, Michael, Lars Kolb, Anika Groß, and Erhard Rahm. "Optimizing Similarity Computations for Ontology Matching-Experiences from GOMMA." InData Integration in the Life Sciences, pp. 81-89. Springer Berlin Heidelberg, 2013
- [15] Jiménez-Ruiz, E., & Grau, B. C. Logmap: Logic-based and scalable ontology matching. In The Semantic Web–ISWC, 2011; (pp. 273-288). Springer Berlin Heidelberg.
- [16] Ernesto Jiménez-Ruiz, Bernardo Cuenca Grau, and Valerie V. Cross. "LogMap family participation in the OAEI 2016." In Proceedings of the 11th International Workshop on Ontology Matching (OM-2016) Colocated with the 15th International Semantic Web Conference (ISWC-2016) (CEUR-WS) , Vol. 1766. Kobe, Japan, 185–189
- [17] Faria, D., Pesquita, C., Santos, E., Palmonari, M., Cruz, I. F., & Couto, F. M. The agreementmakerlight ontology matching system. In On the Move to Meaningful Internet Systems: OTM 2013 Conferences, 2013; (pp. 527-541). Springer Berlin Heidelberg
- [18] Ratnesh Sahay, Ronan Fox, Antoine Zimmermann, Axel Polleres, and Manfred Hauswrith. A methodological approach for ontologising and aligning health level seven (hl7) applications. Availability, Reliability and Security for Business, Enterprise and Health Information Systems, pages 102-117, 2011
- [19] F. Jauro, S. Junaidu, and S. Abdullahi. Falcon-ao++: An improved ontology alignment system. International Journal of Computer Applications, vol. 94, no. 2, 2014
- [20] Hu, Wei, Jianfeng Chen, Gong Cheng, and Yuzhong Qu. "ObjectCoref & Falcon-AO: results for OAEI 2010." In Proceedings of the Fifth International Workshop on Ontology Matching, pp. 158-165. 2010
- [21] Jenders, R. A., and B. Dasgupta. "Assessment of a knowledge-acquisition tool for writing Medical Logic Modules in the Arden Syntax." Proceedings of the AMIA Annual Fall Symposium. American Medical Informatics Association, 1996.
- [22] Ali, Taqdir, et al. "Multi-model-based interactive authoring environment for creating shareable medical knowledge." Computer methods and programs in biomedicine 150 (2017): 41-72
- [23] Taqdir Ali, and Sungyoung Lee, "Reconciliation of SNOMED CT and domain clinical model for interoperable medical knowledge creation", 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2017), Jeju island, Korea, July 11-15, 2017

Problem Statement References

[1] A. Wright, D. F. Sittig, J. S. Ash, S. Sharma, J. E. Pang, and B. Middleton, “Clinical decision support capabilities of commercially-available clinical information systems,” Journal of the American Medical Informatics Association, vol. 16, no. 5, pp. 637–644, 2009
[2] Hamouda, Iman Ben, et al. "Understanding Knowledge Sharing in Health Care System." MCIS. 2015.
[3] CHOURABI, Olfa. "Understanding the barriers to Knowledge sharing in the French healthcare system: a conceptual model and research propositions." (2018).
[4] Sottilare, Robert, et al., eds. Design recommendations for intelligent tutoring systems: authoring tools and expert modeling techniques. Robert Sottilare, 2015.
[5] O'Donnell, Eileen, et al. "A review of personalised e-learning: Towards supporting learner diversity." International Journal of Distance Education Technologies (IJDET) 13.1 (2015): 22-47.
[6] Rezaei, Reza, et al. "Interoperability evaluation models: A systematic review." Computers in Industry 65.1 (2014): 1-23.
[7] Gyrard, Amelie, Christian Bonnet, and Karima Boudaoud. "Domain knowledge Interoperability to build the semantic web of things." W3C Workshop on the Web of Things. 2014.

Publications

- **SCI/ SCIE Journals (12)**
  - First Author- TWO Published
  - Co-Author- 10 Published
- **Local Journal (1)**
  - Co-Author
- **International Conferences (19)**
  - First Author - Seven Publications
  - Co-Author- Twelve Publications
- **Local Conferences (3)**
  - First Author – Three Publications
- **Patents Local (3)**
  - 2 Registered
  - 1 Applied
- **International Patent (1)**
  - 1 Applied

Total Publications = 39

Achievements

Research Achievement Award
Korean Society of Heart Failure – 2018

Computer Methods and Programs in Biomedicine 150 (2017) A1–A2

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Editorial

Two new computational methods for data analysis: A social network analysis-based classifier and the GEEORD SAS module

Two of this month's editor's choice articles involve new computational methods that provide additional options to applied data scientists and statisticians. The third article is an applied research report that addresses the important problem of increasing the efficiency and easy of creating clinical knowledge repositories.

The first editor's choice article, "Social Network Analysis-based Classifier (SNAc): A case study on time course gene expression data," repurposes techniques used in social network analysis to create a classifier for time sequential genomic data [1]. The authors are able to model genomic data in terms of a social network by grouping genes to create 'communities' according to commonly shared characteristics. They calculate distances between genes and connect them using adjacency matrices. Their final model, called SNAc, is an ensemble method which ultimately results from calculating genomic expression differences between nodes. The authors compare their social network-based classifier against six other mainstream machine learning algorithms and find that it outperforms many of them and is not inferior to them. These results indicate that social network analysis may be another lens with which to view time sequential datasets, and the authors close by mentioning that they can extend their classifier so that it becomes a general classifier.

The second editor's choice article is titled, "GEEORD: A SAS macro for analyzing ordinal response variables with repeated measures through proportional odds, partial proportional odds, or non-proportional odds models." [2] Many clinical datasets fit this use case, including before/after studies in which patients are graded on semi-quantitative scaled outcomes. Commonly, proportional odds among subjects is not always the case, and the included SAS modules, PROC Logistic and PROC GENMOD, lack formal tests for proportional odds, which many decrease the robustness of results. Created by a group of statisticians, the free GEEORD macro fills a needed gap in the SAS statistical software suite by providing a formal proportional odds test and adds the capability to fit a non-proportional odds model, a partial proportional odds model, and a proportional odds model. The authors illustrate their macro by using it on real-world datasets.

The final editor's choice article, "Multi-model-based Interactive Authoring Environment for Creating Shareable Medical Knowledge," details one research group's effort to speed up the generation of machine-interpretable decision support rules through the design of an easy-to-use, interactive knowledge authoring tool for non-technical users [3]. The authors designed the Intelligent Knowledge Authoring Tool (I-KAT), essentially a user-friendly way for non-technical physicians to transfer their clinical knowledge into the Arden syntax to generate computer-readable knowledge artifacts. The authors go beyond merely writing software to testing the easy of use compared to the current de facto standard software. They measure the ability of their systems' features to meet all modeling requirements and show that their I-KAT software covered more than twice as many requirements as the de facto standard software and was at least fifteen times more user-friendly. These findings reinforce the lesson that much room for delivering value lies not just in inventing new fundamental innovations (Arden syntax in this case), but making these innovations accessible to the critical audience (in this case physicians) has just as much to do with the realizing the end goal (efficient creation of knowledge repositories).

Guest Editors

Richard Lu
Graduate Institute of Biomedical Informatics, College of Medicine Science and Technology, Taipei Medical University, Taipei, Taiwan; International Center for Health Information Technology (ICHIT), Taipei Medical University, Taipei, Taiwan;

Usman Iqbal
International Center for Health Information Technology (ICHIT), Taipei Medical University, Taipei, Taiwan; Master Program in Global Health and Development, College of Public Health, Taipei Medical University, Taipei, Taiwan; Health Informatics Unit, COMSATS Institute of Information Technology (CIIT), Islamabad, Pakistan

Yu-Chuan (Jack) Li*
Graduate Institute of Biomedical Informatics, College of Medicine Science and Technology, Taipei Medical University, Taipei, Taiwan; International Center for Health Information Technology (ICHIT), Taipei Medical University, Taipei, Taiwan; Chair, Dept. of Dermatology, Wan Fang Hospital, Taipei, Taiwan

*Corresponding author. Address: 250- Wuxing Street, Xinyi District, Taipei 11031, Taiwan
E-mail addresses: jack@tmu.edu.tw, jaak88@gmail.com (Y.-C. (Jack) Li)

Editor's Best Choice of Volume Award
Computer Methods and Programs in Biomedicine

THANK YOU!

Any questions or comments?

Appendix

Summary of Limitations of Existing Works

Knowledge Interoperability

1 Semantic Mapping

Lack of Semantic Mappings, such as explicit and implicit semantics.

2 Low Accuracy

Accuracy for specific size of ontologies - Low accuracy in standard and non-standard ontologies

3 Definition-based Mapping

Lack of definition based mappings. Standard data model and terminology can only map with definitions

Lack of definition-based matching in ontology mapping

Knowledge Shareability

1 Semantic Mapping

Lack of Semantic Mappings, such as explicit and implicit semantics.

2 Low Accuracy

Accuracy for specific size of ontologies - Low accuracy in standard and non-standard ontologies

3 Definition-based Mapping

Lack of definition based mappings. Standard data model and terminology can only map with definitions

Lack of definition-based matching in ontology mapping

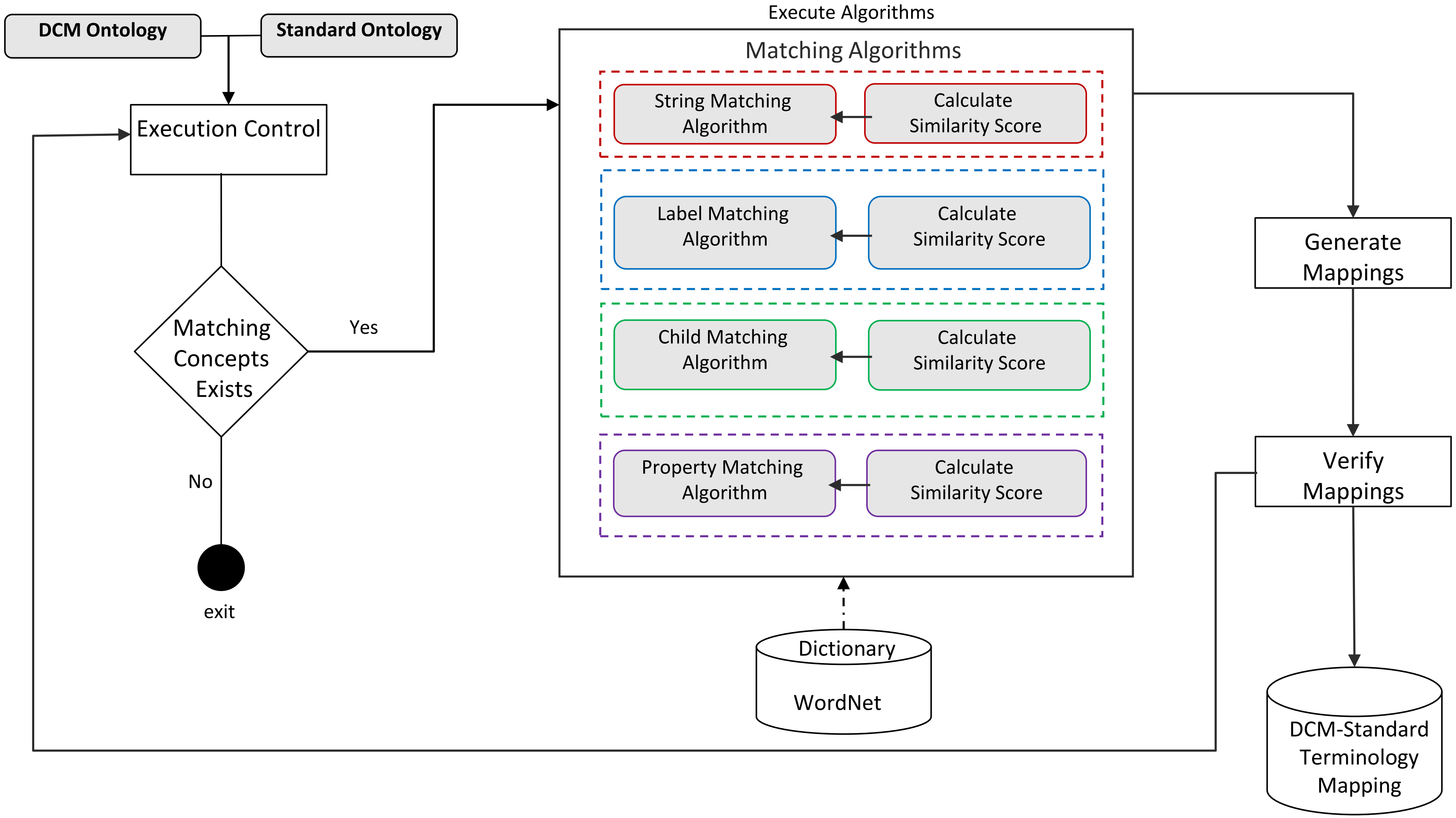
Proposed Methodology

Limitations	Elements	Solutions
Lack of Semantics Interoperability	HL7 Standard vMR, SNOMED CT , and Domain Clinical Model (DCM)	Schema-Data level semantic reconciliation
Shareable knowledge authoring complexity	Domain Clinical Model (DCM) and Medical Logic Module (MLM)	Structure level semantic reconciliation

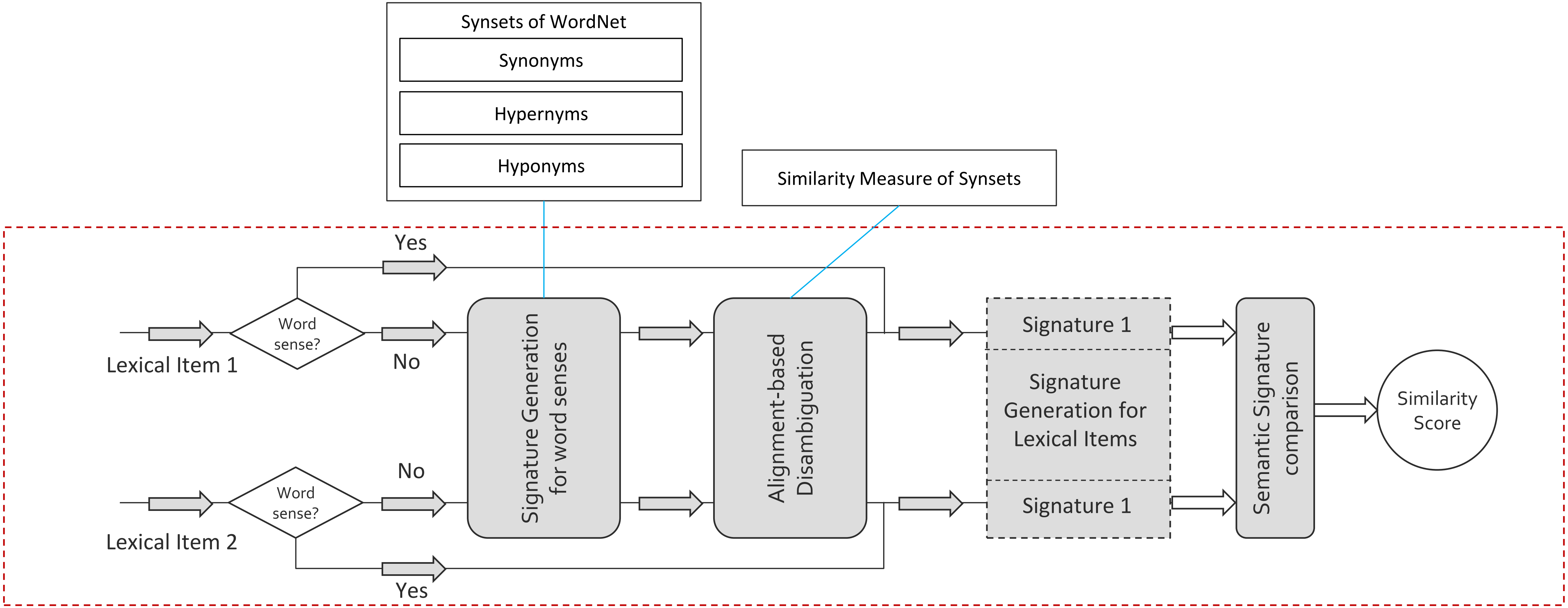
Solution 1: a) DCM-Standard terminology mapping algorithm (As – is)

Existing Approach Limitations

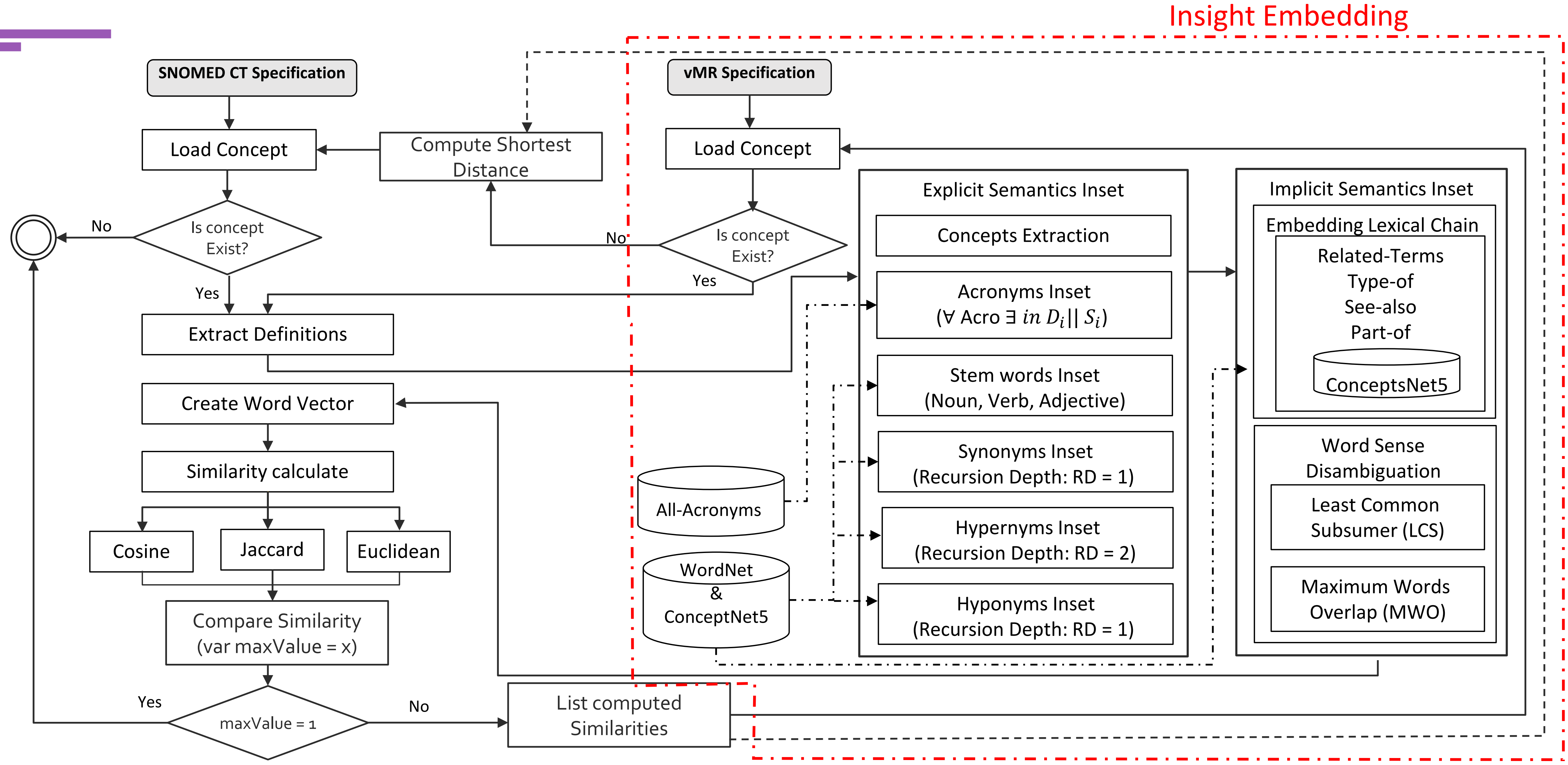
- ❑ Only focus on the internal semantics of ontology.
- ❑ Lack of definition-based mappings
- ❑ These limitations cause to the low accuracy.



Solution 1: b) Standard terminology and data model mapping algorithm



Solution 1: b) Standard terminology and data model mapping algorithm



Solution 1: b) Standard terminology and data model mapping algorithm

Algorithm 1: Standard terminology and data model mapping algorithm

Begin
inputs: $SNOMED\ CT\ Specification, vMR\ Specification$
$A = \{S_1 S_2, \ldots, S_n\}$ // n objects
$B = \{V_1 V_2, \ldots, V_n\}$ // n objects
output: $Mapping\ File$

foreach **Sc in A**
definitionSC = ExtractDefinition(Sc)
InsertExplicitSemantics(definitionSC)
InsertImplicitSemantics(definitionSC)
W^{Sc} **= CreateWordVector(definitionSC) ….** $[\mathbf{0} \quad \mathbf{1} \quad \mathbf{0}]$
foreach **Vc in B**
definitionVc = ExtractDefinition(Vc)
W^{Vc} **= CreateWordVector(definitionVc) ….** $[\mathbf{0} \quad \mathbf{1} \quad \mathbf{0}]$
CosineSimilarity$= \sum_{i=1}^{n} W^{Sc}{}_i . WVc\ {}_i / \sqrt{\sum_{i=1}^{n} W^{Sc}{}_i} . \sqrt{\sum_{i=1}^{n} W^{vc}{}_i}$
⋮
JaccardSimilarity $= \frac{M11}{M01+M10+M11}$
maxValue = CompareSimilarity(CosineSimilarity, ….,JaccardSimilarity)
if maxValue is one Then
Sc is matched with Vc
else
ListComputedSimiliraty = maxValue
end foreach
shortestDistance = ComputeShortestDistance(ListComputedSimiliraty)
if **shortestDistance >= 0.5**
Sc is matched with Vc
end if
end foreach
End

Solution 2: Structure level semantic reconciliation

- **Structure:**
  - It has a standard structure having following categories.
    - Maintenance
    - Library
    - Knowledge
  - Each category contain different slots
  - In knowledge category three core slots are
    - Data
    - Logic
    - Action
- **Syntax**
  - MLM has its own syntax to write meta data and actual rule based on structure, like columns, semi-columns, brackets, curly braces, and key words.

Standard Clinical knowledge representation Using HL7 Arden Syntax MLM

```
maintenance:
        title: Palliative Treatment By Physician 3;;
        mlmname: Palliative Treatment By Physician 3;;
        arden: Arden Syntax V2.7;;
        version: Version 2.7;;
        institution: SKMCH;;
        author: Dr. Physician 3;;
        specialist: Dr. Physician 3;;
        date: 13/01/2015;;
        validation: testing;;
Library:
 purpose: Experimental testing;;
 explanation: Experimental testing;;
 keywords: Oral Cavity;;
 citations:   ;;
Knowledge:
      type: data driven; ;
      data:
        LET varTreatmentIntent = BE Read { Select TreatmentInten from ClientDB }
        ;;
evoke: null_event; ;
        logic:
            if ( varTreatmentIntent is equal to Palliative)
            {
              Conclude true;
            };;
        action:
                WRITE "The recommended treatment plan is Radiotherapy"
                 at stdout_dest;      ; ;
   end;
```

- **Complexity:**
  - The standard structure and syntax of MLM is tedious and difficult task for physicians to remember
    - Structure
    - Syntax
- **Interoperability and Integration**
  - Physicians are bound only for structure and syntax of MLM during creation.
  - While for the internal contents of slots, physicians are free handed to write.
  - It causes diverse concepts and terminologies for Key and values
  - We achieved interoperability and easy integration by integrating SRM with MLM creation.

Solution 2: Structure level semantic reconciliation

If Treatment Intent = Palliative
Then Treatment Plan = Radiotherapy

```
maintenance:
        title: Palliative Treatment By Physician 3;;
        mlmname: Palliative Treatment By Physician 3;;
        arden: Arden Syntax V2.7;;
        version: Version 2.7;;
        institution: SKMCH;;
        author: Dr. Physician 3;;
        specialist: Dr. Physician 3;;
        date: 13/01/2015;;
        validation: testing;;
Library:
 purpose: Experimental testing;;
 explanation: Experimental testing;;
 keywords: Oral Cavity;;
 citations: ;;
Knowledge:
    type: data driven; ;
    data: ProcedureEvents := object [ProcedureEvent];
        ProcedureEventList := read as ProcedureEvents
        { select ProcedureEvent FROM client Where ProcedureEvent.procedureCode IN ("395077000")};
        Recommendation1 := object[ProcedureEvent];
        RecommendationList := (); ;;
evoke: null_event; ;
    logic:
        ProcedureEventListDetail := EXTRACT ATTRIBUTE NAMES ProcedureEventList;
        ProcedureEvent1 := ATTRIBUTE ProcedureEventListDetail[1] FROM ProcedureEventList;
        IF( (ProcedureEvent1.procedureCode = "395077000" And ProcedureEvent1.procedureMethod = "363676003") )
        THEN
            recPart1 := new ProcedureEvent with "108290001";
            rec1 := new Recommendation1 with recPart1;
            recommendationList := recommendationList, rec1;
    action:
            WRITE recommendations.procedureMethod;;
            at stdout_dest; ; ;
  end;
```

SNOMED CT

vMR

Treatment Intent

Radiotherapy

Treatment Intent

Palliative

- **Complexity:**
  - The standard structure and syntax of MLM is already complex, while incorporating interoperability and integration standards enhance the complexity more.
- **Solution: Structure level semantic reconciliation**
  - Proposed solution hides this complexity by automatic generation of MLM.

Solution 1: b) Standard terminology and data model mapping algorithm

Word Sense Disambiguation

Maximum Concepts Overlap (MCO)

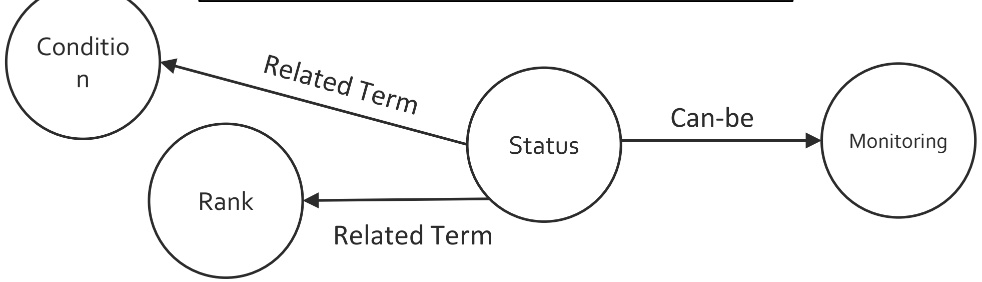
Context

Patient needs A+ blood in **emergency** ward, which is not **available** in blood **bank**

S1: Bank:- sloping land (especially the slope beside a body of water)) "they pulled The canoe up on the bank"; "he sat on the bank of the river and watched the currents

S2: Bank:- a supply or stock held in reserve for **availability** in future (especially in **emergency**))

Lexical Chain



S: Monitor:- e.g Doctor is monitoring the patient's **status**

Algorithm 2: Maximum Concepts Overlap algorithm

Begin
inputs:
W // word
S // Sentence of Context
output: Best – Sense
max-overlap = 0
foreach **sense in senses of** W
signature = set of words in the gloss + set of words in examples of sense
overlap <- COMPUTEOVERLAP (signature, S)
if overlap > max-overlap **Then**
overlap > max-overlap
Best–Sense = **sense**
end foreach
Return (Best–Sense)
End

Results And Evaluation: Solution 2 (System-Centric)

#	Definition	Result 1	Result 2	Result 3
1	Be able to define clinical information models according to a defined technical specification for structuring clinical information in EHR systems	FS	FS	FS
2	Support the semantic interoperability of EHR systems	FS	FS	FS
3	Ensure consistency of information collected by enabling the definition of clinical information models generic enough to be compatible in multiple scenarios through specialisation mechanisms for the additional constraints of each local scenario	FS	FS	FS
4	Definition and validation of the clinical information models according to a formal syntax	FS	FS	FS
5	Import and export clinical information models according to the following formal syntaxes: XML and AML	FS	FS	FS
6	Represent data types according an accepted data type standard (e.g. ISO 21090 standard or a subset of this)	FS	FS	FS
7	Support for version management, tracking changes and past history for each clinical information model	NS	NS	NS
8	Provide an automatic parser for the defined clinical information model	FS	FS	FS
9	Tools will verify that clinical information model and their instances are semantically and syntactically consistent	FS	FS	FS
10	The tool allows the author to create term bindings by connecting with Terminology Servers using (e.g. using CTS2) or another suitable terminology server communication specification	FS	FS	FS
11	Should include an intuitive graphical user interface for navigating large taxonomies	FS	FS	FS
12	Allows the user to assign one or multiple terminology/ontology concept to each node of the clinical information model structure	PS	PS	PS
13	Should include mechanisms that enable users and find a clinical information models in the repository by searching on any of its structured information properties	FS	FS	FS
14	Should export its clinical information model in at least one format that conforms to a published international standard or specification	FS	FS	FS
15	The repository and its services shall maintain a complete and audited version history for all of its clinical information models	NS	NS	NS
16	Should allow collaborative authoring of clinical information models according to the established roles. As well as recording experts and organisation participating in this process	FS	FS	FS
17	Should provide mechanisms to support multiple language translations of an clinical information model	NS	NS	NS
18	Should enable the formal definition of clinical content by domain experts without the need for technical understanding	FS	FS	FS
19	Should ensure the definition of purpose, appropriate description of usage, and precise mention of clinical information model domain	FS	FS	FS
20	Generate documentation for clinician review as MindMaps and Prototype Screens	NS	NS	NS
21	Facilitate the implementation of EHR systems that meet clinical requirements	NS	NS	NS
22	Import and export clinical information models according to Web Ontology Language (OWL)	NS	NS	NS
23	Support the organisational needs relating to the definition process, with coordination capabilities among clinical information modelling experts and clinical teams to provide a common or consensus agreed definition of the clinical information model	FS	FS	FS
24	Support the implementation of governance mechanisms to allow the establishment of an agreed editorial policy, process and quality criteria	NS	NS	NS
25	Promote the clinician adoption with a simplified and guided view well understood by them that guide their participation in the modelling process	FS	FS	FS
26	Define semantic and syntactic patterns in the form of constraints to on the selected Reference Model	FS	FS	FS
27	Provide an automatic testing environment for systems using the defined clinical information model	PS	PS	PS
28	Should allow the definition and import of Semantic patterns	NS	NS	NS

#	Definition	Result 1	Result 2	Result 3
29	Should include visualisation components for viewing complex term relationships	FS	FS	FS
30	Should facilitate the use of the clinical information model to transform/map from existing data	FS	FS	FS
31	Should allow to define transformations of the clinical information models to/from other specifications	PS	PS	PS
32	A repository service should provide a notification service to experts and systems about clinical information model updates, additions and backwards compatibility	FS	FS	FS
33	Where more than one format is supported, requester user or system will be able to nominate the preferred retrieval format	FS	FS	FS
34	Requesters of obsolete versions of an clinical information model shall be provided with a notification that an update (or updates) exist and be able to nominate the version(s) to be returned	PS	PS	PS
35	Allows to subscribe to clinical information model and terminology repositories from national/international regulatory bodies to ensure that is contained version of the clinical knowledge is updated	PS	PS	PS
36	Provide mechanisms for backward compatibility	NS	NS	NS
37	Should provide mechanisms to assign the following roles to experts participating in the clinical information modelling process and document this information in the final clinical information model produced: editor, author and reviewer	FS	FS	FS
38	Should provide mechanisms for document sharing, discussion and wiki with 2.0 capabilities to support the collaborative development	NS	NS	NS
39	Should provide the means to define the clinical and usage scope of the clinical information model in a structured and coded format, in order to be able to check for possible scope overlap with other clinical information model	FS	FS	FS
40	Should implement clinician understandable mechanisms for a guided process for local specialisation and validation purposes	FS	FS	FS
41	Should be able to create prototype screens for domain expert validation of the defined clinical information model auto-generates example GUIs to test the creation of example instances	PS	PS	PS
42	User friendly interface for clinicians including drag and drop capabilities to be able to manage multiple clinical information models easily	FS	FS	FS
43	Editorial role can examine changes, and accept or reject changes	FS	FS	FS
44	Should be easily adapted to using alternative types of (or new versions of) a Reference Model	FS	FS	FS
45	Import/select the Reference Model that will lead underpin the definition	FS	FS	FS
46	Should be able to compare 2 clinical information models covering a similar clinical domain and highlight differences	FS	FS	FS
47	Should allow to rank similar clinical information models	FS	FS	FS
48	Tools should suggest clinical information modellers with candidate terminology/ontology terms based on their semantic underlying model	FS	FS	FS
49	Should request the items to be included in the generic definition of clinical information models according to the maximal data set approach	FS	FS	FS
50	Should provide mechanisms for prioritising data items to be included in local implementations based on minimal data set approach and multiple user needs	FS	FS	FS
51	Should integrate or link to educational material to teach clinicians how to participate either in core and validation domain expert group	FS	FS	FS
52	Should allow to assign or edit the GUI presentation capabilities for local purposes, making possible that clinician/administrator edit the local presentation	FS	FS	FS
53	Tools for ongoing monitoring level of use and acceptance of clinical information models	FS	FS	FS
54	Provide mechanisms for generalisation capabilities	FS	FS	FS
55	Ensure conformance to any relevant licenses or restrictions for use of a clinical information model, and provide appropriate means for potential users of it to be informed of these	FS	FS	FS
56	Should include checkbox to verify that the resultant clinical information model quality has been developed according to the quality metrics defined by editorial role	FS	FS	FS

Comparison table of I-KAT and ArdenSuite with respect to implementation category [Essential: E, Recommended: R, Optional: O].

Priority	Req. Number	I-KAT			ArdenSuite		
		NS	PS	FS	NS	PS	FS
E	IR1			✓		✓	
E	IR2			✓		✓	
E	IR3			✓	✓		
E	IR4			✓			✓
E	IR5			✓			✓
E	IR6			✓			✓
E	IR8			✓			✓
E	IR9			✓		✓	
E	IR10			✓	✓		
E	IR11			✓			✓
E	IR13			✓			✓
E	IR14			✓			✓
E	IR16			✓	✓		
E	IR17	✓			✓		
E	IR18			✓		✓	
E	IR19			✓			✓
R	IR23			✓			✓
R	IR25			✓		✓	
R	IR26			✓	✓		
R	IR27	✓					✓
R	IR29		✓		✓		
R	IR30			✓		✓	
R	IR31		✓				✓
R	IR32			✓	✓		
R	IR33			✓			✓
R	IR34		✓		✓		
R	IR37			✓	✓		
R	IR39			✓		✓	
R	IR40			✓		✓	
R	IR41	✓					✓
R	IR42			✓		✓	
R	IR43			✓	✓		
R	ER57			✓	✓		
R	ER58			✓	✓		
O	IR21			✓		✓	
O	IR45			✓	✓		
O	IR48			✓	✓		
O	IR51			✓			✓
O	IR52	✓				✓	

Results And Evaluation: Solution 2 (Reliability Evaluation)

Reliability Testing

- Three Physicians tested the reliability of the system on 8 compiled MLMs.
- The test was handled in separate environment for each physicians.
- Physicians tested and validated MLMs on 1314 Head and Neck Cancer patient data of the real environment.
- The accuracy of compiled MLM was 100%

	Physician A	Physician B	Physician C
Rule 1	correct	correct	correct
Rule 2	correct	correct	correct
Rule 3	correct	correct	correct
Rule 4	correct	correct	correct
Rule 5	correct	correct	correct
Rule 6	correct	correct	correct
Rule 7	correct	correct	correct
Rule 8	correct	correct	correct

Samples	Physician A Decision Comparison	Physician B Decision comparison	Physician B Decision comparison
1314	100%	100%	100%

End of Document