SALUS

SALUS
Semantic Interoperability Framework

A. Anil SINACI, PhD

SRDC SOFTWARE RESEARCH & DEVELOPMENT CONSULTANCY

SALUS: Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies (http://www.salusproject.eu/)

- A STREP funded under Objective ICT-2011.5.3b) Tools and environments enabling the re-use of electronic health records which aims to
  - Enable effective integration and utilization of electronic health record (EHR) data to improve post-market safety activities on a proactive basis
  - Pilots in Lombardia Region (Italy) and Eastern Saxony (Germany)
    - WHO-UMC and ROCHE is actively involved in pilot studies
- Partners
  - SRDC Ltd, Turkey (coordinator)
  - EUROREC, France
  - WHO- UMC, Sweden
  - OFFIS, Germany
  - AGFA Healthcare, Belgium
  - ERS, Netherlands
  - LISPA, Italy
  - INSERM, France
  - TUD, Germany
  - ROCHE, Switzerland

Motivation I

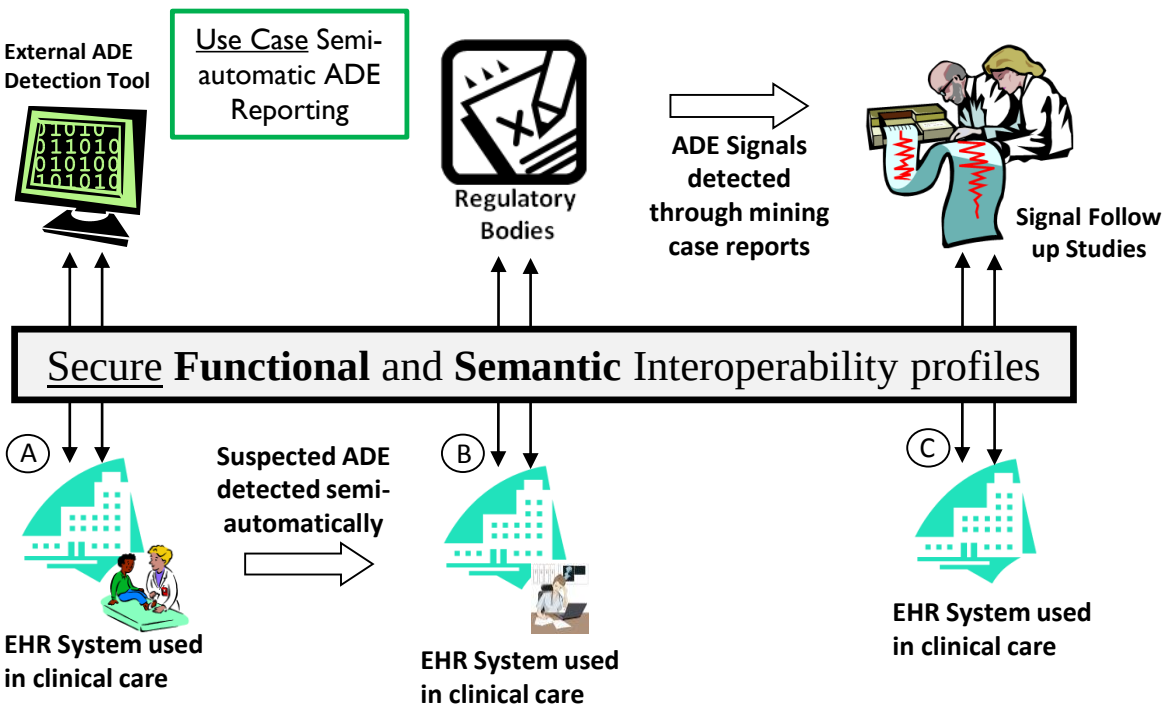
- We address the interoperability gaps between clinical research and clinical care systems for post market safety studies
- Clinical trials are focused and not adequate to ensure comprehensive drug safety
  - Limited size and scope
    - Patients with co-morbidity excluded
    - Mostly no co-medication considered
  - Designed to pick-up immediate common problems not rare adverse events
  - Cannot detect long-term adverse events

Motivation II

- Post market safety studies address this problem, but
  - Reactive based on spontaneous case safety reports
  - It is estimated that medical practitioners report only about 5% of harmful drug side effects
    - Medical professionals do not always see reporting a priority
    - Detecting adverse events may not always be straightforward
  - Approximately 5% of all hospital admissions in Europe are due to an adverse drug reaction (ADR)
  - ADRs are the fifth most common cause of hospital deaths
    - An impact assessment carried out for the European Commission has estimated that ADRs cause 197,000 deaths per year in the EU, at a total cost of €79 billion

An ideal system for ADR surveillance would combine the strengths of case reports with those of EHRs

SALUS



Case Series Characterization Tool

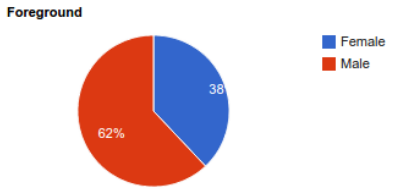
SALUS

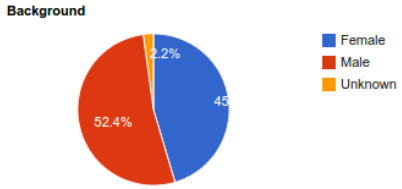
SALUS - Case Series Characterization Tool | Foreground | Background | Statistics | Risk Factor | Configuration | Results | Anil Sinaci

Results urn:3b0810f-e3af-4964-b07d-d5732c67455

Gender Distribution

Gender	Foreground	Background
Female	38% (41)	45% (15788)
Male	62% (67)	52% (18230)
Unknown	0% (0)	2% (755)





LISPA deployment and validation

- ~16 million patients
- ~550 million ambulatory diagnosis records
- ~30 million inpatient diagnosis records
- ~80 million condition records
- ~275 million drug prescriptions
- ~800.000 pregnancy records
- ~35 million vaccination records
- ~2 million allergy records

TUD deployment and validation

- ~945,000 patients
- ~13 million diagnosis
- ~114,000 adverse events
- ~56 million lab results
- ~3.8 million procedures
- ~10,000 immunization
- ~2.2 million medication forms

Average Age

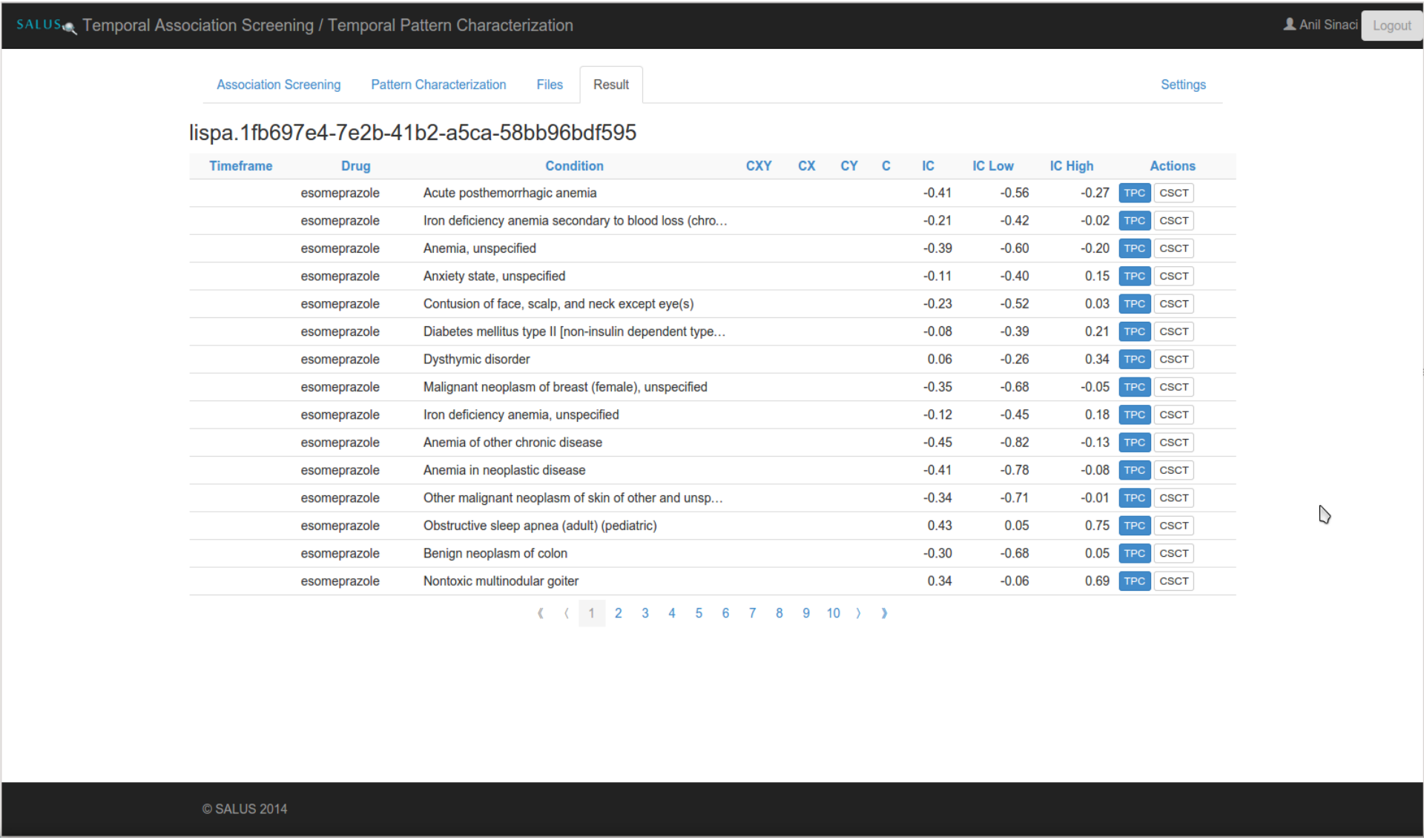
Age Distribution

Country Distribution

Common Conditions

Condition	Foreground	Background
Angiopathy	80% (87)	52% (18303)
Arrhythmia	38% (42)	19% (6834)
Arterial disorder	71% (77)	45% (15968)
Biliary tract disorder	63% (68)	7% (2557)
Bronchial disorder	33% (36)	12% (4500)

Temporal Association Screening



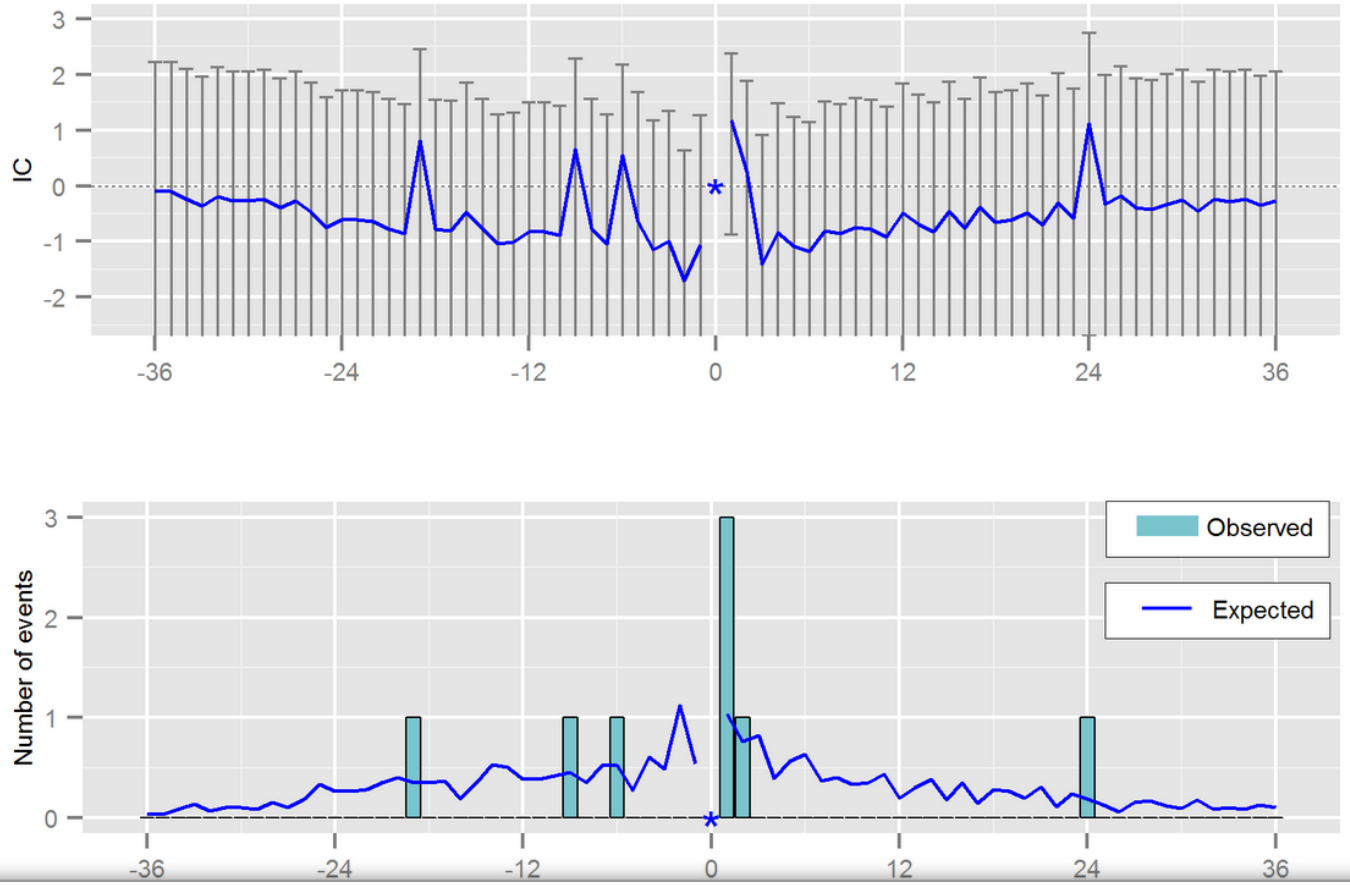
Temporal Pattern Characterization

SALUS Temporal Association Screening / Temporal Pattern Characterization

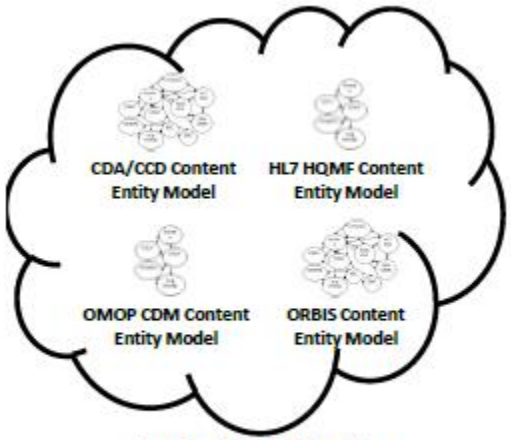
Anil Sinaci Logout

Association Screening | Pattern Characterization | Files | Result | Settings

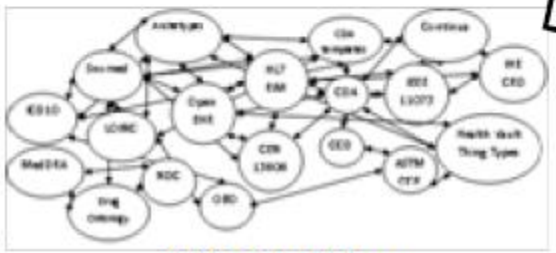
lispa.61e7fb12-e4e8-4779-a750-02c284295ae9



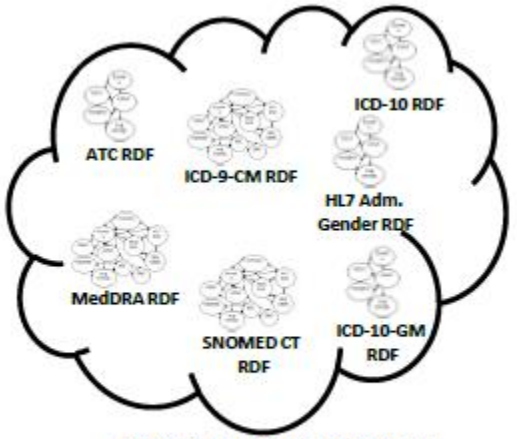
The SALUS Semantic Resource Set
Harmonized "model of meaning" as a whole



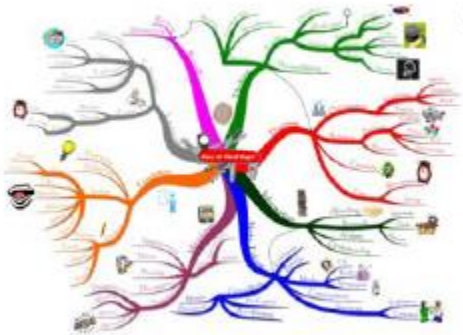
Content Entity Models



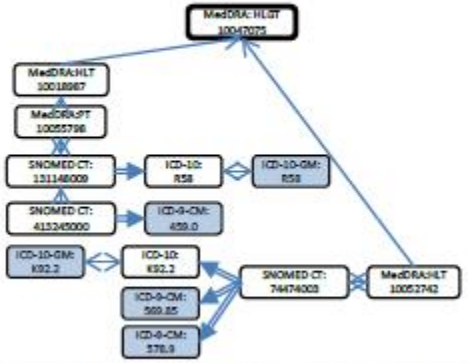
SALUS CIM Ontology



Clinical Terminology Systems as Ontologies

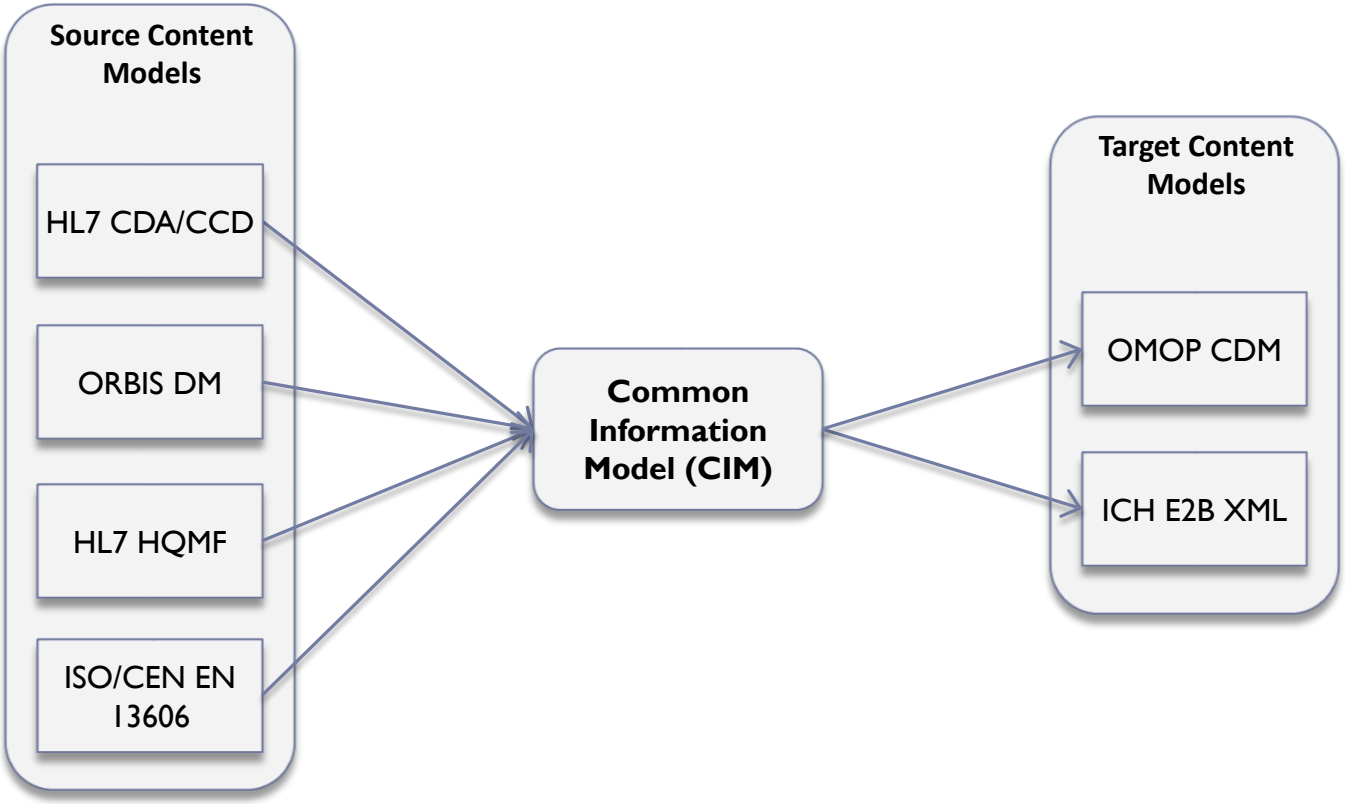


Conversion Rules between the Content Entity Models and the CIM



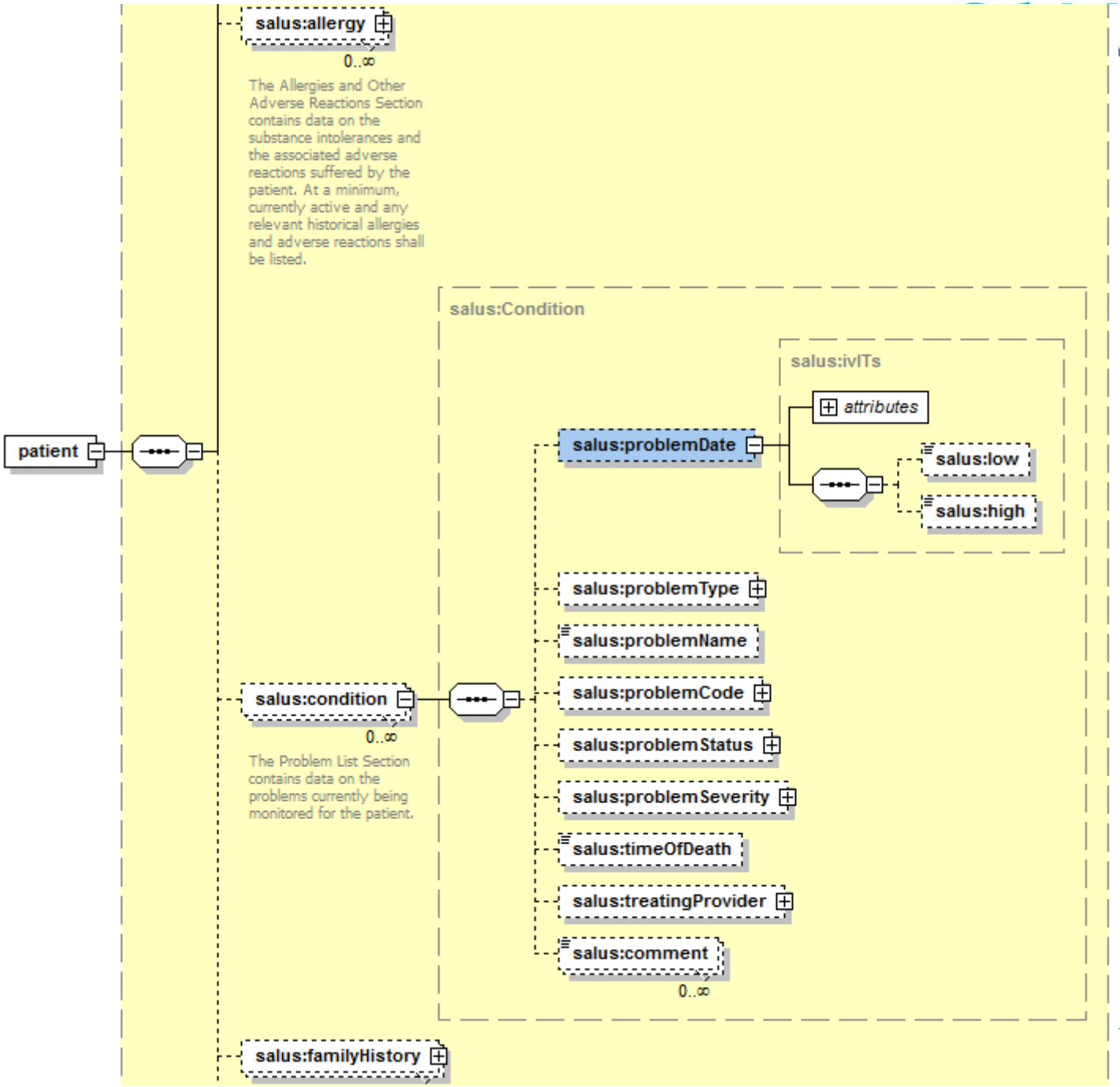
Mappings among Clinical Terminology Systems

The Content Entity Models in the Semantic Resource Set



The international models addressed in the CIM

- HL7/ASTM Continuity of Care Document (CCD) and IHE Patient Care Coordination (PCC) templates, which constitute the source model for the data provided by Lombardy Region
- HITSP C32 and C83 components
- greenCDA representation of HITSP C32
- Consolidated CDA (C-CDA) Templates Guide
- HL7 Clinical Statement Model
- OMOP CDM, which is a target model in two of our pilot application scenarios
- ICH Data Elements for Transmission of Individual Case Safety Reports E2B(R2), which is a target model in our ADE reporting scenario
- Common Data Model of the Mini-Sentinel pilot project
- ISO/CEN EN 13606 archetypes relevant to our scenarios
- HL7 Health Quality Measures Format (HQMF), for representing the population based queries

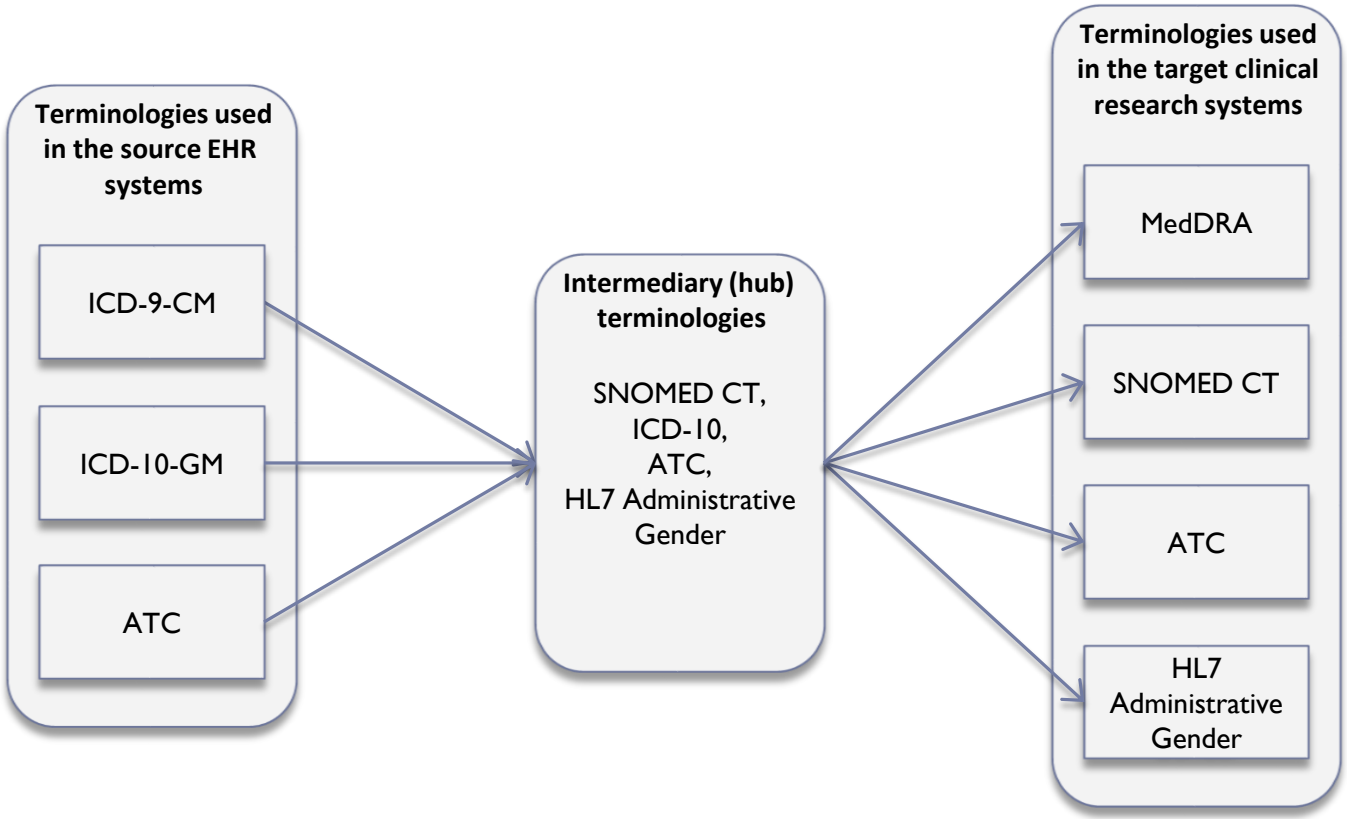

salus:allergy
0..∞
The Allergies and Other Adverse Reactions Section contains data on the substance intolerances and the associated adverse reactions suffered by the patient. At a minimum, currently active and any relevant historical allergies and adverse reactions shall be listed.
salus:Condition
salus:ivlTs
attributes
salus:low
salus:high
patient
salus:problemDate
salus:problemType
salus:problemName
salus:problemCode
salus:problemStatus
salus:problemSeverity
salus:timeOfDeath
salus:treatingProvider
salus:comment
0..∞
salus:condition
0..∞
The Problem List Section contains data on the problems currently being monitored for the patient.
salus:familyHistory

An example CIM Condition instance representing an active asthma problem

```
@prefix foaf:  <http://xmlns.com/foaf/0.1/> .
@prefix salus: <http://www.salusproject.eu/ontology/common-information-model#>.

[ rdf:type salus:Condition ;
  salus:problemCode
    [ rdf:type salus:cd ;
      salus:code "493" ;
      salus:codeSystem "2.16.840.1.113883.6.2" ;
      salus:codeSystemName "ICD-9-CM" ;
      salus:displayName "Asthma"
    ] ;
  salus:problemDate
    [ rdf:type salus:ivlTs ;
      salus:low "2003-08-01T00:00:00"^^xsd:dateTime
    ] ;
  salus:problemName "Asthma" ;
  salus:problemSeverity
    [ rdf:type salus:cd ;
      salus:code "H" ;
      salus:codeSystem "2.16.840.1.113883.5.1063" ;
      salus:codeSystemName "ObservationValue" ;
      salus:displayName "High"
    ] ;
  salus:problemStatus
    [ rdf:type salus:cd ;
      salus:code "55561003" ;
      salus:codeSystem "2.16.840.1.113883.6.96" ;
      salus:codeSystemName "SNOMED CT" ;
      salus:displayName "Active"
    ] ;
  salus:treatingProvider
    [ rdf:type salus:HealthcareProvider ;
      salus:providerID
        [ rdf:type salus:ii ;
          salus:extension "54321678906" ;
          salus:root "2.16.840.1.113883.2.9.4.3.2"
        ] ;
      salus:providerRole
        [ rdf:type salus:cd ;
          salus:code "309345004" ;
          salus:codeSystem "2.16.840.1.113883.6.96" ;
          salus:codeSystemName "SNOMED CT" ;
          salus:displayName "Chest Physicians"
        ] ;
      foaf:familyName "Passerini" ;
      foaf:givenName "Fabiola" ;
      foaf:title "Dr."
    ]
]
```

Using 3rd party ontologies (*In order of preference:* SNOMED CT, schema.org, W3C, Others)

Clinical Terminology Systems as ontologies



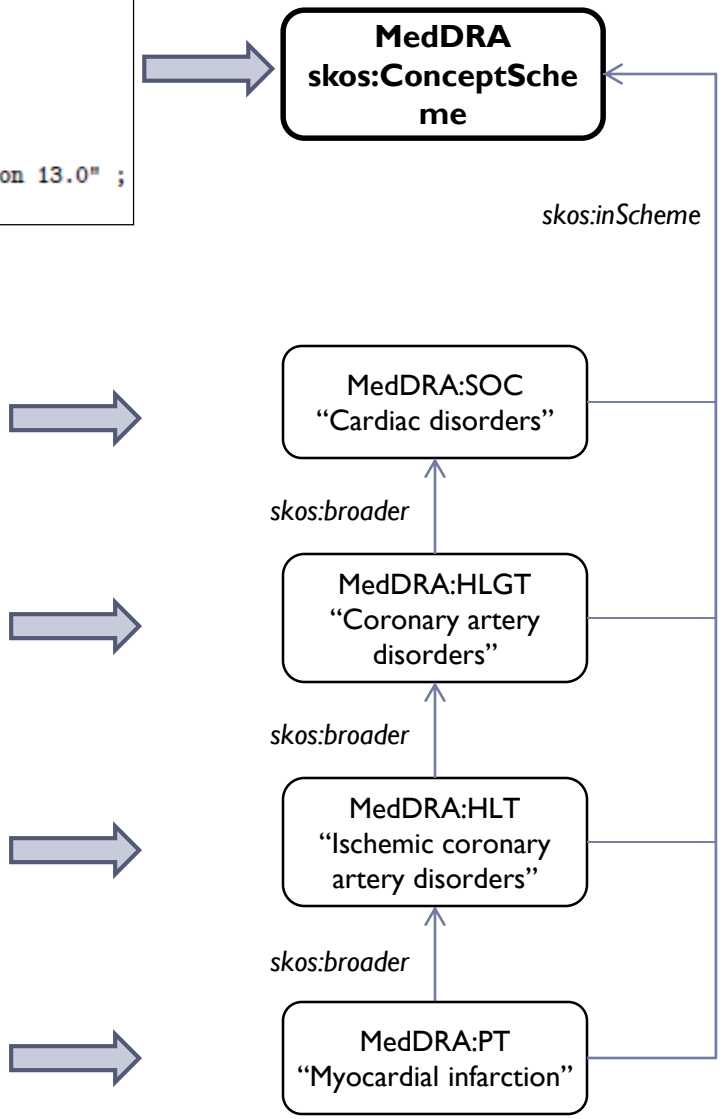
system ontologies in the Semantic Resource Set

SALUS

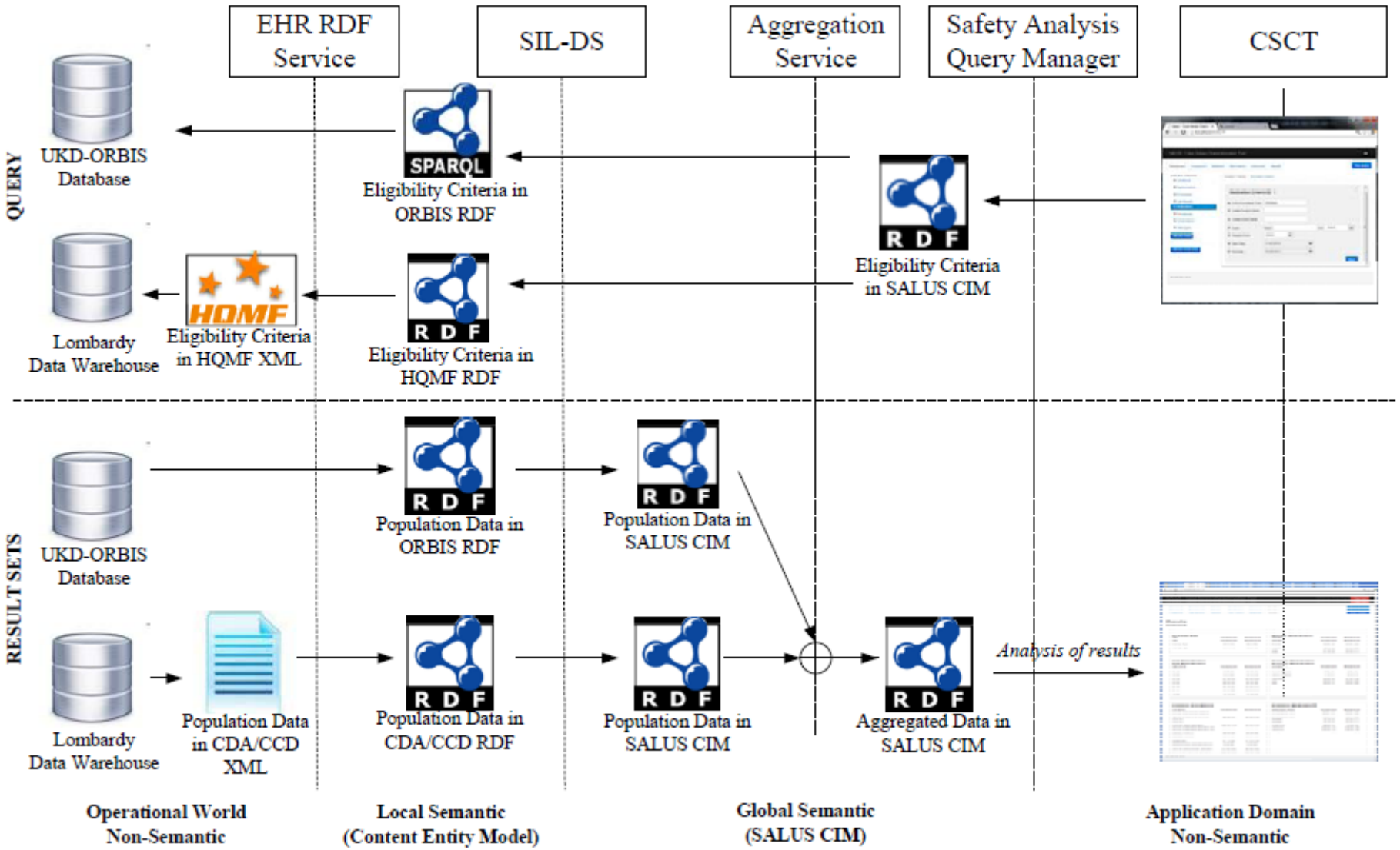
Terminology System	# of concepts
MedDRA	20,856
SNOMED CT Clinical Findings	97,139
ICD-9-CM	21,669
ICD-10	12,318
ICD-10-GM	15,801
WHO ATC	5,718
HL7 AdministrativeGender	3
Total	**173,504**

```
@prefix MDR:     <http://purl.bioontology.org/ontology/MDR/> .
@prefix foaf:    <http://xmlns.com/foaf/0.1/> .
@prefix iso:     <uri:iso.org:9834#> .
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .

<http://purl.bioontology.org/ontology/MDR>
      rdf:type skos:ConceptScheme ;
      rdfs:label "MedDRA" ;
      foaf:name "Medical Dictionary for Regulatory Activities Terminology, Version 13.0" ;
      iso:oid "2.16.840.1.113883.6.163" .
```

```
<http://purl.bioontology.org/ontology/MDR/10007541>
      a       skos:Concept ;
      MDR:level "SOC" ;
      skos:inScheme <http://purl.bioontology.org/ontology/MDR> ;
      skos:notation "10007541" ;
      skos:prefLabel "Cardiac disorders" .

<http://purl.bioontology.org/ontology/MDR/10011082>
      a       skos:Concept ;
      MDR:level "HLGT" ;
      skos:broader <http://purl.bioontology.org/ontology/MDR/10007541> ;
      skos:inScheme <http://purl.bioontology.org/ontology/MDR> ;
      skos:notation "10011082" ;
      skos:prefLabel "Coronary artery disorders" .

<http://purl.bioontology.org/ontology/MDR/10011085>
      a       skos:Concept ;
      MDR:level "HLT" ;
      skos:broader <http://purl.bioontology.org/ontology/MDR/10011082> ;
      skos:inScheme <http://purl.bioontology.org/ontology/MDR> ;
      skos:notation "10011085" ;
      skos:prefLabel "Ischaemic coronary artery disorders" .

<http://purl.bioontology.org/ontology/MDR/10028596>
      a       skos:Concept ;
      MDR:level "PT" ;
      skos:broader <http://purl.bioontology.org/ontology/MDR/10011085> ,
            <http://purl.bioontology.org/ontology/MDR/10065875> ;
      skos:inScheme <http://purl.bioontology.org/ontology/MDR> ;
      skos:notation "10028596" ;
      skos:prefLabel "Myocardial infarction" .
```



The complete transformation and mediation cycle of different content models



```
[]  a salus:EligibilityCriteria ;
    salus:inclusionCriteria [
        a salus:CriteriaGroup ;
        salus:conjunctionCode [ a salus:cd ; salus:code "AND" ] ;
        salus:groupItem _:med , _:cond ;
        salus:negationIndicator "false"^^xsd:boolean
    ].

_:med a salus:Criterion ;
    salus:clinicalStatement [
        a salus:Medication ;
        salus:medicationInformation [
            a salus:MedicationInformation ;
            salus:codedActiveIngredient [
                a salus:cd ;
                salus:code "C08CA05" ;
                salus:displayName "nifedipine" ;
                salus:codeSystem "2.16.840.1.113883.6.73" ;
                salus:codeSystemName "ATC"
            ]
        ]
    ] .

_:cond a salus:Criterion ;
    salus:clinicalStatement [
        a salus:Condition ;
        salus:problemCode [
            a salus:cd ;
            salus:code "10028596" ;
            salus:displayName "Myocardial infarction" ;
            salus:codeSystem "2.16.840.1.113883.6.163" ;
            salus:codeSystemName "MedDRA"
        ] ;
        salus:temporalRelation [
            a salus:TemporalConstraint ;
            salus:targetCriterion _:med ;
            salus:typeCode [ a salus:cd ; salus:code "SAS" ] ;
            salus:pauseQuantity [
                a salus:ivlPq ;
                salus:low [ a salus:pq ; salus:value "2" ; salus:unit "wk" ]
            ]
        ]
    ] .
```

Finally, a salus:CriteriaGroup defined by "AND"ing the criterions below, and assigned as the inclusion criteria

Criterion 1: A salus:Medication instance with active ingredient code referring to nifedipine from the ATC terminology system

Criterion 2: A salus:Condition instance with problem code referring to myocardial infarction (MI) from the MedDRA terminology system

Temporal relation from the condition to medication, indicating that the MI shall "start after start of" (SAS) nifedipine intake, with an allowed time interval of 2 weeks

Formalizing EHR Data: ontmalizer

- While returning results to queries;
  - Input: Population data as CDA/CCD **Content Model** instances (i.e. XML)
  - Output: Population data as CDA/CCD **Content Entity Model** instances (i.e. RDF)

- Ontmalizer performs comprehensive transformations of XML Schemas (XSD) and XML data to RDF/OWL automatically
  - The state of the art free and/or open source tools could not handle complex schemas
  - Only the commercial version of TopBraid Composer is able to handle HL7 CDA R2 XSD
  - We implemented from scratch by using XSOM, Xerces and Jena libraries
  - Tested heavily
  - It is available under ASL at https://github.com/srdc/ontmalizer

```
<observation classCode="OBS" moodCode="EVN">
  <code code="55607006" displayName="Problem" codeSystem="2.16.840.1.113883.6.96"
    codeSystemName="SNOMED CT"/>
  <effectiveTime>
    <low value="20090801"/>
  </effectiveTime>
  <value xsi:type="CD" code="410.0" displayName="Acute myocardial infarction, of anterolateral wall"
    codeSystem="2.16.840.1.113883.6.2" codeSystemName="ICD-9-CM">
    <originalText>Acute myocardial infarction, of anterolateral wall</originalText>
  </value>
</observation>
```

```
<http://www.srdc.com.tr/ontmalizer/instance#INS7778722_POCD_MT000040.Observation_1>
   a <urn:hl7-org:v3#POCD_MT000040.Observation> ;
   v3:code
      [ a v3:CD ;
        v3:code "55607006"^^v3:csDatatype ;
        v3:codeSystem "2.16.840.1.113883.6.96"^^v3:uidDatatype ;
        v3:codeSystemName "SNOMED CT"^^v3:stDatatype ;
        v3:displayName "Problem"^^v3:stDatatype
      ] ;
   v3:effectiveTime
      [ a v3:IVL_TS ;
        v3:low
          [ a v3:IVXB_TS ;
            v3:value "20090801"^^v3:tsDatatype
          ]
      ] ;
   v3:value
      [ a v3:CD ;
        v3:code "410.0"^^v3:csDatatype ;
        v3:codeSystem "2.16.840.1.113883.6.2"^^v3:uidDatatype ;
        v3:codeSystemName "ICD-9-CM"^^v3:stDatatype ;
        v3:displayName "Acute myocardial infarction, of anterolateral wall"^^v3:stDatatype;
        v3:originalText
          [ a v3:ED ;
            v3:textContent "Acute myocardial infarction, of anterolateral wall"^^xsd:string
          ]
      ] .
```

```
<http://www.srdc.com.tr/ontmalizer/instance#INS7778722_POCD_MT000040.Observation_1>
   a <urn:hl7-org:v3#POCD_MT000040.Observation> ;
   v3:code
      [ a v3:CD ;
        v3:code "55607006"^^v3:csDatatype ;
        v3:codeSystem "2.16.840.1.113883.6.96"^^v3:uidDatatype ;
        v3:codeSystemName "SNOMED CT"^^v3:stDatatype ;
        v3:displayName "Problem"^^v3:stDatatype
      ] ;
   v3:effectiveTime
      [ a v3:IVL_TS ;
        v3:low
          [ a v3:IVXB_TS ;
            v3:value "20090801"^^v3:tsDatatype
          ]
      ] ;
   v3:value
      [ a v3:CD ;
        v3:code "410.0"^^v3:csDatatype ;
        v3:codeSystem "2.16.840.1.113883.6.2"^^v3:uidDatatype ;
        v3:codeSystemName "ICD-9-CM"^^v3:stDatatype ;
        v3:displayName "Acute myocardial infarction, of anterolateral wall"^^v3:stDatatype;
        v3:originalText
          [ a v3:ED ;
            v3:textContent "Acute myocardial infarction, of anterolateral wall"^^xsd:string
          ]
      ] .
```

```
_:t0
   a salus:Condition ;
   salus:problemType
      [ rdf:type salus:cd ;
        salus:code "55607006" ;
        salus:codeSystem "2.16.840.1.113883.6.96" ;
        salus:codeSystemName "SNOMED CT" ;
        salus:displayName "Problem"
      ] ;
   salus:problemCode
      [ rdf:type salus:cd ;
        salus:code "410.0" ;
        salus:codeSystem "2.16.840.1.113883.6.2" ;
        salus:codeSystemName "ICD-9-CM" ;
        salus:displayName "Acute myocardial infarction, of anterolateral wall"
      ] ;
   salus:problemDate
      [ rdf:type salus:ivlTs ;
        salus:low "2009-08-01T00:00:00"^^xsd:dateTime
      ] ;
   salus:problemName "Acute myocardial infarction, of anterolateral wall".
```

Terminology Reasoning Service

- The EHR sources on the clinical care side and the end-users on the clinical research side use very different medical terminology systems for coding medical data

- In order to overcome the terminology reasoning challenge, we achieved the following:
  - Representation of the terminology systems as ontologies within the SALUS Semantic Resource Set,
  - Utilizing reliable terminology system mapping resources,
  - Automatically linking coded patient data with terminology system ontologies, and
  - Purpose specific materialization for high performance

Terminology mapping resources that are utilized in the Semantic Resource Set

Source System	Target System	Type of Mapping	Number of Mappings	Mapping Resource
MedDRA	SNOMED CT	exact match	10,648	OntoADR of the PROTECT project; manual improvement of UMLS mappings by PROTECT experts
ICD-9-CM	SNOMED CT	exact or broad match	16,819	OMOP Vocabulary; created manually by OMOP experts
ICD-10-CM	SNOMED CT	exact or broad match	59,122	OMOP Vocabulary; created manually by OMOP experts
ICD-10-GM	ICD-10	exact match	12,318	Identical codes in both systems
ICD-9-CM	SNOMED CT	close match	43,086	BioPortal; manual review by SALUS experts before inclusion
ICD-10-CM	SNOMED CT	close match	45,022	BioPortal; manual review by SALUS experts before inclusion

SALUS Common Data Elements

- The total number of the identified CDEs is 163

Data Element	
Name	Description
Patient.ID.II	Identifier of the patient
Patient.Title.String	Title/prefix of the patient
Patient.GivenName.String	Given name of the patient
Patient.FamilyName.String	Family name of the patient
Patient.Gender.CD	Gender of the patient
Allergy.AdverseEventType.CD	Coded type of the allergy / intolerance / adverse event (e.g. drug allergy, food intolerance)
Allergy.TimeInterval.IVLTS	Effective time interval of the allergy / intolerance / adverse event
Allergy.Product.CD	Product (i.e. substance) that causes the allergy / intolerance / adverse event (e.g. egg protein, dust, nifedipine)
Allergy.Reaction.Condition	The condition which occur as a reaction to the allergy / intolerance / adverse event; can be any condition

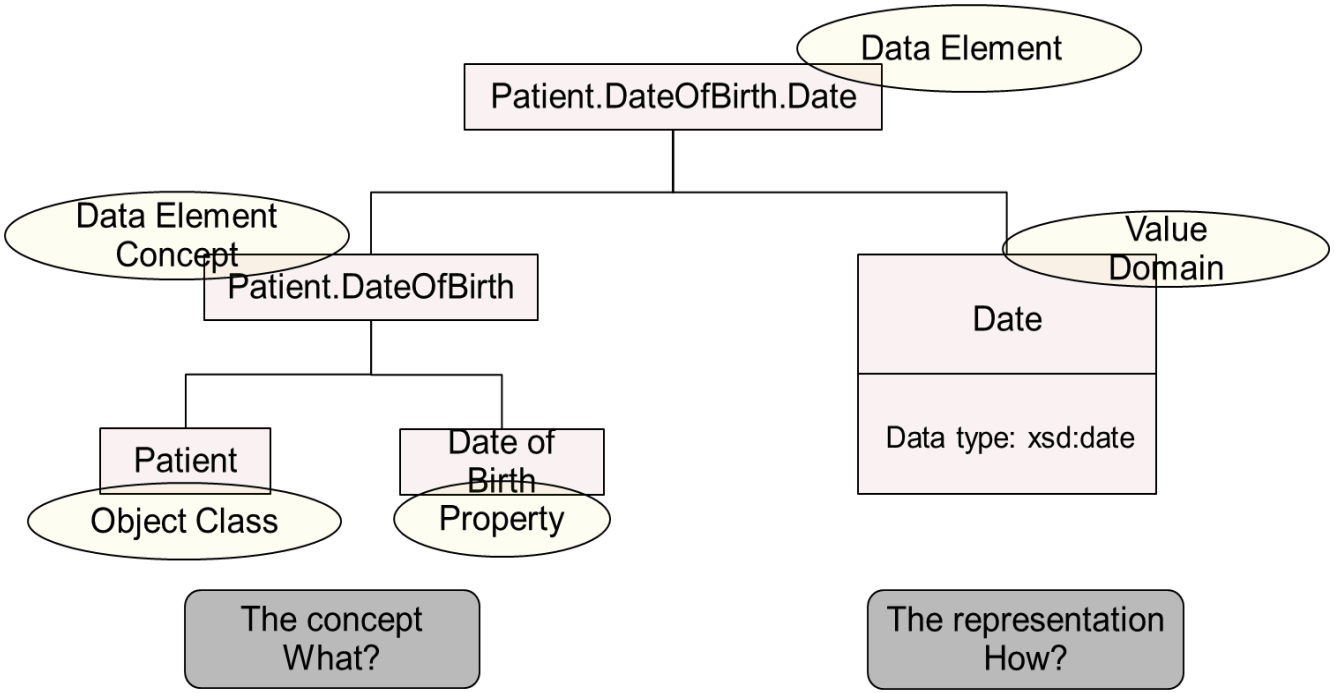
SALUS Semantic MDR

- The design and implementation of CDE Repository go beyond the requirements of SALUS interoperability framework
  - *Federated Semantic Metadata Repository*
- During the elicitation of SALUS CDEs
  - several other ***common data element models*** have been analysed
  - one of the major deficiencies is that most of them are published through PDF documents or spread sheets
    - not accessible in a ***machine-processable*** way
  - HITSP C154, C32
  - FHIM CIM
  - S&I CEDD
  - CDISC CDASH, SDTM
  - BRIDG CDM
  - Intermountain Healthcare CEM
  - Mini-Sentinel CDM
  - i2b2

it is not practical to expect all of these diverse initiatives and projects to stick to the same common model, and to use the same set of CDE

ISO/IEC 11179

- Family of specifications (6 parts) for metadata registries to increase the interoperability of applications with the use of data elements
- A relational metamodel
  - Generic: any data element model can be represented through regardless of the level of granularity



Federated Semantic MDRs

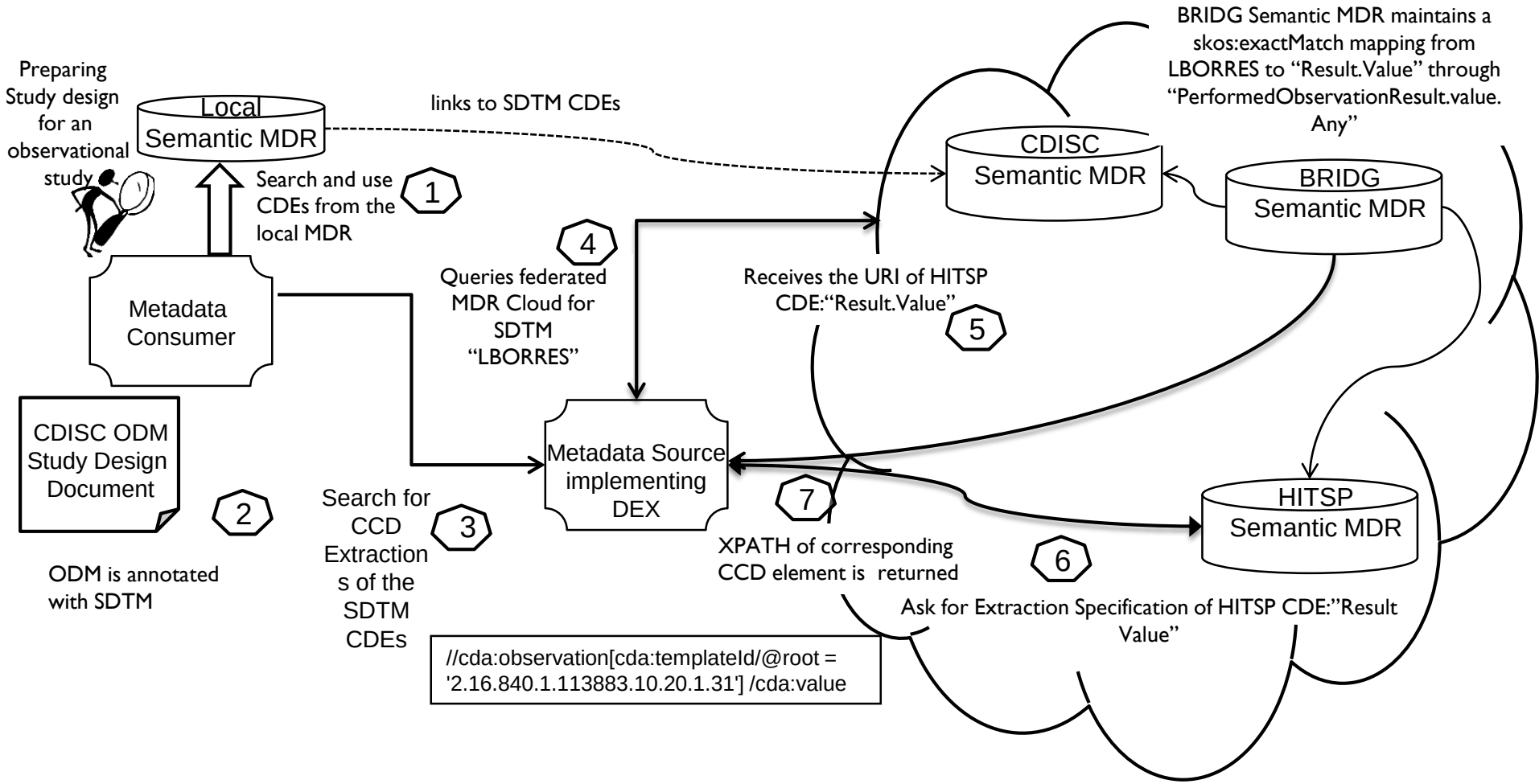
- Maintain & Manage
  - CDEs
  - the relations between CDEs
  - the components of CDEs
  - the relations between the components of CDEs
- Different CDEs from different Content Models
  - their relations and mappings are managed semantically
- A set of CDEs with lots of relations – Semantic Resource Set
  - The relations can be through the LOD cloud
- The relations may point to native representations of the Content Models
  - Extraction Specification
  - **IHE DEX Profile**

IHE DEX

- For the reuse of EHRs for clinical research
  - E.g. CCD → CDASH annotated ODM
- Can be achieved through existing IHE profiles
  - RFD, CRD, Redaction
  - The problem: ***one size fits all*** – XSLT mappings
- Power of an MDR
  - apply mappings earlier in the process
    - During the form design, data elements of the form have already been mapped to the corresponding elements in the EHR export
  - *The MDR to maintain the exact correspondences between the research and healthcare data elements*

- DEX is to support ***study feasibility***, ***patient eligibility and recruiting***, ***adverse event reporting***, ***retrospective observational studies*** as well as ***case report form pre-population***
  - existing standards for patient summaries – ASTM/HL7 CCD

An example Execution in SALUS



SALUS

Thank you for listening…

Questions

A. Anil SINACI, PhD

SRDC SOFTWARE RESEARCH & DEVELOPMENT CONSULTANCY