UTHealth | School of Biomedical Informatics
The University of Texas Health Science Center at Houston

APPLYING ONTOLOGY AND SEMANTIC WEB TECHNOLOGIES TO CLINICAL AND TRANSLATIONAL STUDIES

Cui Tao, PhD
Assistant Professor of Biomedical Informatics
University of Texas Health Science Center at Houston
School of Biomedical Informatics

Ontology & Semantic Web

History of the Semantic Web

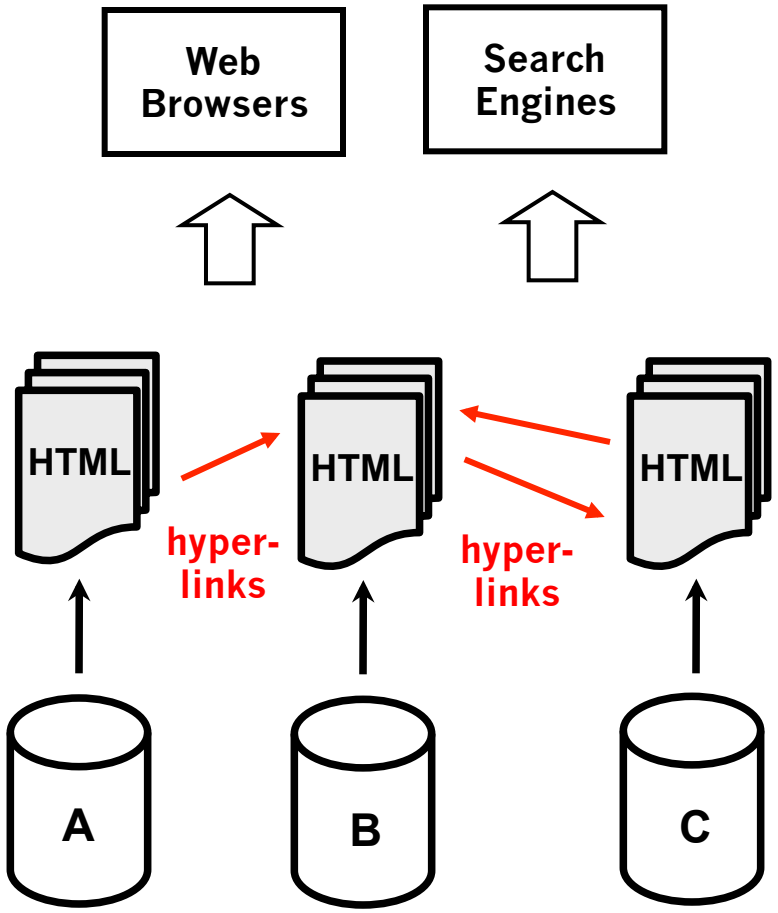
- Web was "invented" by Tim Berners-Lee (amongst others)
- TBL's original vision of the Web was much more ambitious than the reality of the existing (syntactic) Web:

"... a goal of the Web was that, if the interaction between person and hypertext could be so intuitive that the **machine-readable** information space gave an accurate representation of the state of people's thoughts, interactions, and work patterns, then **machine analysis** could become a very powerful management tool, seeing patterns in our work and facilitating our working together through the typical problems which beset the management of large organizations."

- TBL (and others) have since been working towards realising this vision, which has become known as the Semantic Web
  - E.g., article in May 2001 issue of Scientific American…

Credit: Ian Horrocks and **Alan Rector**

The Syntactic Web (Web of Documents)

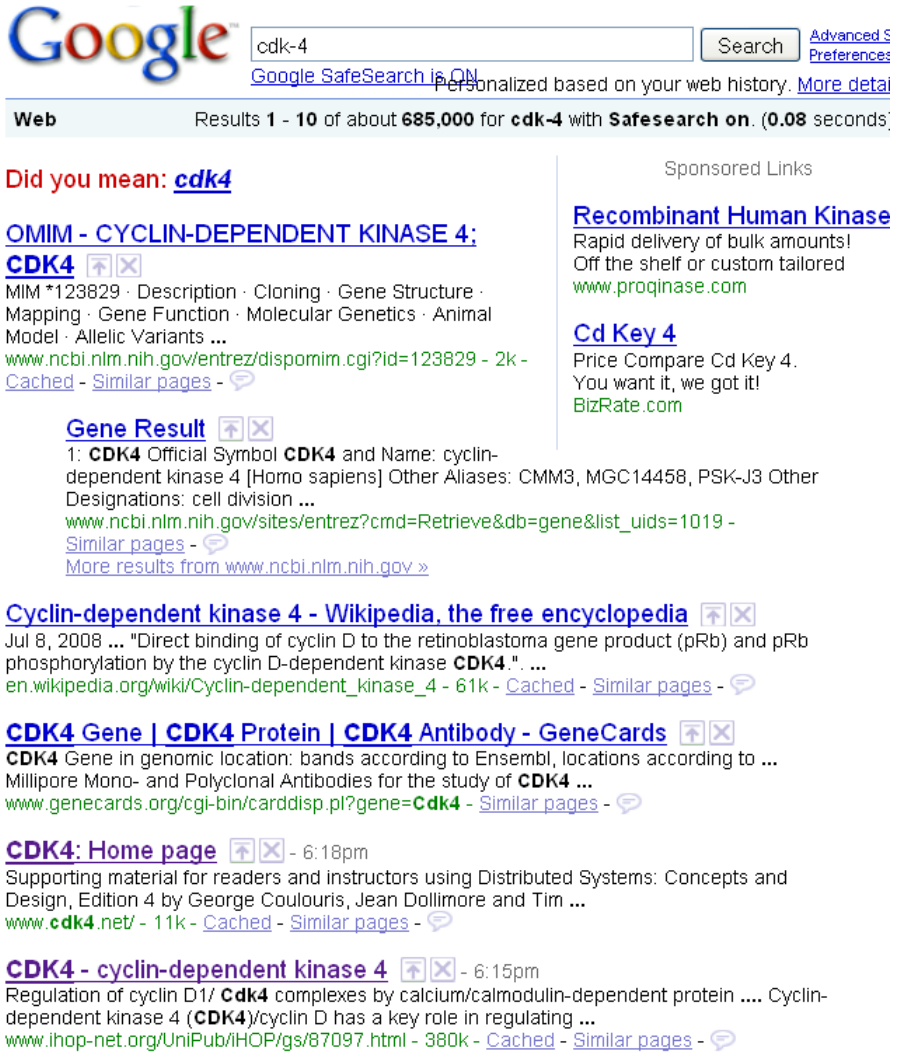


Source Chris Bizer

- **Single information space**
- **Built on URIs**
  - **globally unique IDs**
  - **retrieval mechanism**
- **Built on Hyperlinks**
  - **are the glue that holds everything together**

Search by Search Engines

Find the protein
and the animo-acids
information for gene “cdk-4"



The Answer is Here

Home | Genome | Blast / Blat | WormMart | Batch Sequences | Markers | Genetic Maps | Submit | Searches | Site Map

Find: WBGene00000406 Anything

WormBase — The Biology and Genome of C. elegans.

Gene Summary | Locus Summary | Sequence Summary | Protein Summary | EST Alignments | Genome Browser | Genetic Map | Nearby Genes | Bibliography | Tree Display | XML | Schema | Acedb Image

Gene Summary for cdk-4

Specify a gene using a gene name (unc-26), a predicted gene id (R13A5.9), or a protein ID (CE02711) cdk-4

[identification] [location] [function] [expression] [gene ontology] [alleles] [similarities] [reagents] [bibliography]

Identification

IDs:

Main name	Sequence name	Other name(s)	WB Gene ID
cdk-4 - (*Cyclin-Dependent Kinase family*) via person evidence: Michael Krause	F18H3.5	NM_077855 (inferred automatically) NM_001029420 (inferred automatically) XO136 (inferred automatically)	WBGene00000406

Concise Description: cdk-4 encodes two isoforms of a cyclin-dependent serine/threonine protein kinase orthologous to human CDK4 and CDK6 (OMIM:123829 and OMIM:603368, mutated in cutanous malignant melanoma) which complex with D-type cyclins to regulate progression through the G1 phase of the cell cycle; CDK-4 activity is essential for G1 progression in postembryonic blast cells and as a result, cdk-4 mutant animals generally arrest during larval stages; the lethality generated by cdk mutations, also seen be suppressed by mutations in lin-35/Rb suggesting that, as in other organisms, LIN-35/Rb may be a major target o and hypodermal lineages during mid-to-late embryogenesis, with postembryonic expression detected in hypoderma and cells of the somatic gonad, the vulva, and the intestine. [details]

NCBI KOGs*: Protein kinase PCTAIRE and related kinases [KOG0594]

Species: *Caenorhabditis elegans*

⊞ Other sequences

NCBI: [AceView: XO136]

Gene model(s):

Gene Model	Status	Nucleotides (coding/transcript)	Protein	Amino Acids
F18H3.5a [1, 2]	confirmed by cDNA(s)	1029/2854 bp	WP:CE18608	342 aa
F18H3.5b [1, 2, 3]	partially confirmed by cDNA(s)	1221/1704 bp	WP:CE28918	406 aa

⊞ Footnotes

⊞ Other Notes

⊞ History

Location Genetic Position: X:12.69 +/- 0.007 cM [mapping data]

Find the protein and the animo-acids information for gene “cdk-4"

What is the Problem?

Consider a typical web page:

- **Markup consists of:**
  - **rendering information (e.g., font size and color)**
  - **Hyper-links to related content**
- **Semantic content is accessible to humans but not (easily) to computers…**

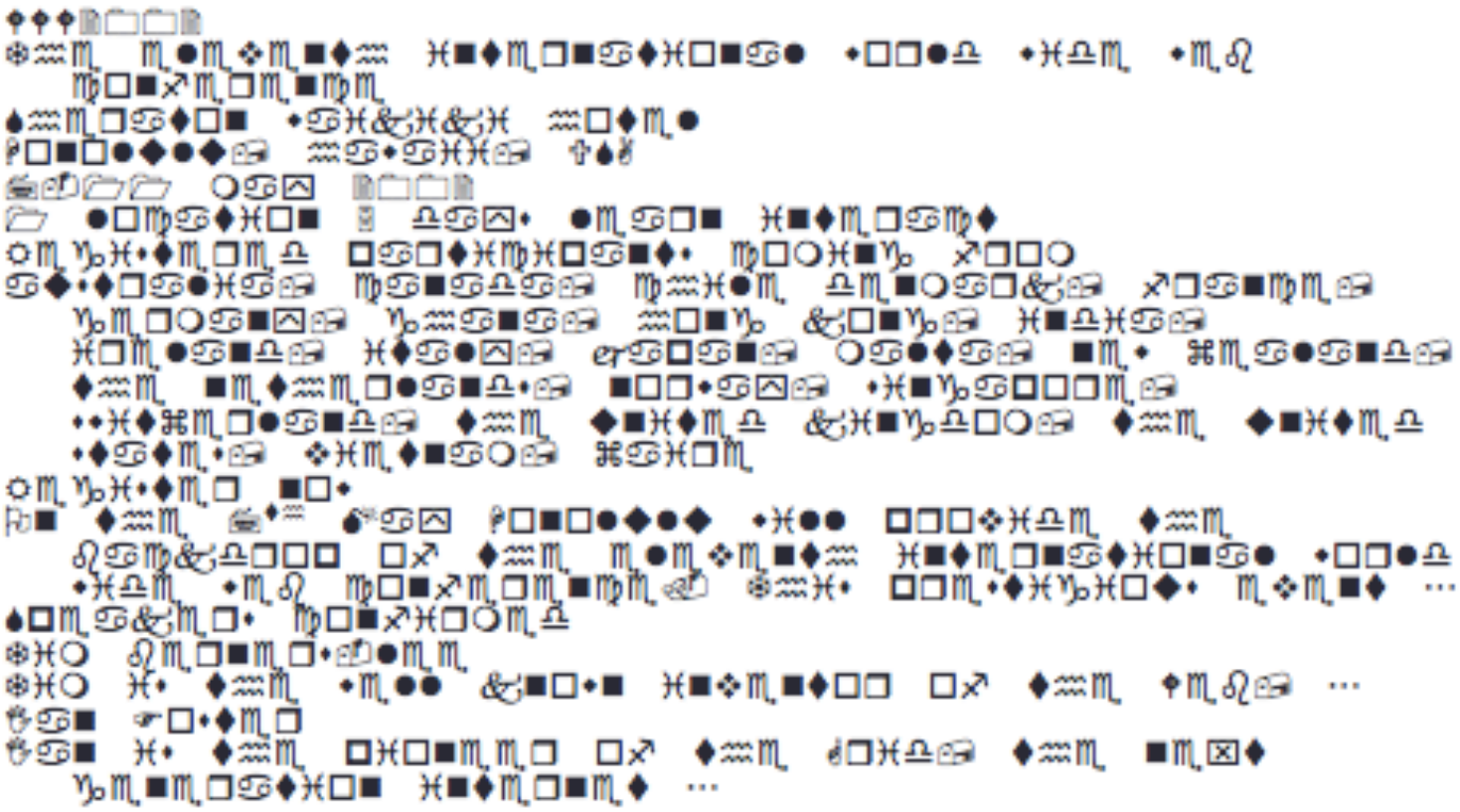
Credit: Ian Horrocks and **Alan Rector**

What information can we see…

WWW2002
The eleventh international world wide web conference
Sheraton waikiki hotel
Honolulu, hawaii, USA
7-11 may 2002
1 location 5 days learn interact
Registered participants coming from
australia, canada, chile denmark, france, germany, ghana, hong kong, india, ireland, italy, japan, malta, new zealand, the netherlands, norway, singapore, switzerland, the united kingdom, the united states, vietnam, zaire
Register now
On the 7th May Honolulu will provide the backdrop of the eleventh international world wide web conference. This prestigious event …
Speakers confirmed
Tim berners-lee
Tim is the well known inventor of the Web, …
Ian Foster
Ian is the pioneer of the Grid, the next generation internet …

Credit: Ian Horrocks and **Alan Rector**

What information can a machine see…



Credit: Ian Horrocks and **Alan Rector**

Need to Add “Semantics”

- Semantic annotation with respect to a domain ontology
- Ontology is the philosophical study of the nature of being, existence or reality in general, as well as the basic categories of being and their relations
- In computer science and information science, an ontology is

  a formal representation of the knowledge:
  - concepts within a domain
  - the relationships between these concepts
  - constraints
  - It is used to
    - describe the domain
    - reason about the properties of that domain
    - consistency checking

Clinical Informatics using Ontologies

- Big and Complex Data: large-scale deployment of Electronic Health Record (EHR)
- Highly diverse
- New opportunities of secondary use of EHR for clinical and translational studies
- Interoperability of EHR data, clinical knowledge, and application in healthcare IT
- Ontologies and semantic web technologies offer potential solutions


HealthIT.gov
Newsroom
Multimedia
Providers & Professionals
Patients & Families
Policy Researchers & Implementers
HealthIT.gov › About ONC
Print | Share
Newsroom
Health IT in the News
News Releases
Events
Fact Sheets
Infographics
Recent Updates
About ONC
The Office of the National Coordinator for Health Information Technology (ONC) is at the forefront of the administration's health IT efforts and is a resource to the entire health system to support the adoption of health information technology and the promotion of nationwide health information exchange to improve health care. ONC is organizationally located within the Office of the Secretary for the U.S. Department of Health and Human Services (HHS).
ONC is the principal federal entity charged with coordination of nationwide efforts to implement and use the most advanced health information technology and the electronic exchange of health information. The position of National Coordinator was created in 2004, through an Executive Order, and legislatively mandated in the Health Information Technology for Economic and Clinical Health Act (HITECH Act) of 2009.
Media Questions
Contact Peter Ashkenaz if you have media questions. Your queries will be addressed within one business day.
Get On-the-Ground Support
Email:
Peter.Ashkenaz@hhs.gov
Telephone: (202) 260-6342

The SHARP

AREA ONE: Security and Health Information Technology – The University of Illinois at Urbana-Champaign is helping develop technologies and policy recommendations that reduce privacy and security risks and increase public trust.

AREA TWO: Patient-Centered Cognitive Support – Innovative cognitive research is being led by the University of Texas, Houston to harness the power of health IT to integrate and support physician reasoning and decision-making as providers care for patients.

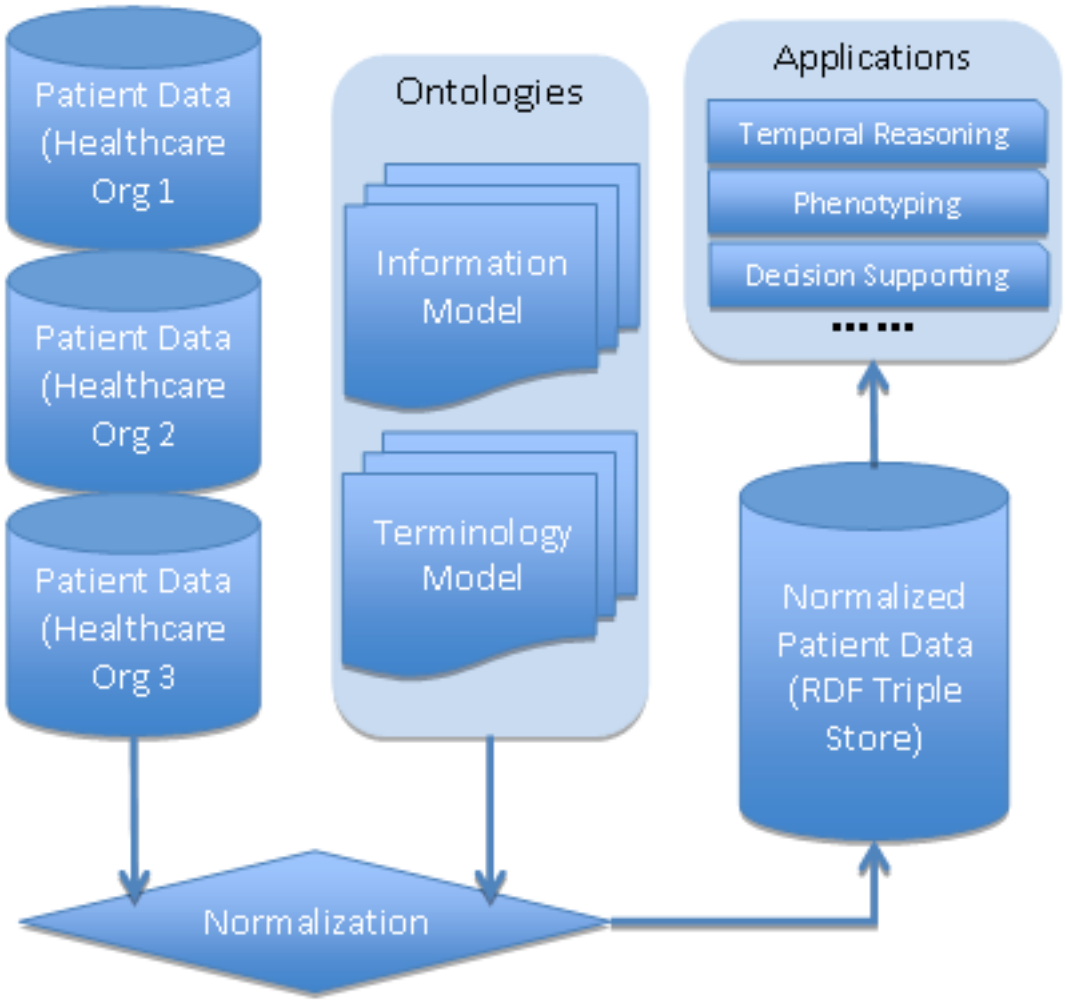
AREA THREE: Health Care Application and Network Design – Harvard University is leading platform-based research to create new and improved system designs that facilitate information exchange while ensuring the accuracy, privacy, and security of electronic health information.

AREA FOUR: Secondary Use of EHR Information – Mayo Clinic of Medicine is developing strategies to improve the overall quality of healthcare by leveraging existing EHR data to generate new, environmentally appropriate, best practice suggestions.

Outline

- Semantic Web representation of EHR data
  - Model representation
  - Instance representation
  - Semantic reasoning
- Phenotyping application
- Temporal Relation Modeling, Extraction, and Reasoning (TIMER)

EHR Data Representation



Tao C, et al. Toward semantic web based knowledge representation and extraction from electronic health records. CIKM 2011

Standard Model Representation & Ontology Normalization

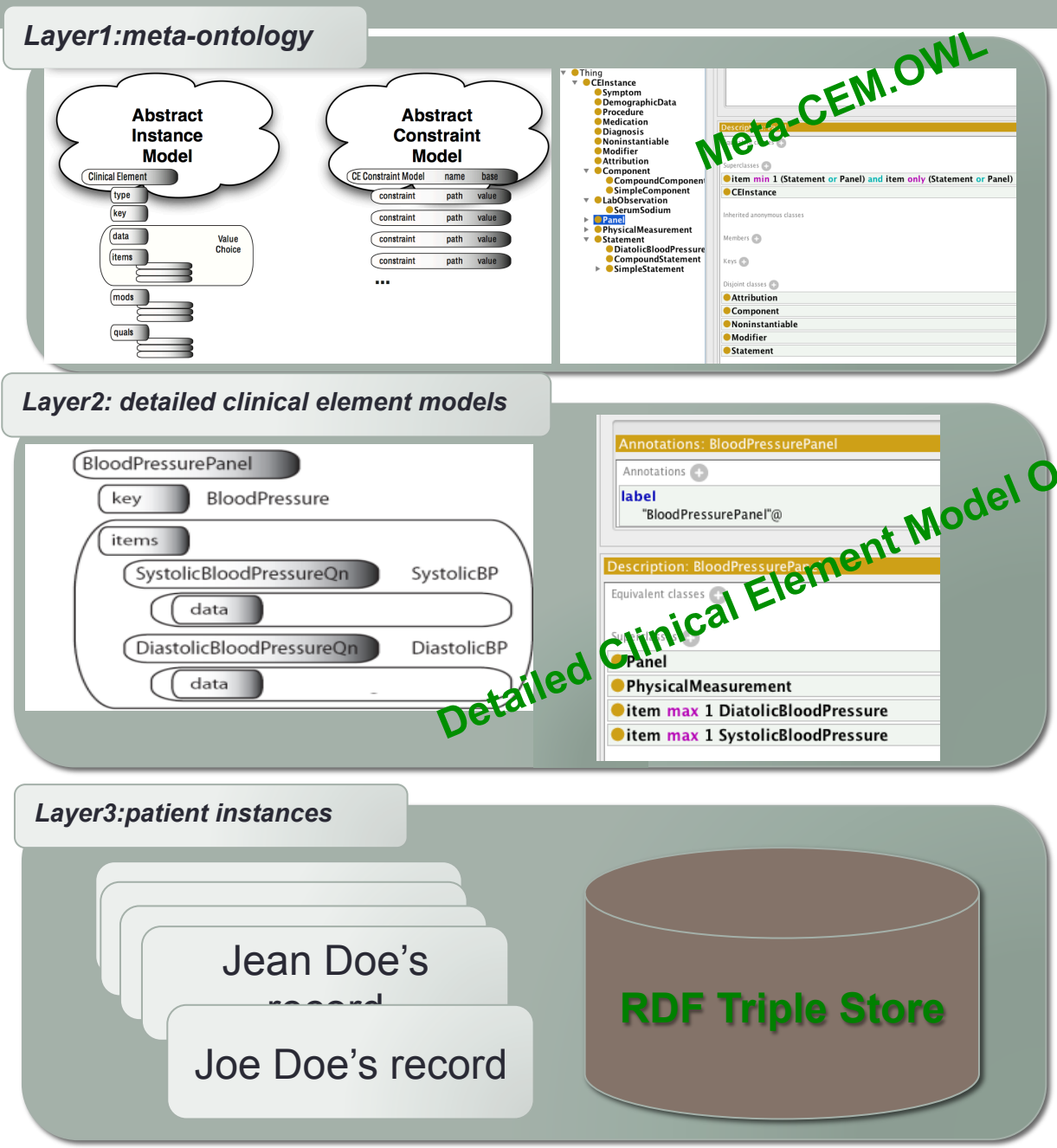
- Terminology model:
  - Represent domain knowledge
  - SNOMED CT, LOINC, RxNorm, etc
  - Labetalol is an anti-hypertension drug
- Information model:
  - Provides a consistent architecture for representing healthcare and clinical application-specific concepts in EHR systems
  - Clinical Element Model （CEM)
  - Blood pressure measurements:
    - Systolic and diastolic blood pressure measurements
    - Device
    - Position
    - Inpatient/outpatient

Clinical Element Model (CEM)

- Information model used in SHARPn (Strategic Health IT Advanced Research Projects - Secondary Use of EHR Data)
- Ensure semantic interoperability for:
  - Data representation
  - Data interpretation
  - Data exchange within and across heterogeneous sources and applications
- Represented in CEML/CDL (Constraint Definition Language)
  - Define syntax and grammar
  - Not semantics

CEM-OWL

- Explicit and formal semantic representation
- Web Ontology Language (OWL) :
  - Define relationships
  - Define classes
  - Define constraints
- Consistency checking
- Link to other domain terminologies
- Semantic reasoning
- Directly using semantic web tools

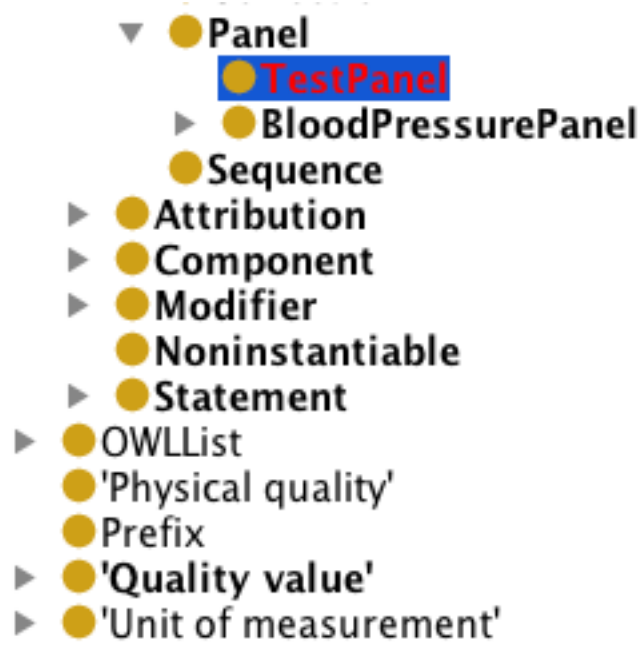


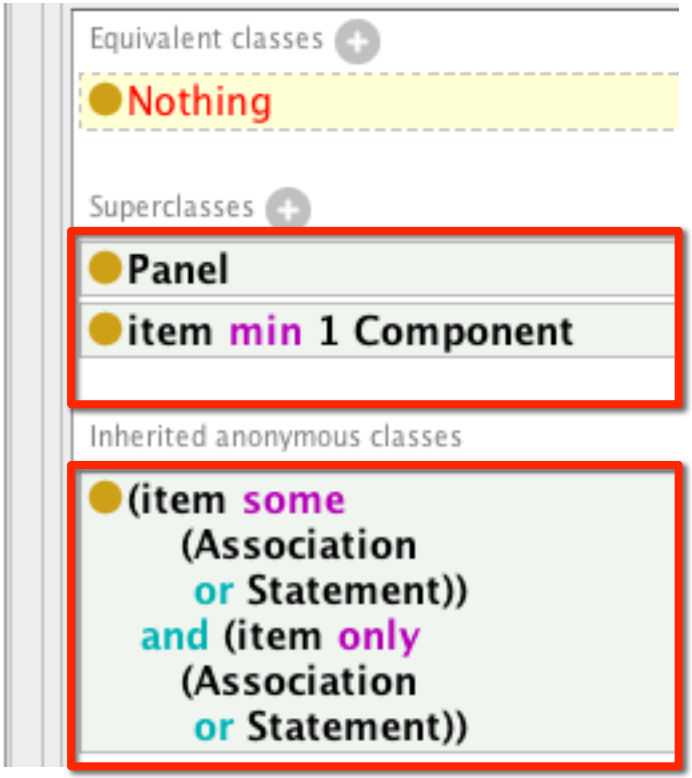
Tao C, et al, A Semantic-Web Oriented Representation of the Clinical Element Model for Secondary Use of EHR Data, JAMIA. 2012

Semantic Reasoning

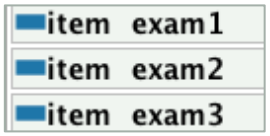
- Consistency checking
  - Cardinality constraints
  - Data types
  - Property allowed domains and ranges
  - Permissible values in value sets
- Classification and reasoning

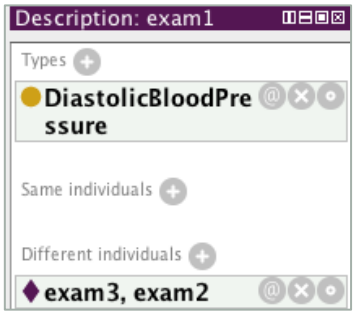
Consistency Checking

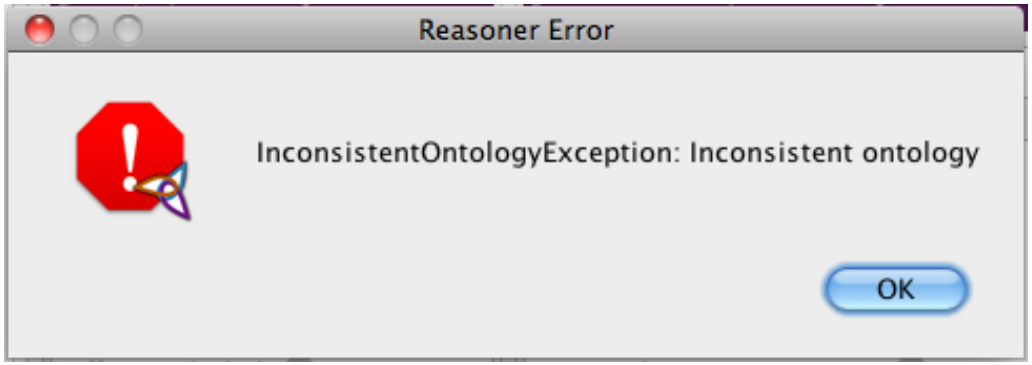




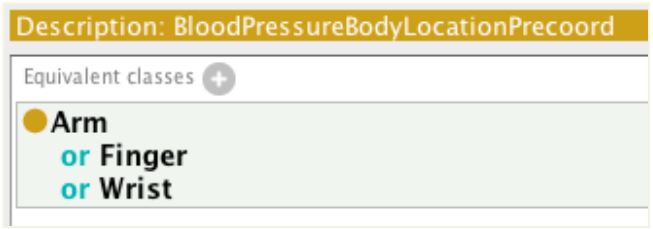
Consistency Checking

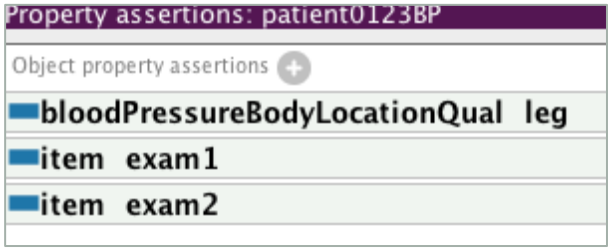


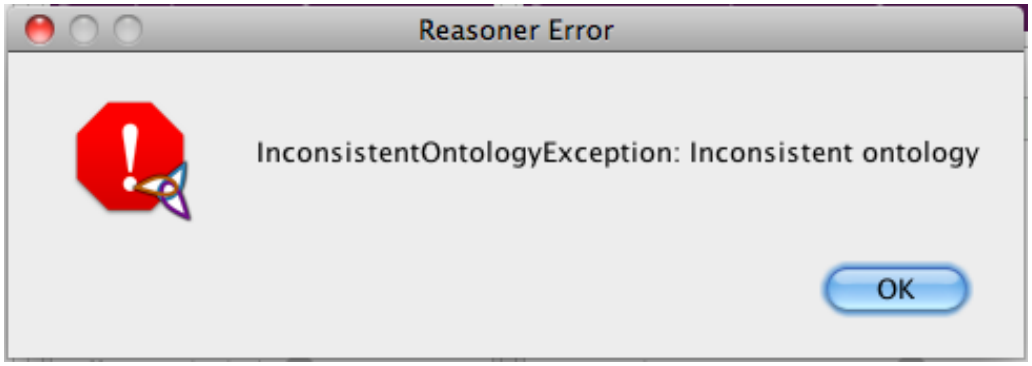




Consistency Checking



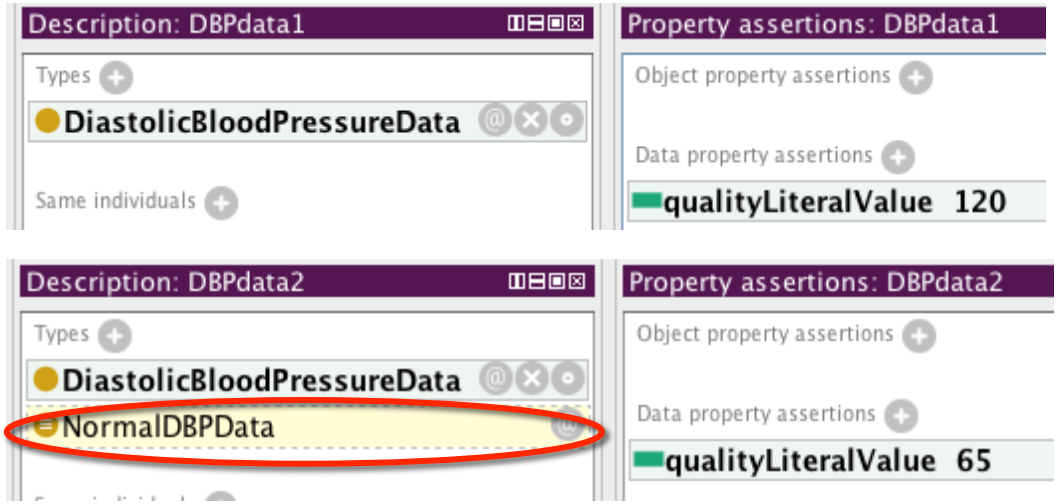




Automatic Classification

Semantic Definition of Normal DBP Data

Two DBP Measurements

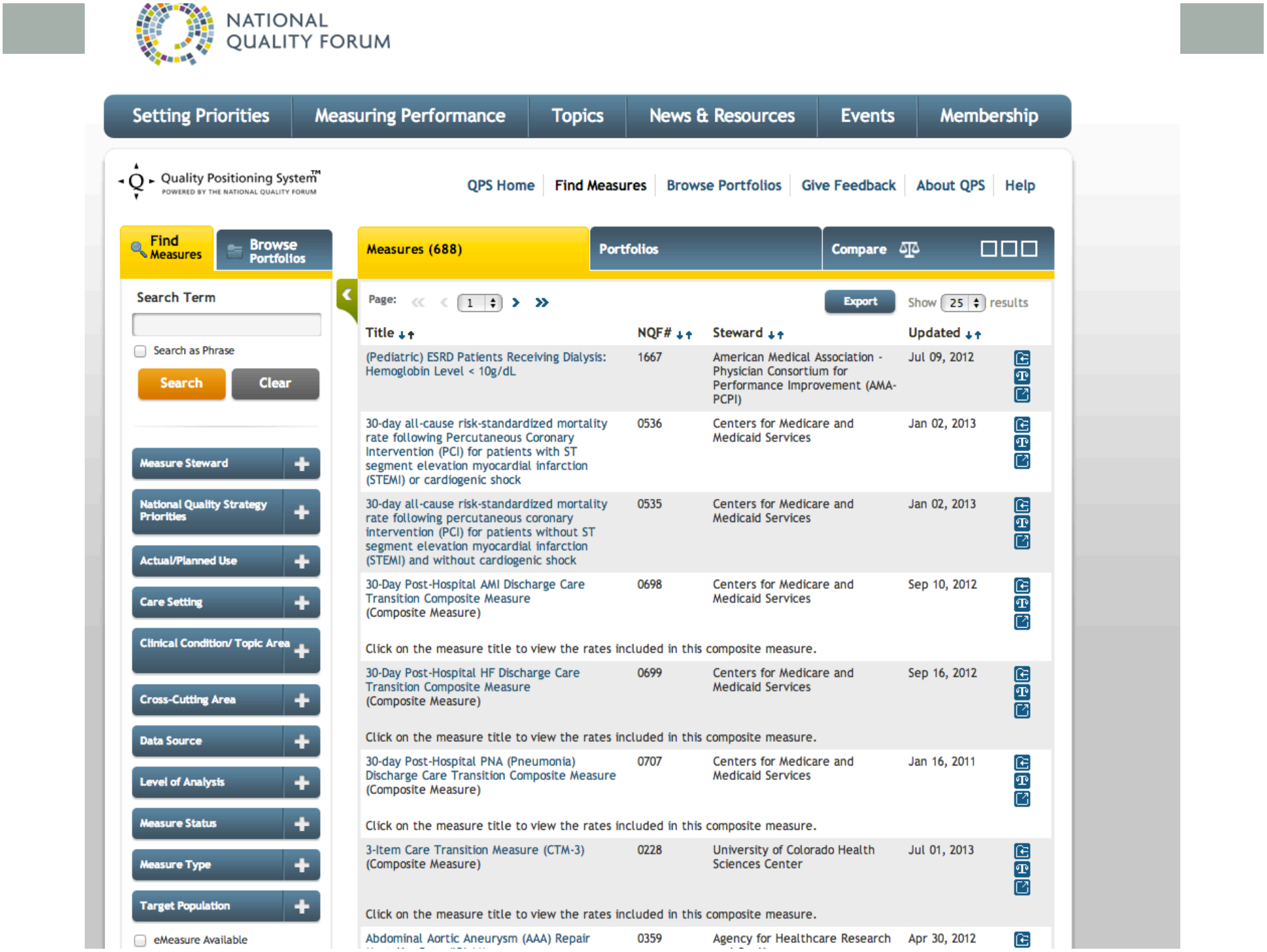


Outline

- Semantic Web representation of EHR data
  - Model representation
  - Instance representation
  - Semantic Reasoning
- Phenotyping application
- Temporal Relation Modeling, Extraction, and Reasoning (TIMER)

Electronic health records (EHRs) driven phenotyping

- Overarching goal
  - To develop automated techniques and algorithms that operate on normalized EHR data to identify cohorts of potentially eligible subjects on the basis of disease, symptoms, or related findings
- National Quality Forum (NQF)
  - More than 600 meaningful use quality measures
  - Quality Data Model (QDM)
    - A structure and grammar to represent quality measures in a standardized format
    - XML-based
    - Not executable


NATIONAL QUALITY FORUM
Setting Priorities
Measuring Performance
Topics
News & Resources
Events
Membership
Quality Positioning System™
POWERED BY THE NATIONAL QUALITY FORUM
QPS Home
Find Measures
Browse Portfolios
Give Feedback
About QPS
Help
Find Measures
Browse Portfolios
Search Term
Search as Phrase
Search
Clear
Measure Steward
National Quality Strategy Priorities
Actual/Planned Use
Care Setting
Clinical Condition/ Topic Area
Cross-Cutting Area
Data Source
Level of Analysis
Measure Status
Measure Type
Target Population
eMeasure Available
Measures (688)
Portfolios
Compare
Page: 1
Export
Show 25 results
Title
NQF#
Steward
Updated
(Pediatric) ESRD Patients Receiving Dialysis: Hemoglobin Level < 10g/dL
1667
American Medical Association - Physician Consortium for Performance Improvement (AMA-PCPI)
Jul 09, 2012
30-day all-cause risk-standardized mortality rate following Percutaneous Coronary Intervention (PCI) for patients with ST segment elevation myocardial infarction (STEMI) or cardiogenic shock
0536
Centers for Medicare and Medicaid Services
Jan 02, 2013
30-day all-cause risk-standardized mortality rate following percutaneous coronary intervention (PCI) for patients without ST segment elevation myocardial infarction (STEMI) and without cardiogenic shock
0535
Centers for Medicare and Medicaid Services
Jan 02, 2013
30-Day Post-Hospital AMI Discharge Care Transition Composite Measure (Composite Measure)
0698
Centers for Medicare and Medicaid Services
Sep 10, 2012
Click on the measure title to view the rates included in this composite measure.
30-Day Post-Hospital HF Discharge Care Transition Composite Measure (Composite Measure)
0699
Centers for Medicare and Medicaid Services
Sep 16, 2012
Click on the measure title to view the rates included in this composite measure.
30-day Post-Hospital PNA (Pneumonia) Discharge Care Transition Composite Measure (Composite Measure)
0707
Centers for Medicare and Medicaid Services
Jan 16, 2011
Click on the measure title to view the rates included in this composite measure.
3-Item Care Transition Measure (CTM-3) (Composite Measure)
0228
University of Colorado Health Sciences Center
Jul 01, 2013
Click on the measure title to view the rates included in this composite measure.
Abdominal Aortic Aneurysm (AAA) Repair
0359
Agency for Healthcare Research
Apr 30, 2012

Phenotype Use Case Example

Resistant Hypertension Phenotyping Algorithm

Has two outpatient (if possible) measurements of Systolic blood pressure>140 or Diastolic blood pressure > 90 at least one month after taking antihypertensive drugs.

Sample Patient Data

Patient was put on Labetalol 100mg starting from March 1st.

Follow-up BP in 5 weeks: SBP=156, DBP=110

Today's BP measurements: SBP=148, DBP=112 (note date: May 1)

Extract and Represent Clinical Elements & Constraints

Resistant Hypertension Phenotyping Algorithm

Has two outpatient (if possible) measurements of Systolic blood pressure>140 or Diastolic blood pressure > 90 at least one month after taking antihypertensive drugs.

Sample Patient Data

Patient was put on Labetalol 100mg starting from March 1st.

Follow-up BP in 5 weeks: SBP=156, DBP=110

Today's BP measurements: SBP=148, DBP=112 (note date: May 1)

Example: Diabetes & Lipid Mgmt. - I

Example: Diabetes & Lipid Mgmt. - I

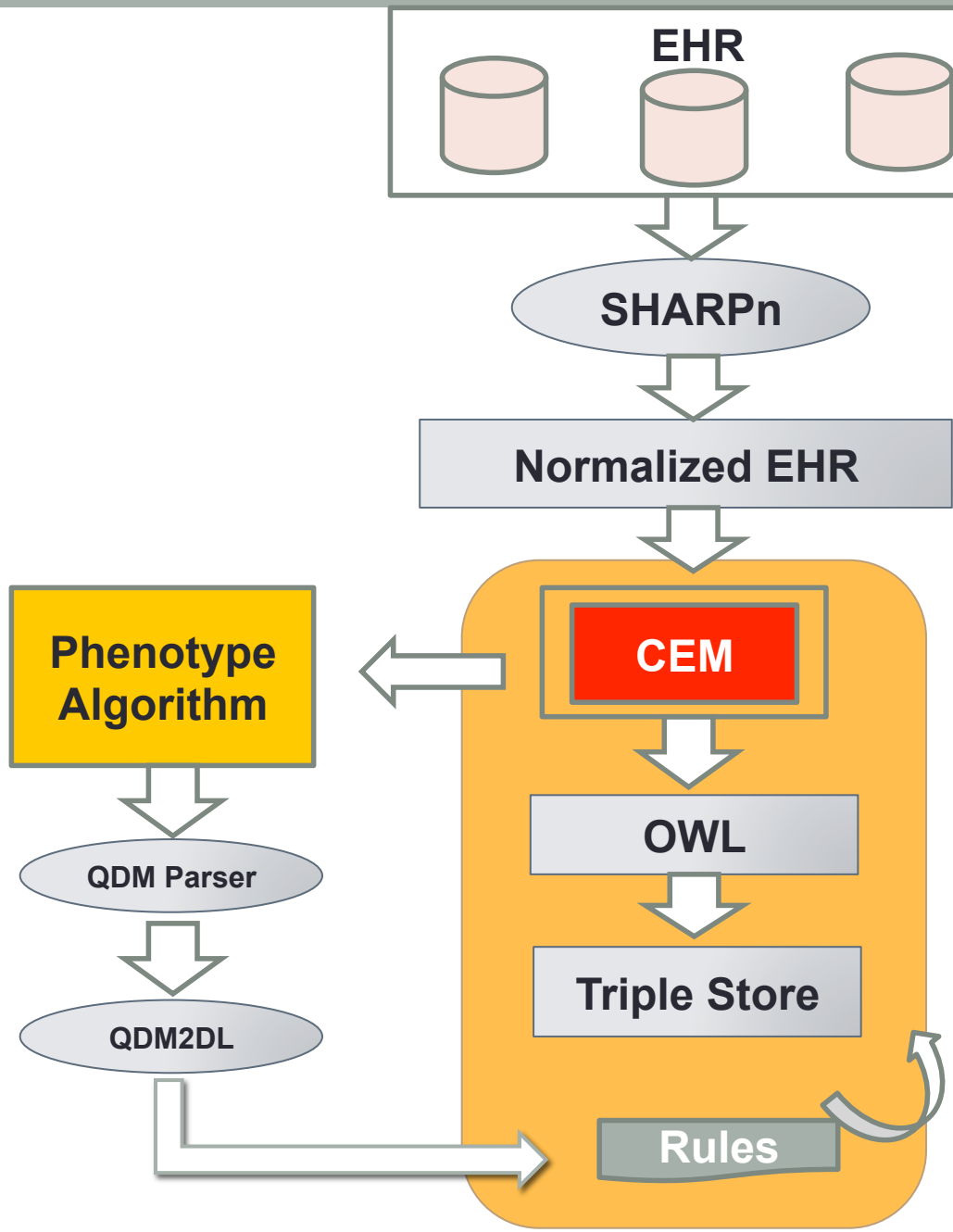
Human readable HTML

Population criteria

- **Initial Patient Population =**
  - AND: "Patient characteristic: birth date" >= 17 year(s) and <= 74 year(s) starts before start of "Measurement period"
- **Denominator=**
  - AND: "Initial Patient Population"
  - AND:
    - OR:
      - AND:
        - OR: "Encounter: Encounter acute inpatient or ED"
        - OR:
          - AND: >= 2 count(s) of
            - AND: "Encounter: Encounter non-acute inpatient and outpatient"
          - AND: FIRST:"Encounter: Encounter non-acute inpatient and outpatient" starts before start of SECOND :"Encounter: Encounter non-acute inpatient and outpatient"
      - AND: "Diagnosis active: diabetes"
    - OR:
      - OR: "Medication order: Medications indicative of diabetes"
      - OR: "Medication dispensed: Medications indicative of diabetes"
      - OR: "Medication active: Medications indicative of diabetes"
    - <= 2 year starts before or during "Measurement end date"

Phenotyping Application

- Make the measures executable on normalized EHR!
- RDF representation of normalized EHR data
  - Demographic information, laboratory records, medication history, diagnosis, encounters, and symptom descriptions
  - RDF triple store
  - Modeled by CEM-OWL
- SWRL/DL representation of the NQF measures
  - UIMA-based QDM parser
  - QDM2DL API



Shen F, Li D, Liu H, Pathak J, Chute CG, Tao C. A SWRL Implementation of NQF Measures for Querying Electronic Healthcare Records in RDF Triples. Submitted to AMIA-CRI, 2013

1. Value Set Restriction
NQF Operation: Encounter: Encounter Office & Outpatient Consult

Query (class expression)

```
PatientExternalId and itemFor some
(Encounter and (EncounterHasCode some
 (0001_EncounterCode)) and EncounterHasCodeSys some
(0001_EncounterCodeSys))
```

2. Logic Operation
NQF Operation:
AND: "Diagnosis, Active: diabetes"
AND:
OR: "Encounter: Encounter acute inpatient or ED"
OR: "Encounter: Encounter non-acute inpatient and outpatient"

Query (class expression)

```
PatientExternalId and itemFor some
(AdministrativeDiagnosis and DiagnosisHasCode some
(0064_DiagnosisCode) and DiagnosisHasCodeSys some
(0064_DiagnosisCodeSys))
AND
(PatientExternalId  and itemFor some
(Encounter and EncounterHasCode some
(0064_EncounterCode) and EncounterHasCodeSys some
(0064_EncounterCodeSys))
OR
(PatientExternalId  and itemFor some
(Encounter and EncounterHasCode some
(0064_EncounterCode) and EncounterHasCodeSys some
(0064_EncounterCodeSys))))
```

3.1. Negation Function
NQF Operation: AND NOT: "Diagnosis, Active: diabetes"

Query (class expression)

```
PatientExternalId and itemFor some
(AdministrativeDiagnosis and DiagnosisHasCode some
 (0056_DiagnosisCode) and DiagnosisHasCodeSys some
(0056_DiagnosisCodeSys))
```

3.2. Count Function
NQF Operation: Count >= 2 of:"Encounter: Encounter Office & Outpatient Consult"

Query (class expression)

```
PatientExternalId and itemFor some (Encounter and Count some
 (integer [>="2" ^^integer]))
```

3.3. Order Function
NQF Operation: AND: FIRST:"Encounter: Encounter non-acute inpatient, outpatient, or ophthalmology" >= 1 day(s) starts before start of SECOND :"Encounter: Encounter non-acute inpatient, outpatient, or ophthalmology"

Encounter(?enc1), hasVisitDate(?enc1,?value), PatientExternalId(?pid), item(?enc1, ?pid), Encounter(?enc2), hasVisitDate(?enc2,?value2), PatientExternalId(?pid), item(?enc2, ?pid), subtractDayTimeDurations(?duration,?value,?value2),lessThanOrEqual(?duration,1)
→
PatientExternalId(?pid)

4.1 Time comparison between variable and constant without time arithmetic

NQF Operation: Encounter: Encounter ambulatory including pediatrics" during "Measurement period"

Measurement begin time = 2012-01-01T00:00:00
Measurement end time = 2012-12-31T59:59:59

Query (class expression)

```
PatientExternalId and itemFor some
(SecondaryUsePatient and item some
(BirthDate and data some
(TS and qualityLiteralValue some
dateTime[<="2010-12-31T59:59:59" ^^dateTime])))
```

4.2. Time comparison between variable and constant with time arithmetic

NQF Operation:
AND: "Patient Characteristic: birth date" >= 2 year(s) starts before start of "Measurement period

Time Stamp = 2012-01-01T00:00:00 – 2 years
= 2010-01-01T00:00:00

Query (class expression)

```
PatientExternalId and itemFor some
(SecondaryUsePatient and item some
(BirthDate and data some
(TS and qualityLiteralValue some
dateTime[<="2010-12-31T59:59:59" ^^dateTime])))
```

4.3. Time comparison between two variables without time arithmetic

NQF Operation: "Diagnosis, Active: Asthma" starts before or during ("Encounter: Encounter Office & Outpatient Consult"

AdministrativeDiagnosis(?xatom), AttribRecordedTime (?xatttime), PatientExternalId(?pid), item(?xatom, ?pid), TS(?ts),att(?xatom,?xatttime),data(?xatttime,?ts),TS(?ts),qualityLiteralValue(?ts,?value),Encounter(?enc), hasVisitDate(?enc,?value2), PatientExternalId(?pid), item(?enc, ?pid),lessThanOrEqual(?value,?value2) -> PatientExternalId(?pid)

4.4. Time comparison between two variables with time arithmetic

NQF Operation: "Medication, Active: pharyngitis antibiotics"
<= 30 day(s) starts before start of ("Encounter: Encounter ambulatory including pediatrics"

SecondaryUseNotedDrug (?xatom), PatientExternalId(?pid), item(?xatom,?pid),StartTime(?stime0),TS(?ts0), data(?stime0,?ts0),qualityLiteralValue(?ts0,?value),Encounter(?enc), hasVisitDate(?enc,?value2), PatientExternalId(?pid), item(?enc, ?pid),subtractDayTimeDurations(?duration,?value2,?value),lessThanOrEqual(?duration,30) →PatientExternalId(?pid)

NQF Human Readable Statement

[1] AND:
- [1.1] OR:
  - [1.1.1] AND: "Diagnosis, Active: diabetes"
  - [1.1.2] AND:
    - [1.1.2.1] OR: "Encounter: Encounter acute inpatient or ED"
    - [1.1.2.2] OR:
      - [1.1.2.2.1] AND: Count >= 2 of:AND: "Encounter: Encounter non-acute inpatient, outpatient, or ophthalmology"
      - [1.1.2.2.2] AND: FIRST:"Encounter: Encounter non-acute inpatient, outpatient, or ophthalmology" >= 1 day(s) starts before start of [1.1.2.2.2.1] SECOND :"Encounter: Encounter non-acute inpatient, outpatient, or ophthalmology"
- [1.2] OR: "Medication, Dispensed: Medications indicative of diabetes"
- [1.3] <= 2 year(s) starts before or during "Measurement end date"

Rule Representation

1-1.1-1.1.1 (OWL-DL)

1-1.1-1.1.2-1.1.2.1-1.3 (OWL-DL)

1-1.1-1.1.2-1.1.2.2-1.1.2.2.1-1.3 (OWL-DL)

1-1.1-1.1.2-1.1.2.2-1.1.2.2.2-1.1.2.2.2.1-1.3 (SWRL)

1-1.2-1.3 (OWL-DL)

```
( PatientExternalId and itemFor some (AdministrativeDiagnosis and
DiagnosisHasCode  some ( 0056_DiagnosisCode) and DiagnosisHasCodeSys
some (0056_DiagnosisCodeSys) and qualityLiteralValue some (dateTime[<=
"2011-12-31T00:00:00"^^dateTime])
and
    (( PatientExternalId and itemFor some (Encounter and EncounterHasCode
    some (0056_EncounterCodeIP) and EncounterHasCodeSys  some
    (0056_EncounterCodeSysIP)) and hasVisitDate some (dateTime[<= "2011-
    12-31T00:00:00"^^dateTime]))
    or
        (( PatientExternalId and itemFor some (Encounter and
        EncounterHasCode  some (0056_EncounterCodeNIP) and
        EncounterHasCodeSys  some (0056_EncounterCodeSysNIP)) and Count
        some (Integer [>="2" ^^Integer]) and hasVisitDate some (dateTime[<=
        "2011-12-31T00:00:00" ^^dateTime]))
        and
        Encounter(?enc1), hasVisitDate(?enc1, ?value), PatientExternalId(?pid),
        item(?enc1,?pid), Encounter(?enc2), hasVisitDate(?enc2, ?value2),
        PatientExternalId (?pid), item(?enc2, ?pid),
        subtractDayTimeDurations(?duration, ?value, ?value2),
        lessThanOrEqual(?duration,1) -> PatientExternalId (?pid)
or
    ( PatientExternalId and itemFor some (SecondaryUseNotedDrug and
    DrugHasCode some (0056_MedicationDispenseCode) and DrugHasCodeSys
    some (0056_MedicationDispenseCodeSys) and (qualityLiteralValue some
    (dateTime[<= "2011-12-31T00:00:00"^^dateTime]))
```

Outline

- Semantic Web representation of EHR data
  - Model representation
  - Instance representation
  - Semantic Reasoning
- Phenotyping application
- Temporal Relation Modeling, Extraction, and Reasoning (TIMER)

Introduction

- Time is essential in clinical research
  - Uncover temporal pattern
    - Disease level
    - Patient level
  - Explain past events
  - Predict future events
- Challenges
  - Vast amount of data
  - Data embedded in narratives
  - Many temporal relations are not *explicitly* stated in the clinical narratives, but rather needs to be inferred

Temporal Relation Reasoning (Example)

- *Patient's INR value is below normal* (Event 1) today. (note date: 01/26/07)
- He has had the *chills and body aches* (Event 2) before *the abnormal test*. (Event 3)" (note date: 01/26/07)
- On Jan. 30, 2007, patient started *Coumadin dosing plan of 1.0 mg* (Event 4).(note date: 02/09/07)
- *Question: "did the patient experience body aches before he started the Coumadin dosing plan?"*

Temporal Relation Reasoning (Example)

- *Patient's INR value is below normal* (Event 1) today. (note date: 01/26/07)
- He has had the *chills and body aches* (Event 2) before *the abnormal test*. (Event 3)" (note date: 01/26/07)
- On Jan. 30, 2007, patient started *Coumadin dosing plan of 1.0 mg* (Event 4).(note date: 02/09/07)
- *Question: "did the patient experience body aches before he started the Coumadin dosing plan?"*

Event1 = Event3

Temporal Relation Reasoning (Example)

- *Patient's INR value is below normal* (Event 1) today. (note date: 01/26/07)
- He has had the *chills and body aches* (Event 2) before *the abnormal test*. (Event 3)" (note date: 01/26/07)
- On Jan. 30, 2007, patient started *Coumadin dosing plan of 1.0 mg* (Event 4).(note date: 02/09/07)
- *Question: "did the patient experience body aches before he started the Coumadin dosing plan?"*

Event1 = Event3
+
Event2 before Event3
→
Event2 before Event1

Temporal Relation Reasoning (Example)

- *Patient's INR value is below normal* (Event 1) today. (note date: 01/26/07)
- He has had the *chills and body aches* (Event 2) before *the abnormal test*. (Event 3)" (note date: 01/26/07)
- On Jan. 30, 2007, patient started *Coumadin dosing plan of 1.0 mg* (Event 4).(note date: 02/09/07)
- *Question: "did the patient experience body aches before he started the Coumadin dosing plan?"*

Event1 = Event3
+
Event2 before Event3
→
Event2 before Event1

Event1 01/26/07 +
Event4 01/30/07
→
Event1 before Event4

Temporal Relation Reasoning (Example)

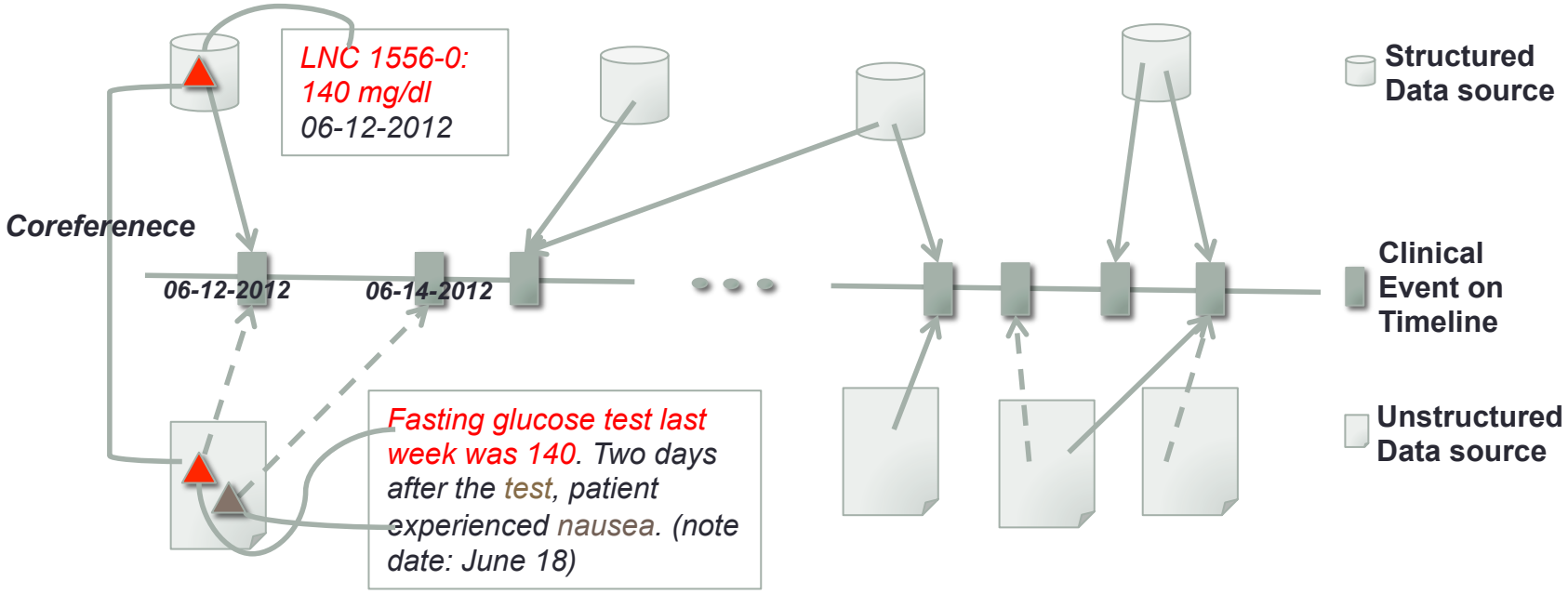
- *Patient's INR value is below normal* (Event 1) today. (note date: 01/26/07)
- He has had the *chills and body aches* (Event 2) before *the abnormal test*. (Event 3)" (note date: 01/26/07)
- On Jan. 30, 2007, patient started *Coumadin dosing plan of 1.0 mg* (Event 4).(note date: 02/09/07)
- *Question: "did the patient experience body aches before he started the Coumadin dosing plan?"*

Event1 = Event3
+
Event2 before Event3
→
Event2 before Event1

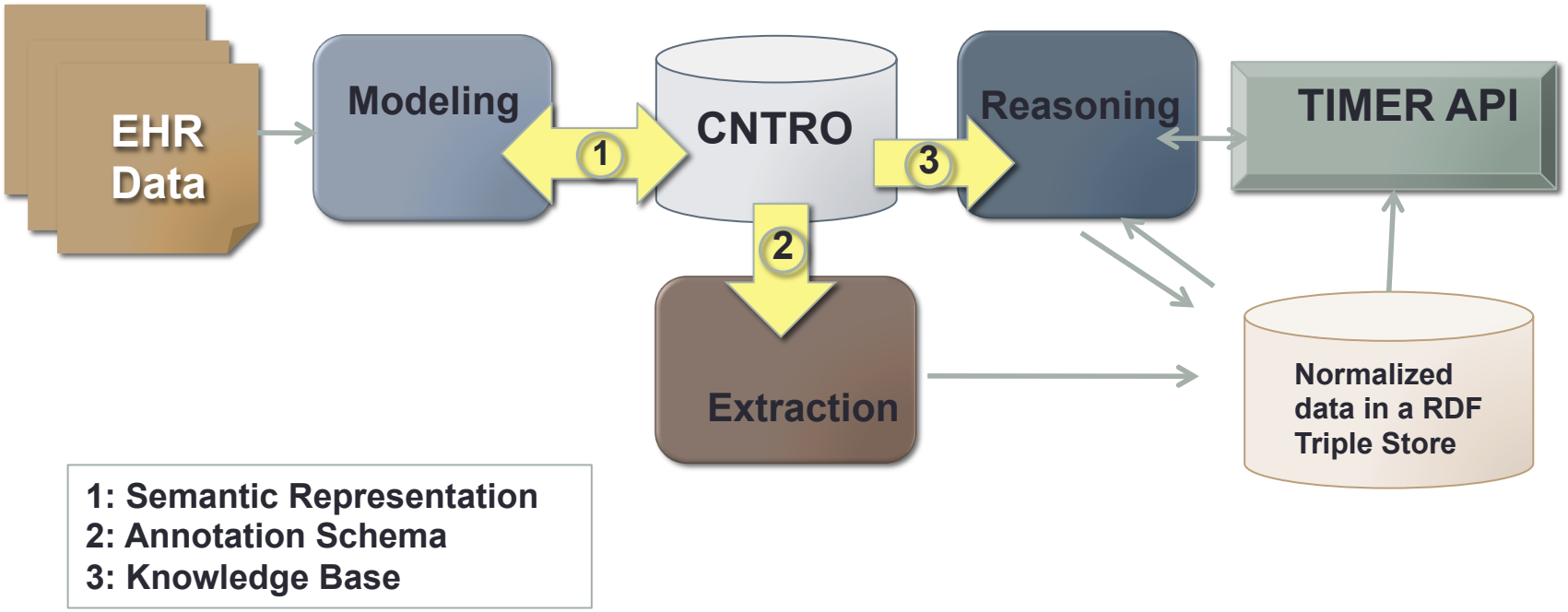
Event1 01/26/07 +
Event4 01/30/07
→
Event1 before Event4

Event2 before Event4

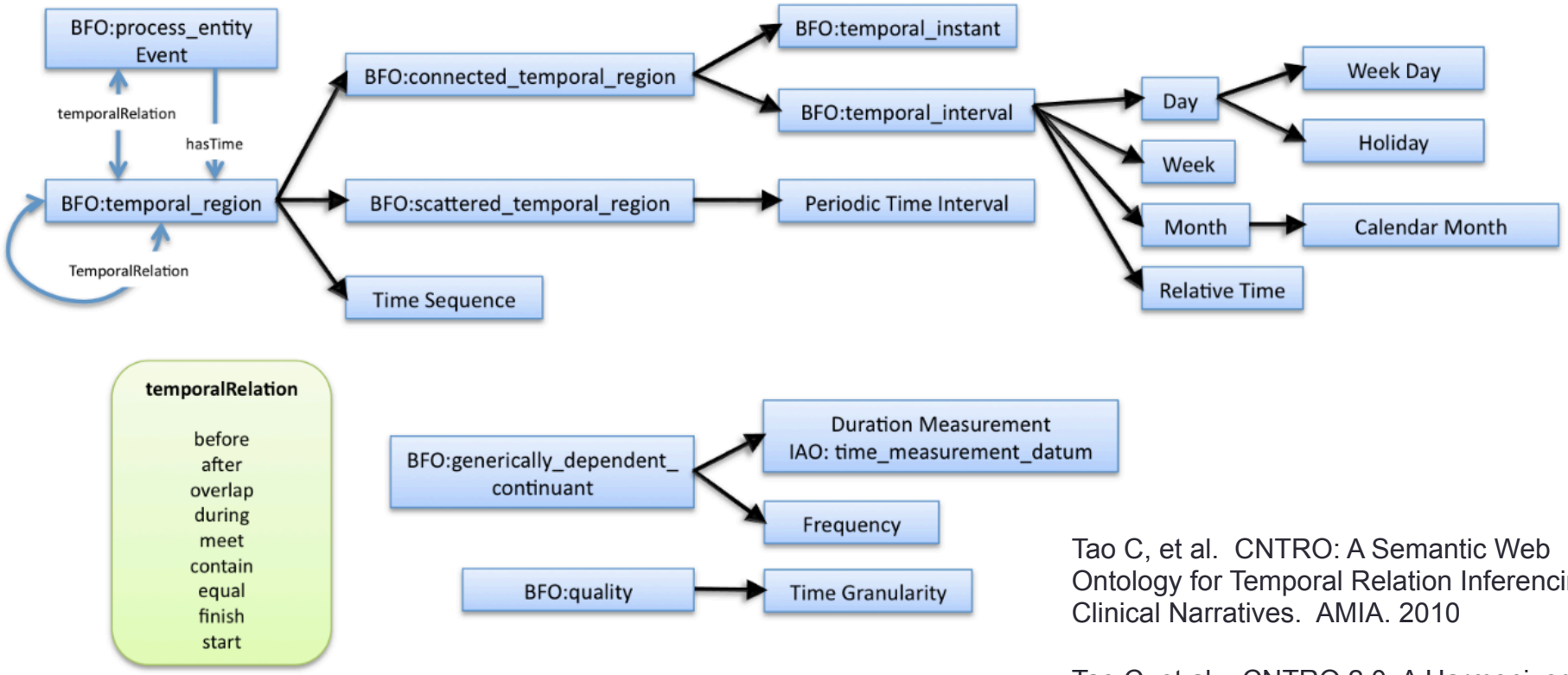
Temporal Reasoning Example



TIMER (Temporal Information Modeling, Extraction, & Reasoning)



CNTRO
Clinical Narrative Temporal Relation Ontology

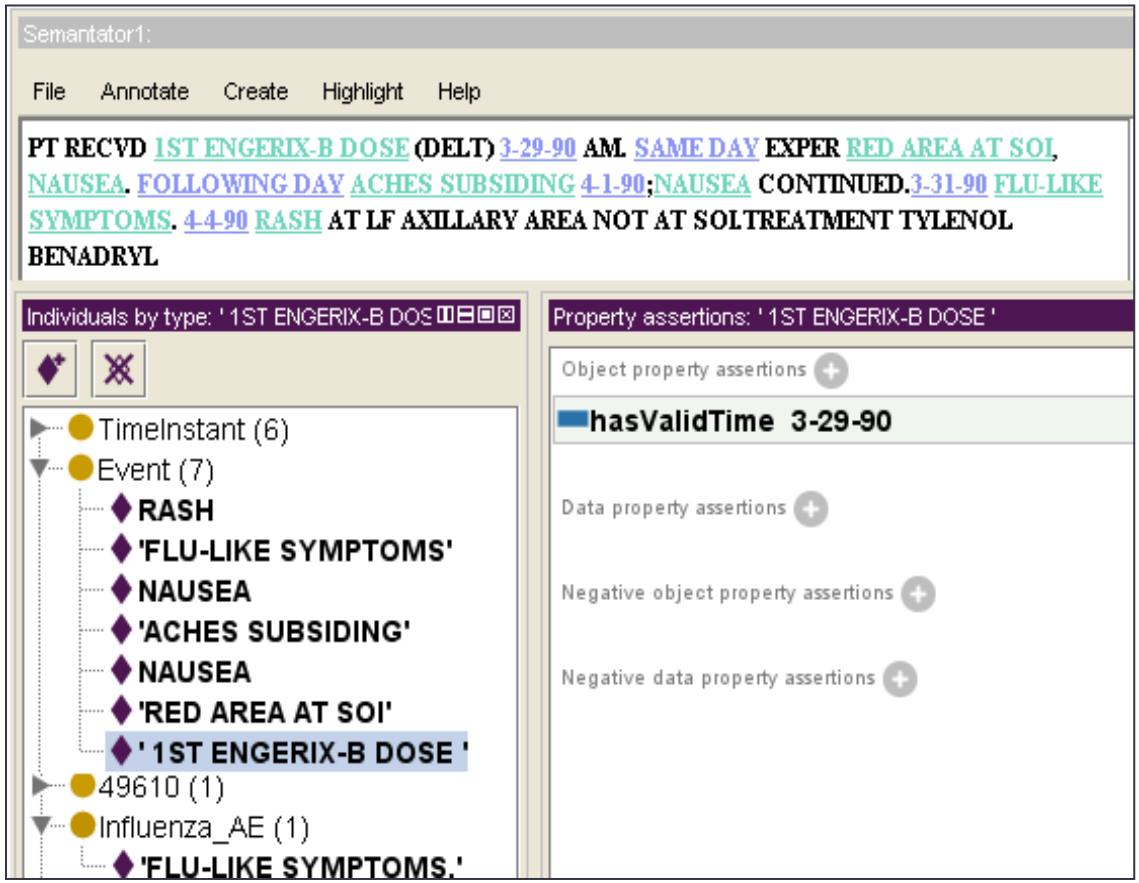


Available at BioPortal

Tao C, et al. CNTRO: A Semantic Web Ontology for Temporal Relation Inferencing in Clinical Narratives. AMIA. 2010

Tao C, et al.. CNTRO 2.0: A Harmonized Semantic Web Ontology for Temporal Relation Inferencing in Clinical Narratives. AMIA-CRI. 2011

Semantator: Information Extraction & Semantic Annotation



Song D, Chute CG, Tao C. Semantator: annotating clinical narratives with semantic web ontologies. AMIA-CRI. 2012

Tao C, et al. Semantic Annotator for Converting Biomedical Text to Linked Data, submitted to Journal of Biomedical Informatics

Information Extraction & Semantic Annotation

Semantator:

- A GUI for users to browse, query, & edit annotated results in the original context
- Manual annotation
- Semi-automatic annotation
- Reasoning: consistency checking
- Inter-annotator agreements

http://informatics.mayo.edu/CNTRO/index.php/Semantator

Temporal Relation Reasoning

- Temporal Representation Normalization
- OWL DL Reasoning
- SWRL-based Reasoning

Tao C, et al. Time-Oriented Question Answering from Clinical Narratives Using Semantic-Web Techniques, ISWC 2010

Implementation Status

- *findEvent(searchText)*
  - returns a list of events that match the searching criteria.
- *GetEventFeature(event, featureflag)*
  - returns a specific time feature for a given event.
  - Sample query:
    - When was the patient diagnosed with diabetes?
    - When did the patient start his chemotherapy?

Implementation Status

- *getDurantionBetweenEvents(event1, event2)*
  - returns the time interval between two events.
  - Sample query: How long after the patient was diagnosed colon cancer did he start the chemotherapy?
- *getDuration(event)*
  - returns the duration of a given event.
  - Sample query: How long did the symptoms of rectal bleeding last?

Implementation Status

- *getTemporalRelationType(event1, event2)*
  - returns the temporal relations between two events if it can be retrieved or inferred.
  - Sample query: Was the CT scan after the colonoscopy?
- *getTemporalRelationType(event1, time)*
  - returns the temporal relations between an event and a specific time if it can be inferred or retrieved.
  - Sample query: Is there any behavior change within a week of the test?

Implementation Status

sortEventsByTemporalRelationsOrTimeline(events)

- returns the order (timeline) of a set of events.
- sample query:
  - What is the tumor status timeline as indicated in the patient's radiology note?
  - What is the treatment timeline as recorded in oncology notes?
  - When was the first colonoscopy done?
  - When was the most recent glucose test?

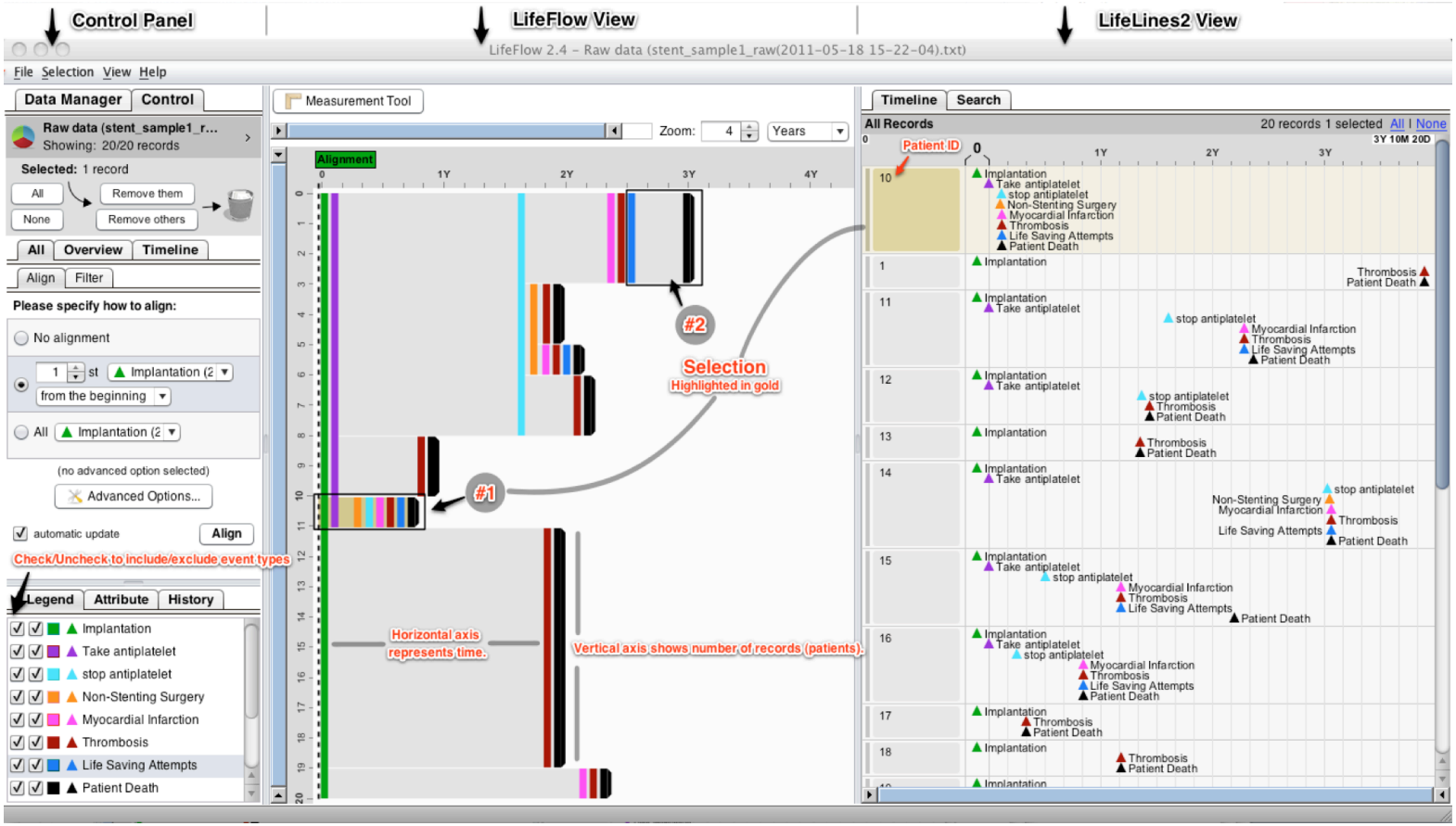
TIMER Application

- Late stent thrombosis (LST) adverse event
- Complaint files from Manufacturer and User Facility Device Experience (MAUDE) database
- Detect potential temporal patterns within complaint files of similar adverse events

Clark KK, Sharma DK, Chute CG, Tao C. Application of a temporal reasoning framework tool in analysis of medical device adverse events. AMIA 2011

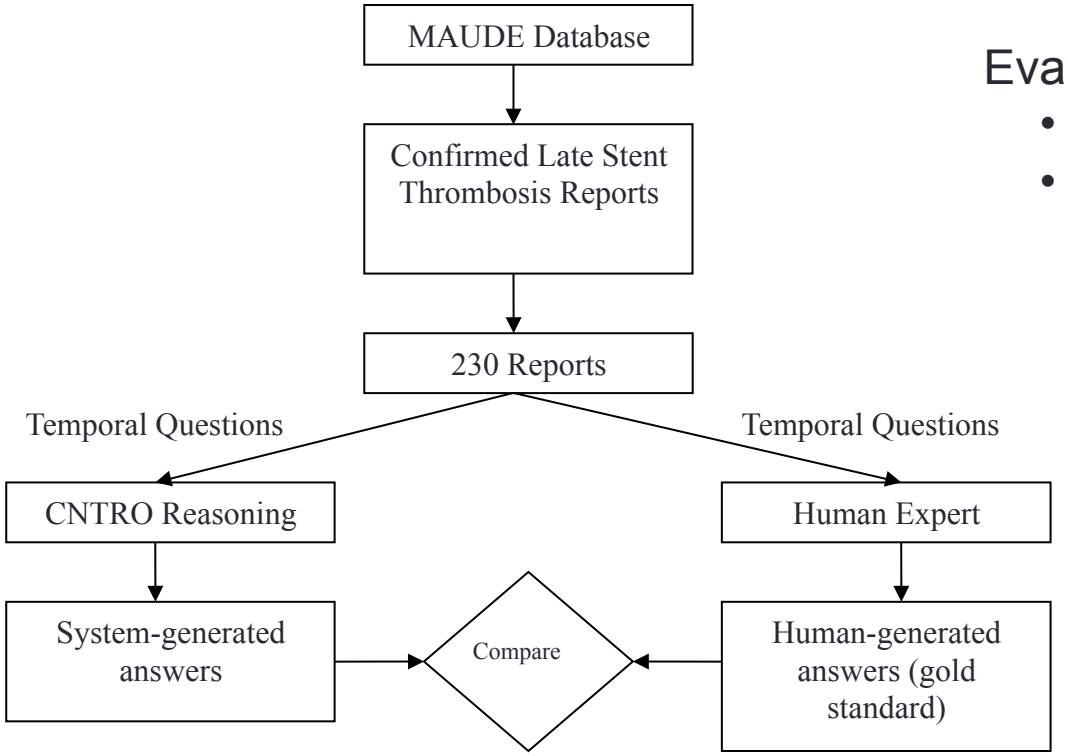
Clark KK, Sharma D, Qin R, Chute CG and Tao C. Ontology-based temporal analysis for medical device adverse event— a use case study on Late Stent Thrombosis. SWAT4LS 2013

Visualization: Connect with LifeFlow



Tao C, Wongsuphasawat K, Plaisant C, Shneiderman B, et al. Towards event sequence representation, reasoning and visualization for EHR data. IHI'12 - Proceedings of the 2nd ACM SIGHIT IHI. 2012

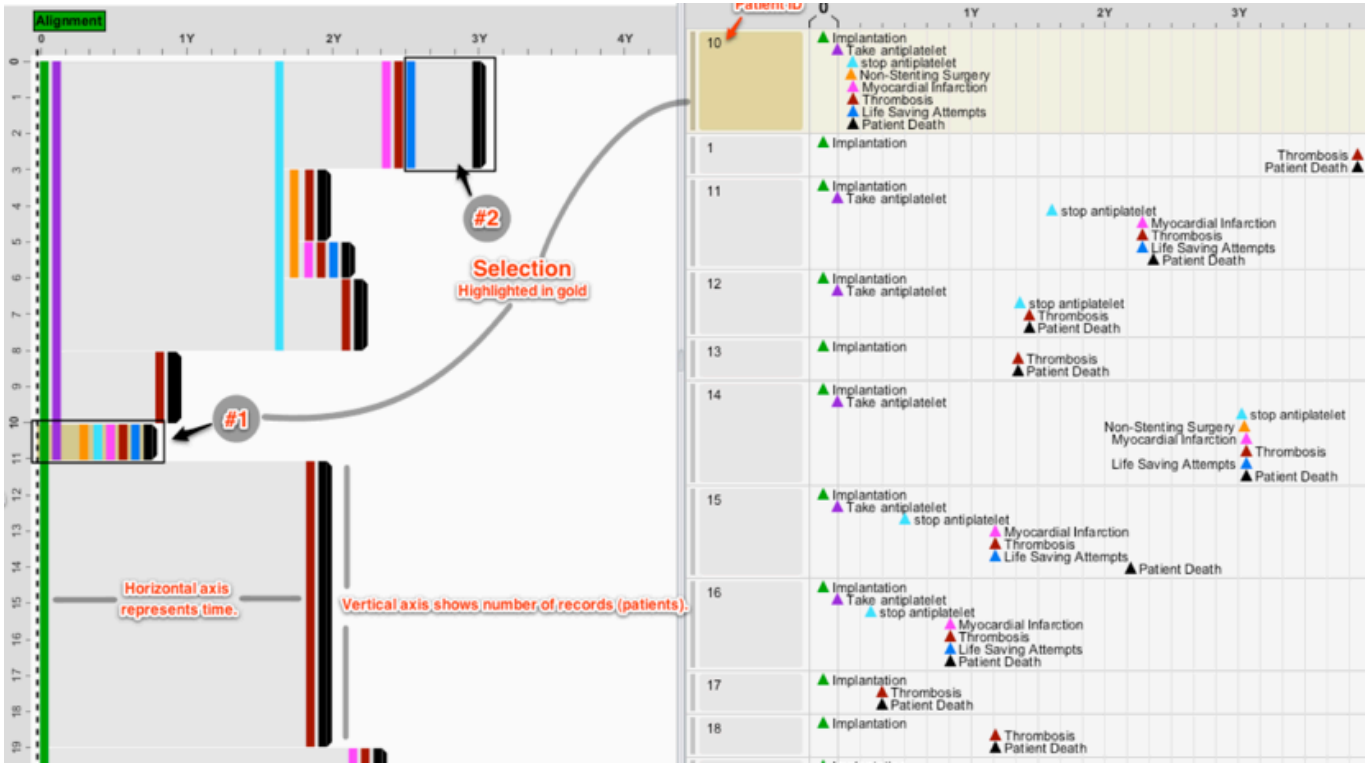
TIMER Application



Evaluation Results:

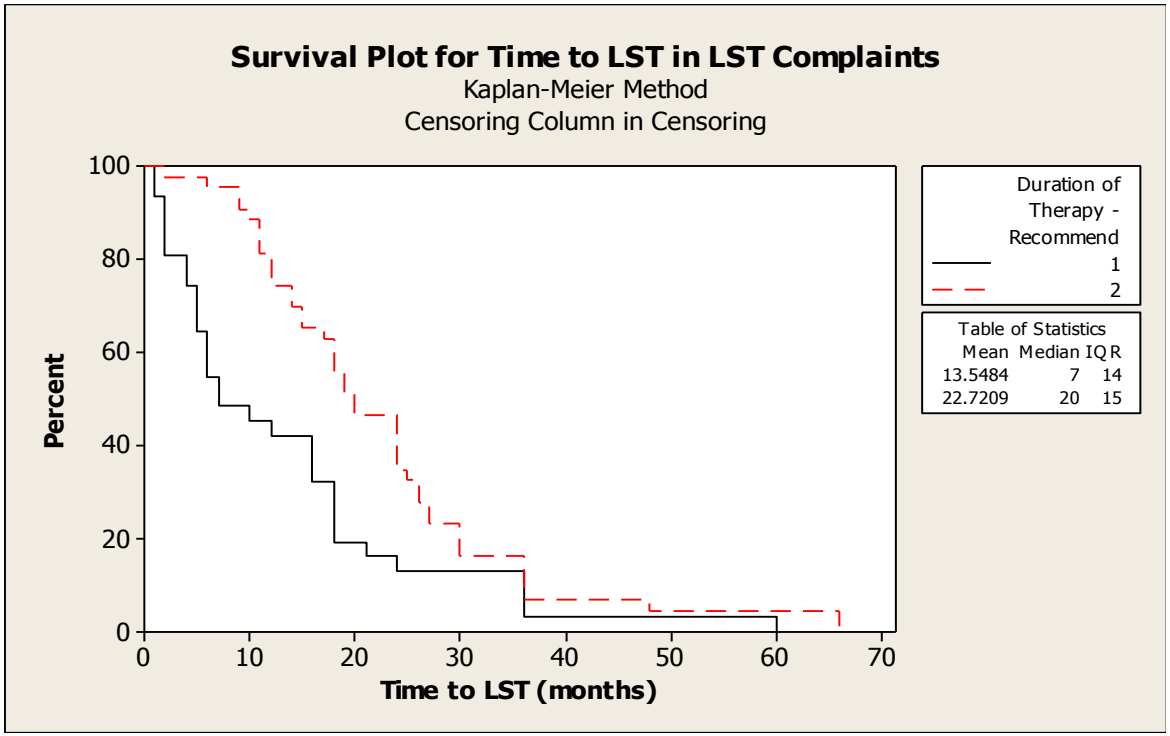
- Timeline: 87.4%
- Duration calculation: 89%

TIMER Application: Temporal Pattern



Event sequence pattern matches previously identified common temporal patterns within LST AEs.

TIMER Application: Statistical Analysis



Survival analysis of shorter duration of antiplatelet therapy (group 1) and longer duration of antiplatelet therapy (group 2) in late stent thrombosis complaints

- Consistent with other relevant studies
- Guidance on Antiplatelet Therapy After Stenting is 12 months
- Early discontinuation could associate with significantly higher rates of LST

Summary

- Ontologies and semantic web technologies can provide a viable and interoperable solution for
  - Modeling of clinical data
  - Conducting scalable querying over the data
  - Inferring new knowledge
- Supports clinical and translational research
  - Decision support
  - Phenotyping
  - Biomedical data network analysis


WHO
WHAT
WHERE
WHEN
WHY
HOW
QUESTIONS
ANSWERS