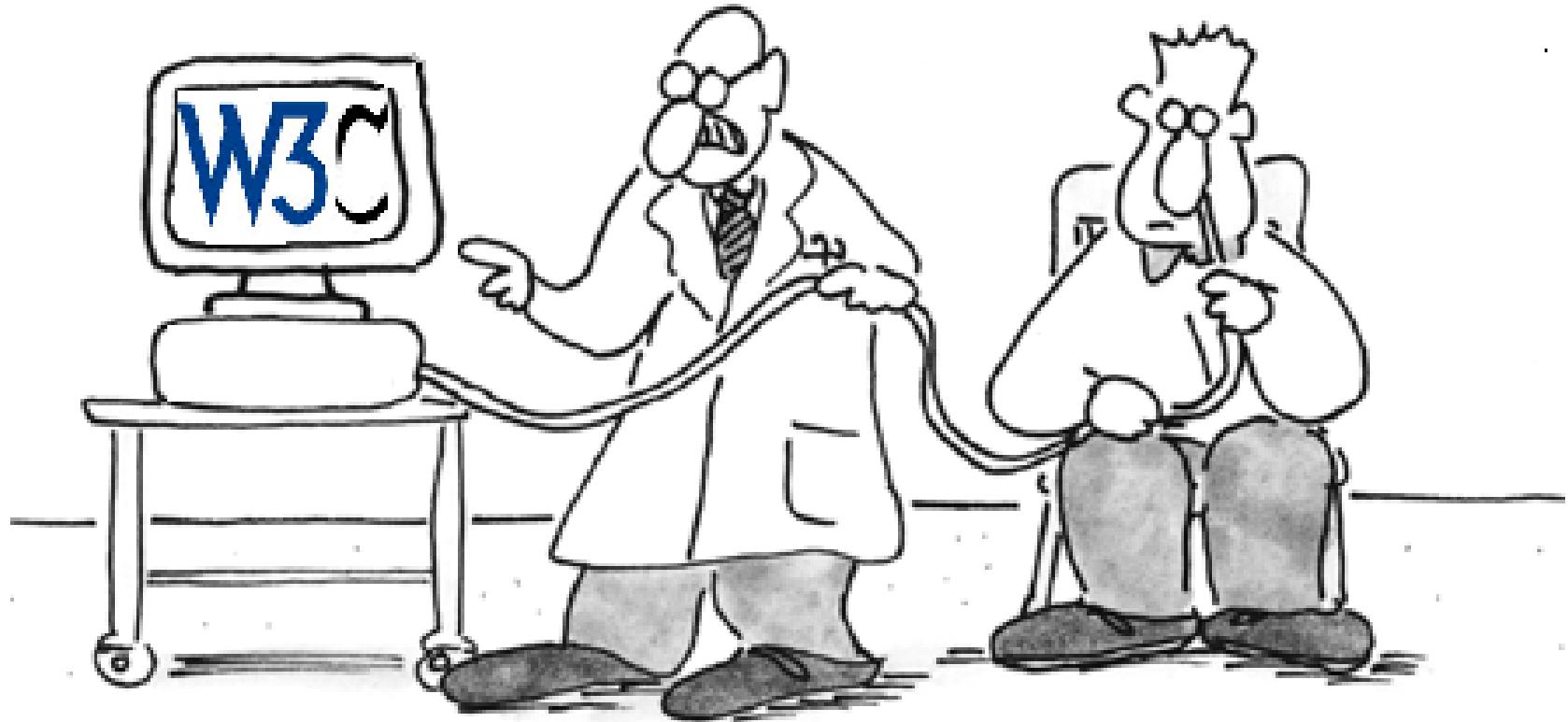


A Semantic Web Approach to Integrative Biosurveillance

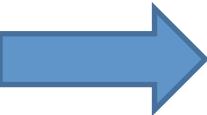


Narendra Kunapareddy, UTHSC
Zhe Wu, Ph.D., Oracle

This talk:

- Translational BioInformatics and Information Integration Dilemma
- Case Study: Public Health Preparedness
- Our Vision
- Our Implementation
- Challenges

This talk:

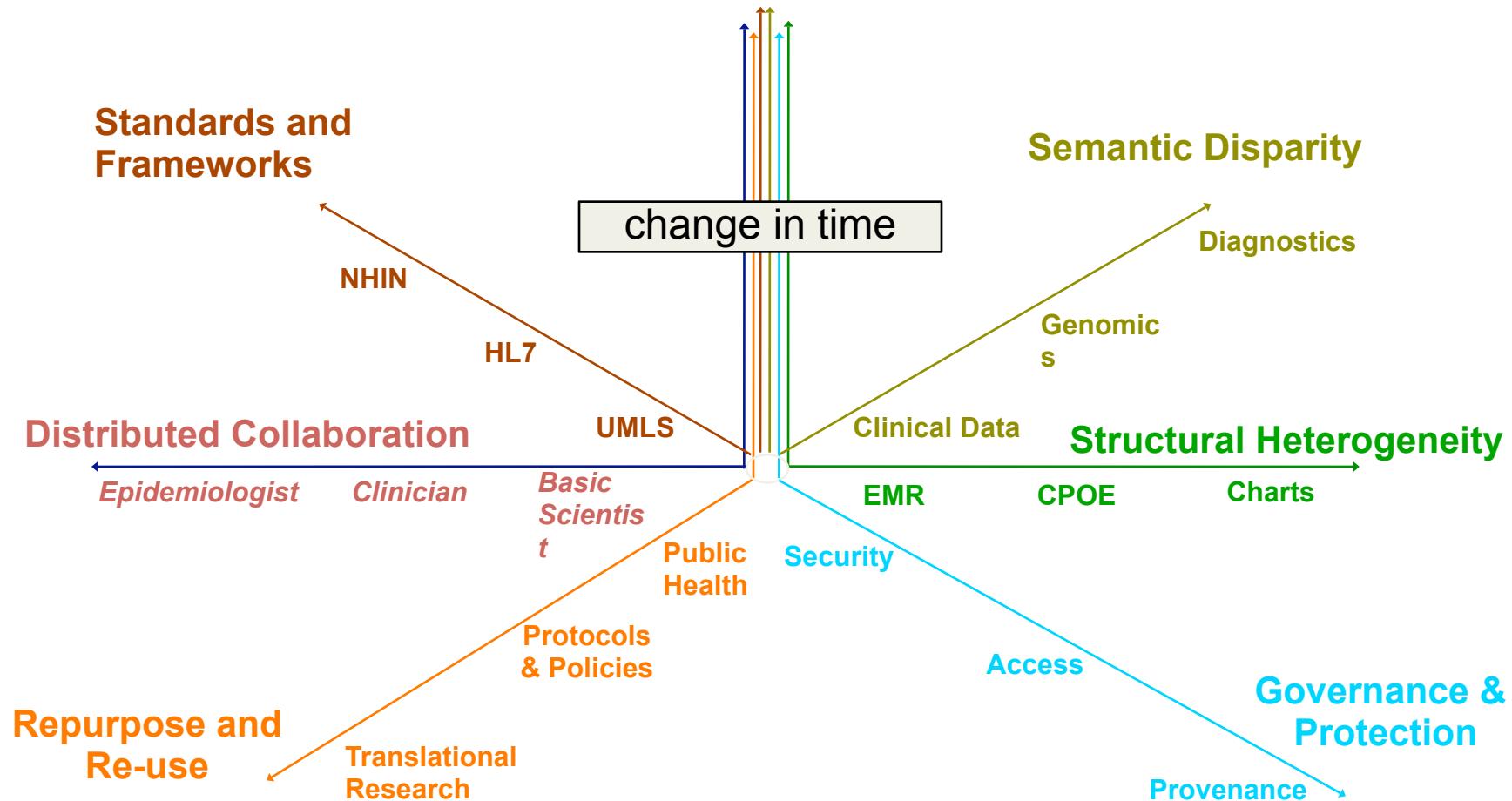


Translational BioInformatics and Information Integration Dilemma

- Case Study: Public Health Preparedness
- Our Vision
- Our Implementation
- Challenges

Information Integration Dilemma

Schema Change, Semantic Drift, Framework Update
Dynamic Environment, New Hypothesis, Governance Planning



This talk:

- Translational Bioinformatics and Information Integration Dilemma



- Case Study: Public Health Preparedness

- Our Vision
- Our Implementation
- Challenges

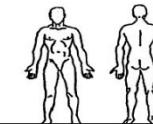
Public Health Preparedness



MedicalRecords 5.1

INSALUD ÁREA 11		PARTE INTERCONSULTA	
NOMBRE Y APELLICIOS DOMICILIO TELÉFONO		SEÑO APELLICIOS TELÉFONO	
Nº APLICACIÓN S.S.		Nº APLICACIÓN S.S.	
CENTRO DE ATENCIÓN PRIMARIA CÓDIGO		CENTRO DE ATENCIÓN PRIMARIA CÓDIGO	
DIA DE LA CONSULTA MES AÑO		HORA SALA CONSULTA NUEVA SUSCIAL	
RMI PREFERENCIA		RMI PREFERENCIA	
OBSERVACIONES		OBSERVACIONES	
MOTIVO DE CONSULTA		MOTIVO DE CONSULTA	
RESUMEN DE SINTOMATOLOGÍA, EXPLORACIÓN Y PRUEBAS, IMPRESIÓN DIAGNÓSTICA Y TRATAMIENTO PREVIO			
Riego sanguíneo de la fractura de la muñeca. Fractura de la muñeca. No se siente dolor cuando se dobla la muñeca.			
JUICIO CLÍNICO		JUICIO CLÍNICO	
INFORME DEL ESPECIALISTA			
RESUMEN EXPLORACIÓN, PRUEBAS COMPLEMENTARIAS, DIAGNÓSTICO Y TRATAMIENTO			
3 MAR 2003 a. fract. distal de la muñeca. Atendido			

MEDICAL RECORD - NURSING HISTORY AND ASSESSMENT			
18. Additional Assessment Data.			
Admission:	TPR	EP	WT
WHT			
ADMITTED VIA: <input type="checkbox"/> Ambulatory <input type="checkbox"/> Wheelchair <input type="checkbox"/> Other ACCOMPANIED BY:			
GENERAL APPEARANCE:			
DISPERSIVE APPARENCE: CHANGE IN PAST 3 MONTHS:			
MENTAL STATUS:			
COPING BEHAVIORS:			
EMERGENCY NEEDS:			
<input type="checkbox"/> No Problem <input type="checkbox"/> Isolation <input type="checkbox"/> Direct Observation <input type="checkbox"/> At Risk for Falls: <input type="checkbox"/> Yes <input type="checkbox"/> No Refer to RAMC OP 693 on DA 4827 <input type="checkbox"/> At Risk for Skin Breakdown: <input type="checkbox"/> Yes <input type="checkbox"/> No Refer to RAMC OP 634 (PCP 710) PATIENT/SIGNIFICANT OTHER REVIEWED FOR MEDICATIONS: <input type="checkbox"/> Yes <input type="checkbox"/> No PATIENT/SIGNIFICANT OTHER REVIEWED FOR WARD RULES: <input type="checkbox"/> Yes <input type="checkbox"/> No			



Medical Records

General		Doctors	Medication	History																																																							
<table border="1"> <tr> <td colspan="2">Personal Information</td> <td colspan="2">Known allergies</td> <td>Notes</td> </tr> <tr> <td colspan="2"> Frank Smith. Name 543 SW Frewel, Jackson, WI, 81474 Address 657-897-5485 Telephone 745-85-8743 Social Security Number Married 06-23-1945 Birth date Marital Status </td> <td colspan="2"> a) Trees b) Grass </td> <td>a) b)</td> </tr> <tr> <td colspan="2"></td> <td>6' 1"</td> <td>210</td> <td>120/98 123</td> </tr> <tr> <td colspan="2"></td> <td>Height</td> <td>Weight</td> <td>Blood Pressure Cholesterol Blood</td> </tr> <tr> <td colspan="2">Medical Insurance</td> <td>03/31/1974</td> <td colspan="2">Effective Date</td> </tr> <tr> <td colspan="2"> Etna Medical. Name 122 Main, Carlsbad, CA 98009 Address 876-323-1234 Group H4J76 HJ-L-35 Phone Group Name Policy Name </td> <td>533 SW Grant, Portland OR 75646</td> <td colspan="2">Dental Insurance</td> </tr> <tr> <td colspan="2"></td> <td>Address</td> <td>03/31/1974</td> <td>Effective Date</td> </tr> <tr> <td colspan="2"></td> <td>Dental Dental Name</td> <td>Dental Dental Name</td> <td></td> </tr> <tr> <td colspan="2"></td> <td>800-323-1234</td> <td>Dental Group Phone</td> <td>111-654, Policy Na</td> </tr> <tr> <td colspan="5">In case of an emergency</td> </tr> <tr> <td colspan="5"> Name Address Phone Ralph Gaylor 122 Main, Carlsbad, CA 98009 503-555-6484 J. Write 533 SW Grant, Portland OR 75646 503-555-6485 Dr. Doris Smith 122 Main, Carlsbad, CA 98009 503-555-6487 J. Write 533 SW Grant, Portland OR 75646 503-555-6489 </td> </tr> </table>					Personal Information		Known allergies		Notes	Frank Smith. Name 543 SW Frewel, Jackson, WI, 81474 Address 657-897-5485 Telephone 745-85-8743 Social Security Number Married 06-23-1945 Birth date Marital Status		a) Trees b) Grass		a) b)			6' 1"	210	120/98 123			Height	Weight	Blood Pressure Cholesterol Blood	Medical Insurance		03/31/1974	Effective Date		Etna Medical. Name 122 Main, Carlsbad, CA 98009 Address 876-323-1234 Group H4J76 HJ-L-35 Phone Group Name Policy Name		533 SW Grant, Portland OR 75646	Dental Insurance				Address	03/31/1974	Effective Date			Dental Dental Name	Dental Dental Name				800-323-1234	Dental Group Phone	111-654, Policy Na	In case of an emergency					Name Address Phone Ralph Gaylor 122 Main, Carlsbad, CA 98009 503-555-6484 J. Write 533 SW Grant, Portland OR 75646 503-555-6485 Dr. Doris Smith 122 Main, Carlsbad, CA 98009 503-555-6487 J. Write 533 SW Grant, Portland OR 75646 503-555-6489				
Personal Information		Known allergies		Notes																																																							
Frank Smith. Name 543 SW Frewel, Jackson, WI, 81474 Address 657-897-5485 Telephone 745-85-8743 Social Security Number Married 06-23-1945 Birth date Marital Status		a) Trees b) Grass		a) b)																																																							
		6' 1"	210	120/98 123																																																							
		Height	Weight	Blood Pressure Cholesterol Blood																																																							
Medical Insurance		03/31/1974	Effective Date																																																								
Etna Medical. Name 122 Main, Carlsbad, CA 98009 Address 876-323-1234 Group H4J76 HJ-L-35 Phone Group Name Policy Name		533 SW Grant, Portland OR 75646	Dental Insurance																																																								
		Address	03/31/1974	Effective Date																																																							
		Dental Dental Name	Dental Dental Name																																																								
		800-323-1234	Dental Group Phone	111-654, Policy Na																																																							
In case of an emergency																																																											
Name Address Phone Ralph Gaylor 122 Main, Carlsbad, CA 98009 503-555-6484 J. Write 533 SW Grant, Portland OR 75646 503-555-6485 Dr. Doris Smith 122 Main, Carlsbad, CA 98009 503-555-6487 J. Write 533 SW Grant, Portland OR 75646 503-555-6489																																																											

Registered To: Unregistered

◀◀ 1 of 1 ▶▶

New Delete

100 Browse

```

<word_entry aspect='ned'>
<headword_lemma>bavit</headword_lemma>
<frame_entry frame_index='1'>
<gloss>poskytovat zábavu</gloss>
<example>bavit přátele svým vyprávěním</example>
<frame_slots>
<slot functor='ACT' type='obl'>
<form type='direct_case' case='1'></form>
</slot>
<slot functor='PAT' type='obl'>
<form type='direct_case' case='4'></form>
</slot>
<slot functor='MEANS' type='typ'>
<form type='direct_case' case='7'></form>
</slot>
</frame_slots>
</frame_entry>

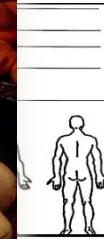
<frame_entry frame_index='2'>
<gloss>zajimati</gloss>
<example>to mě bavit; ty mě nebavíš</example>
<control>ACT</control>
<frame_slots>
<slot functor='ACT' type='obl'>
<form type='direct_case' case='4'></form>
</slot>
<slot functor='PAT' type='obl'>
<form type='direct_case' case='1'></form>
<form type='infinitive'></form>
<form type='subord_conj' subord_conj_lemma='že'></form>
</slot>
</frame_slots>
</frame_entry>
</word_entry>

```

Public Health Preparedness



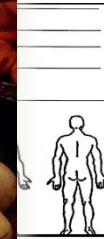
INSALUD ÁREA 11		PARTE INTERCONSULTA		
NOMBRE Y APELLIDOS		SEXO	M. HISTÓRICA/C.U.	
TELÉFONO			AP.	
Nº AFILIACIÓN S.S.		A.E.		
CENTRO DE ATENCIÓN PRIMARIA		CÓDIGO:	F. FIRMAS	
			D.R.O.	



Public Health Preparedness



INSALUD ÁREA 11		PARTE INTERCONSULTA	
NOMBRE Y APELLICIOS DOMICILIO TELÉFONO		SEÑOR _____ A.P. _____ C.P. _____ CALLE _____ N.º _____ COL. _____ C.P. _____ FAX _____	
N.º APLICACIÓN S.S.		DE MEDICAMENTO A.P. _____ A.E. _____	
CENTRO DE ATENCIÓN PRIMARIA		CÓDIGO	
		C.I.A. _____	
		C.I.A. _____	



SNOMED

NNDS

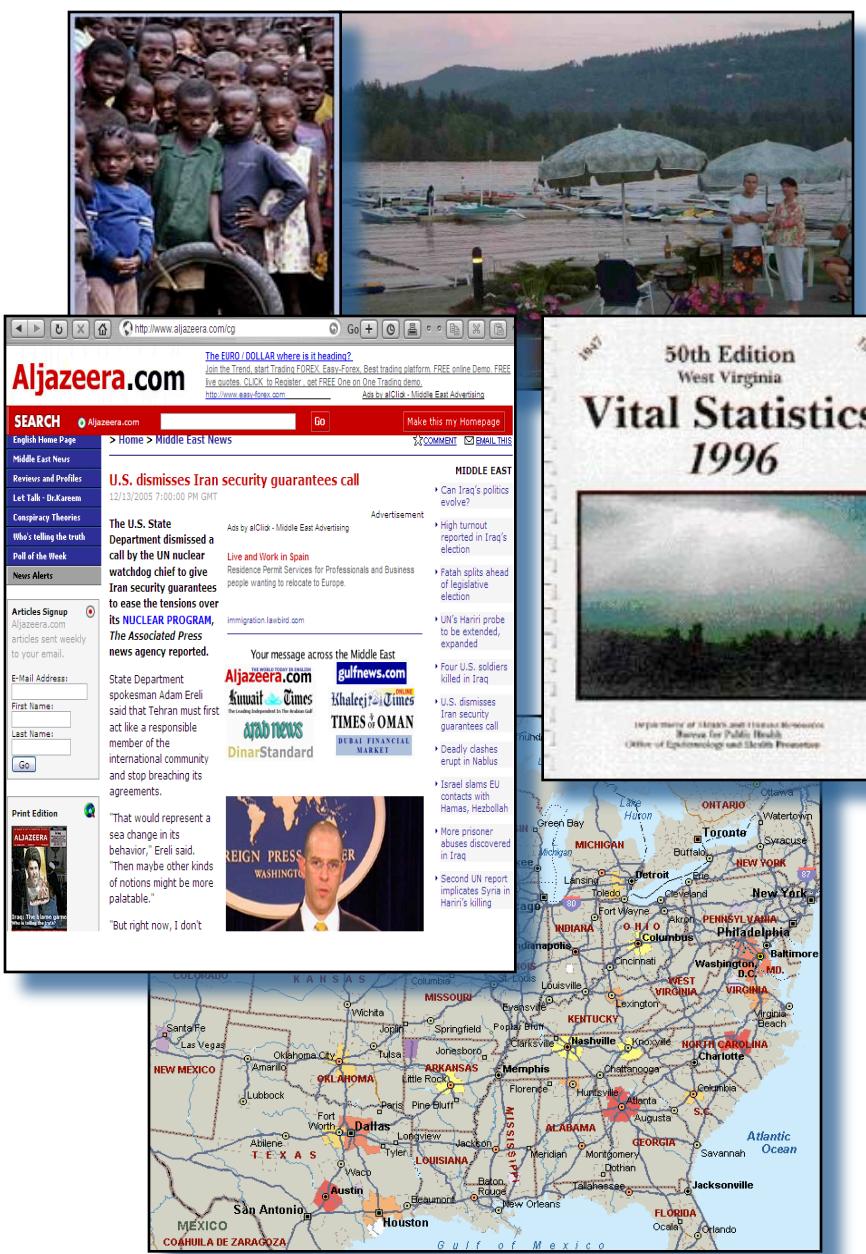


LOINC

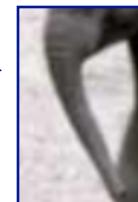
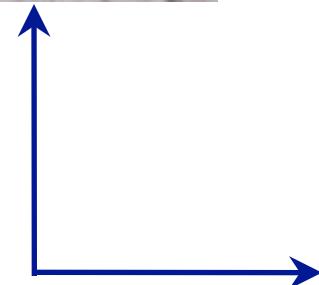
PHIN



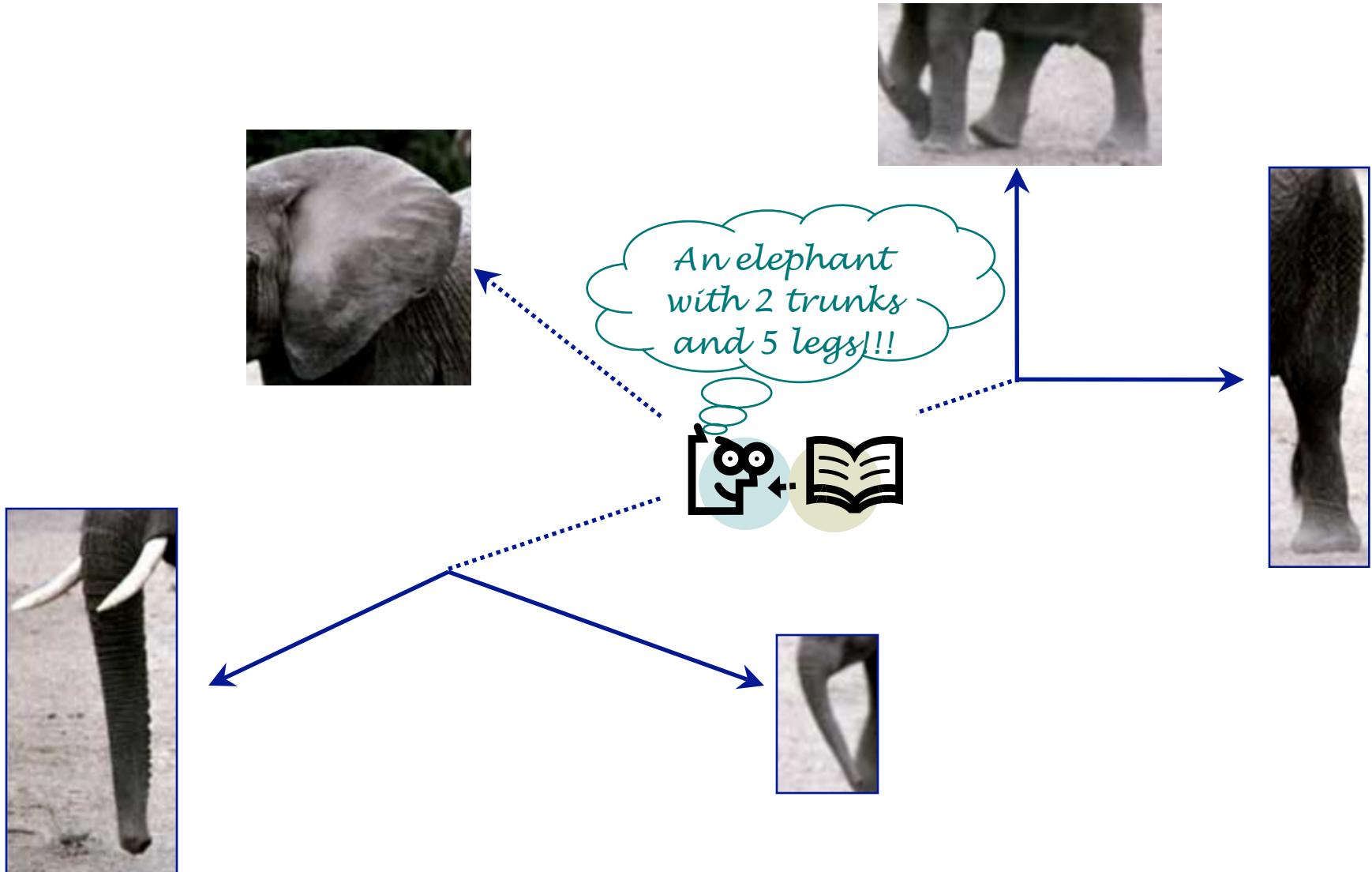
Context is important



State of the art



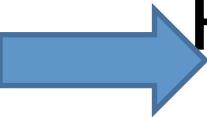
State of the art





This talk:

- Translational Bioinformatics and Information Integration Dilemma
- Case Study: Public Health Preparedness
- **Our Vision**
- Our Implementation
- Challenges

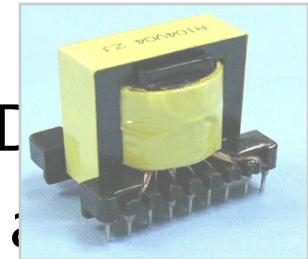


the Solution Framework

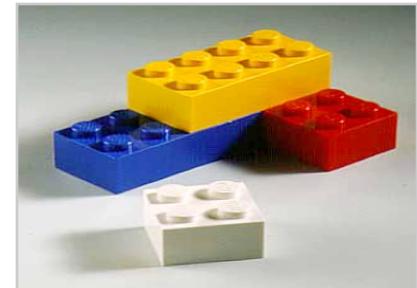
- Resource Definition Framework (RDF) to enable unified Information Representation



- Ontologies (OWL) and Computer Reasoning (D) (OWL-DL) To enable knowledge representation and reasoning



- Services Oriented Architecture Dynamic interoperability and reuse



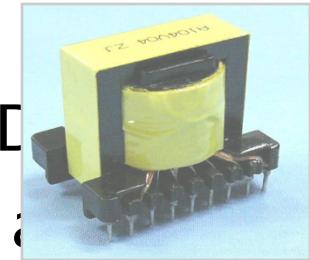
the Solution Framework

- Resource Definition Framework (RDF) to enable unified Information Representation

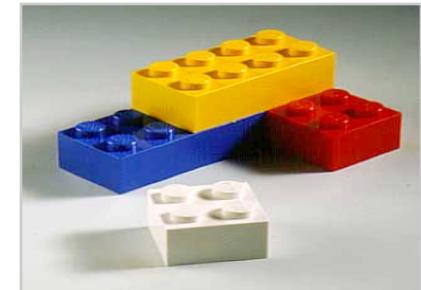
Integrative, transdisciplinary, agile, collaborative



- Ontologies (OWL) and Computer Reasoning (DL) (OWL-DL) To enable knowledge representation and reasoning



- Services Oriented Architecture Dynamic interoperability and reuse



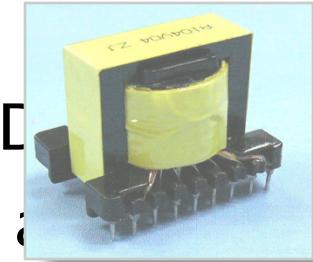
the Solution Framework

- Resource Definition Framework (RDF) to enable unified Information Representation



Integrative, transdisciplinary, agile, collaborative

- Ontologies (OWL) and Computer Reasoning (D) (OWL-DL) To enable knowledge representation and reasoning



Context aware, knowledge based, agile, transdisciplinary, collaborative

- Services Oriented Architecture Dynamic interoperability and reuse



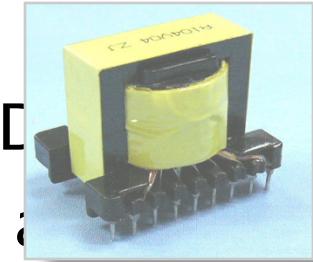
the Solution Framework

- Resource Definition Framework (RDF) to enable unified Information Representation



Integrative, transdisciplinary, agile, collaborative

- Ontologies (OWL) and Computer Reasoning (D) (OWL-DL) To enable knowledge representation and reasoning



Context aware, knowledge based, agile, transdisciplinary, collaborative

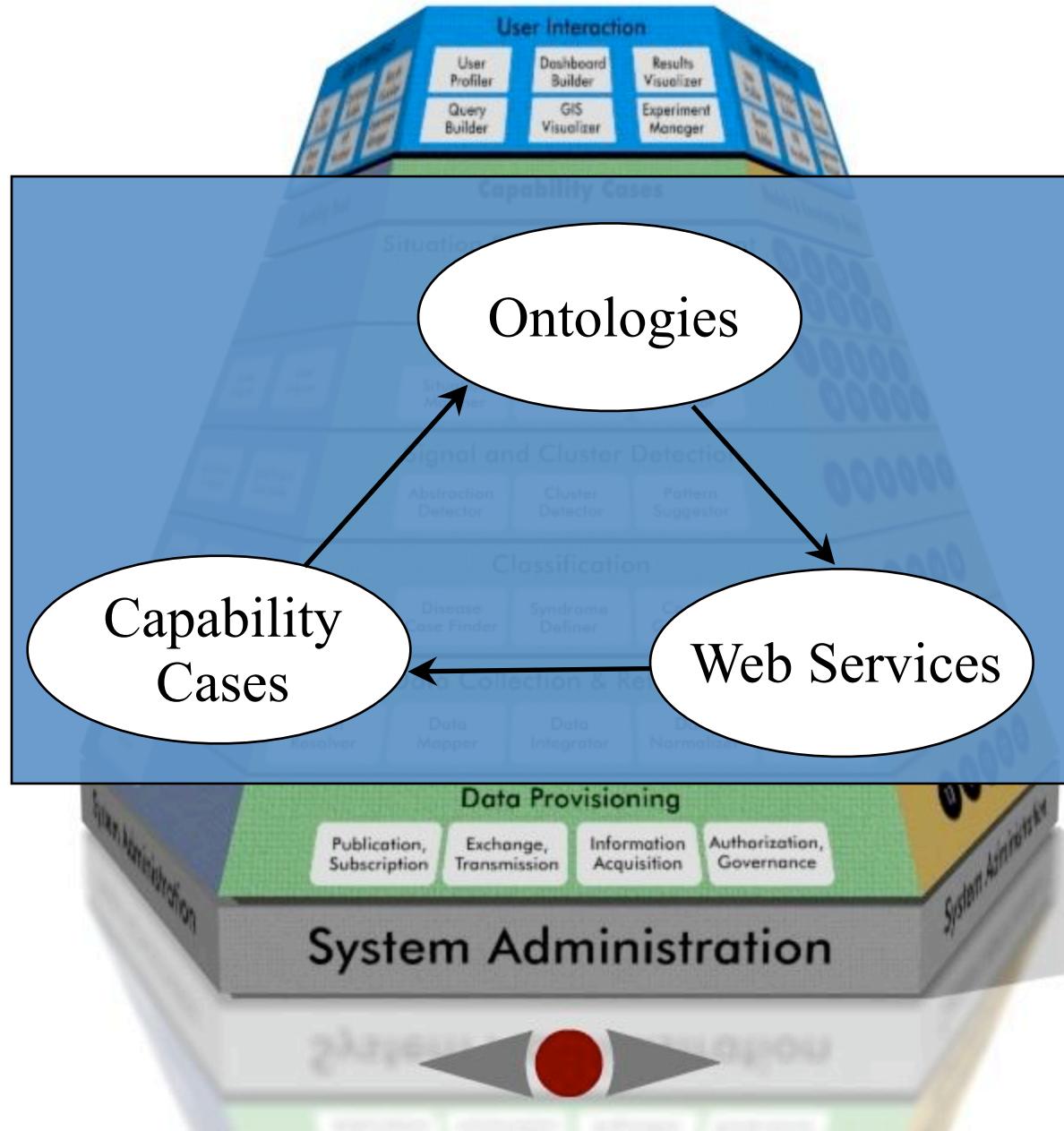
- Services Oriented Architecture

Dynamic interoperability and reuse

Agile, interoperable, collaborative and distributed



Dimensions of SARA



This talk:

- Translational Bioinformatics and Information Integration Dilemma
- Case Study: Public Health Preparedness
- Our Vision
- **Our Implementation**
- Challenges



Data Sources – 1

- Triage Data
 - Patient Demographics (Age, Ethnicity, Gender)
 - Vital Signs (T, RR, PR, PO2)
 - Chief Complaints

Data Sources – 2

- Nurse Notes
 - Vital Signs,
 - Complete Review of Systems: General, Respiratory, Neurological, Gastrointestinal, Dermatological, etc
 - Past Medical and Surgical HX
 - Medications, Past Medications, Home Medications
 - Interventions, Procedures
 - Outcome
 - Discharge and Disposition
 - Past Medical and Surgical HX

Data Sources – 1

- Triage Data
 - Patient Demographics (Age, Ethnicity, Gender)
 - Vital Signs (T, RR, PR, PO2)
 - Chief Complaints
- Nurse Notes
 - Automated submissions through HTTP
 - Accounts for about %30 Houston ED visits
 - Data transmission every 10 minutes or less
 - Over 250,000 concepts, 82 million instances and growing
 - Interventions, Procedures
 - Outcome
 - Discharge and Disposition
 - Past Medical and Surgical HX

Data Sources – 3

Texas Commission for Environmental Quality (TCEQ)

- Pollution Parameters
 - CO, SO₂, H₂S, NO, NO₂, O₃, TNMOC, CH₄, ...
- Meteorological Parameters
 - Temperature (Outdoor , Dew Point)
 - Relative Humidity,
 - Radiation (Solar, Ultraviolet, Net Radiation)
 - Barometric Pressure,
 - Precipitation, ...
- Chromatography Data
 - Ethane, Methylcyclopentane, 1,2,4-Trimethylbenzene, Ethylene, 2,4-

Data Sources – 3

Texas Commission for Environmental Quality (TCEQ)

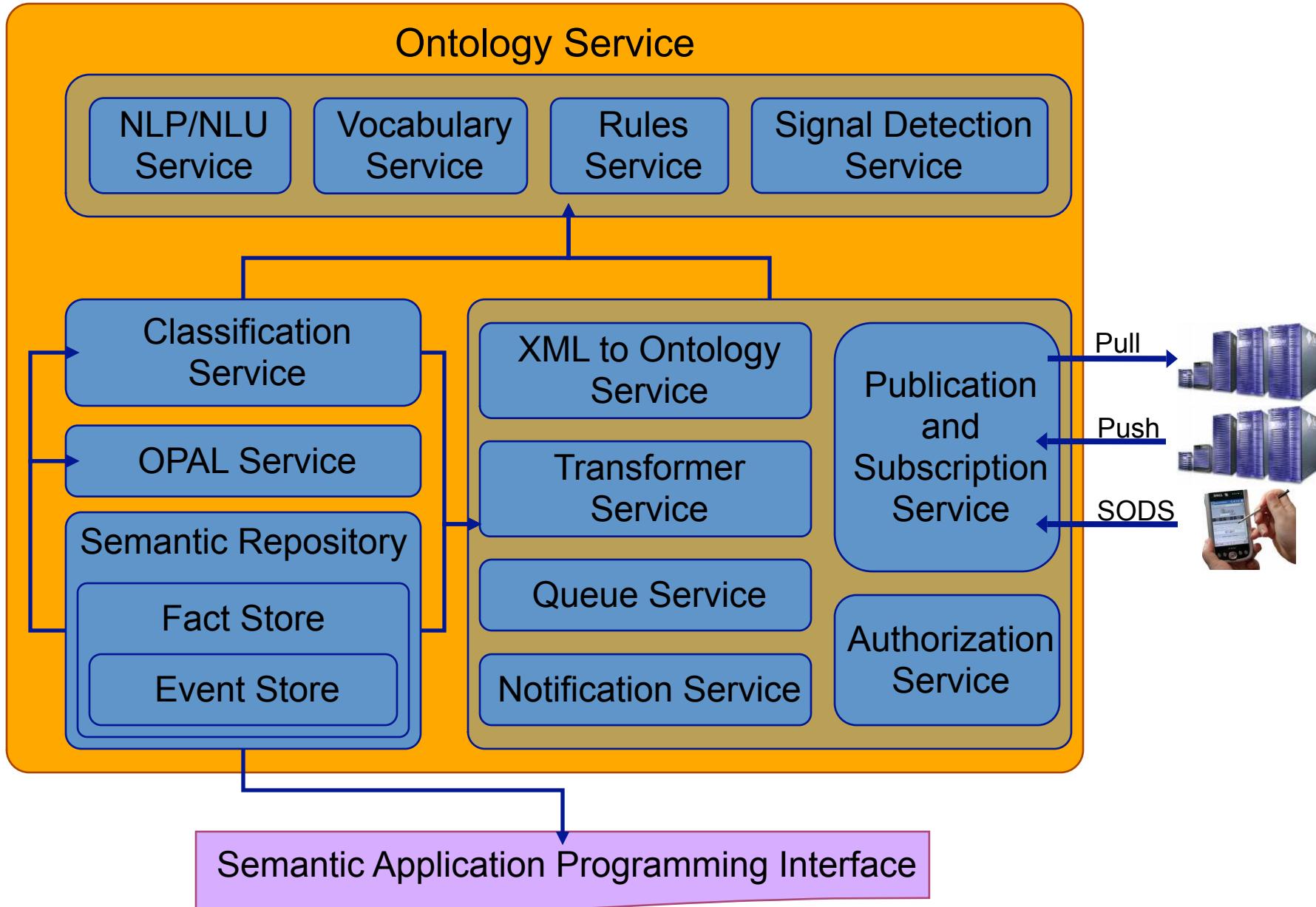
- Pollution Parameters
 - From 18 locations 2 sensors each
 - Data Transmission from TCEQ hourly, ...,
- Measurements
 - 250 concepts on each message
 - Air Quality indices calculated twice daily
 - Temperature (Outdoor, Dew Point)
 - Relative Humidity,
 - Radiation (Solar, Ultraviolet, Net Radiation)
 - Barometric Pressure,
 - Precipitation, ...
- Chromatography Data
 - Ethane, Methylcyclopentane, 1,2,4-Trimethylbenzene, Ethylene, 2,4-

```
<Patient VisitId="130996" PersonId="85627293" Gender="Male" Race="Caucasian / White" Age="61 years" InitTemperature="Temp 97.2F Oral" InitBloodPressure="BP 152 / 89"  
InitPulse="Pulse 104bpm" InitPulseOximetry="Pulse Ox 96% Flow RA Room Air">  
  <Complaint Value="Other" Custom_Text="Bleeding Toe"/>  
  - <Charts>  
    - <Chart Type="Nurse">  
      - <Section Name="Observations">  
        - <Choice Choice_Name="" Choice_Display_Name="Activity" StrikeOut="False">  
          <Choice2 Choice_Name="Resting quietly." Choice_Display_Name="Resting quietly" StrikeOut="False"/>  
          <Choice2 Choice_Name="No apparent distress." Choice_Display_Name="No apparent distress" StrikeOut="False"/>  
        </Choice>  
        - <Choice Choice_Name="" Choice_Display_Name="Awaiting" StrikeOut="False">  
          <Choice2 Choice_Name="Awaiting disposition." Choice_Display_Name="Disposition" StrikeOut="False"/>  
        </Choice>  
      </Section>  
      <Section Name="Presentation"/>  
      <Section Name="Historical"/>  
      <Section Name="Screening"/>  
      <Section Name="Tri. Assess"/>  
      - <Section Name="Vitals">  
        <Choice Choice_Name="Vitals" Choice_Display_Name="Vitals" StrikeOut="False" Plugin_Name="VitalSigns" Plugin_Value="BP 158 / 93"/>  
        <Choice Choice_Name="Vitals" Choice_Display_Name="Vitals" StrikeOut="False" Plugin_Name="VitalSigns" Plugin_Value="BP 159 / 78"/>  
      </Section>  
      <Section Name="Assessment"/>  
      <Section Name="Interventions"/>  
      - <Section Name="Procedures">  
        <Choice Choice_Name="No Procedures done." StrikeOut="False"/>  
      </Section>  
      - <Section Name="Medication">  
        - <Choice Choice_Name="Meds follow up." Choice_Display_Name="Meds follow up" StrikeOut="False">  
          <Choice2 Choice_Name="No adverse reaction" Choice_Display_Name="No adverse reaction" StrikeOut="False"/>  
        </Choice>  
      </Section>  
      - <Section Name="Outcome">  
        - <Choice Choice_Name="" Choice_Display_Name="Condition" StrikeOut="False">  
          <Choice2 Choice_Name="Condition: stable" Choice_Display_Name="Stable" StrikeOut="False"/>  
          <Choice2 Choice_Name="Condition: good" Choice_Display_Name="Good" StrikeOut="False"/>  
        </Choice>  
      </Section>
```

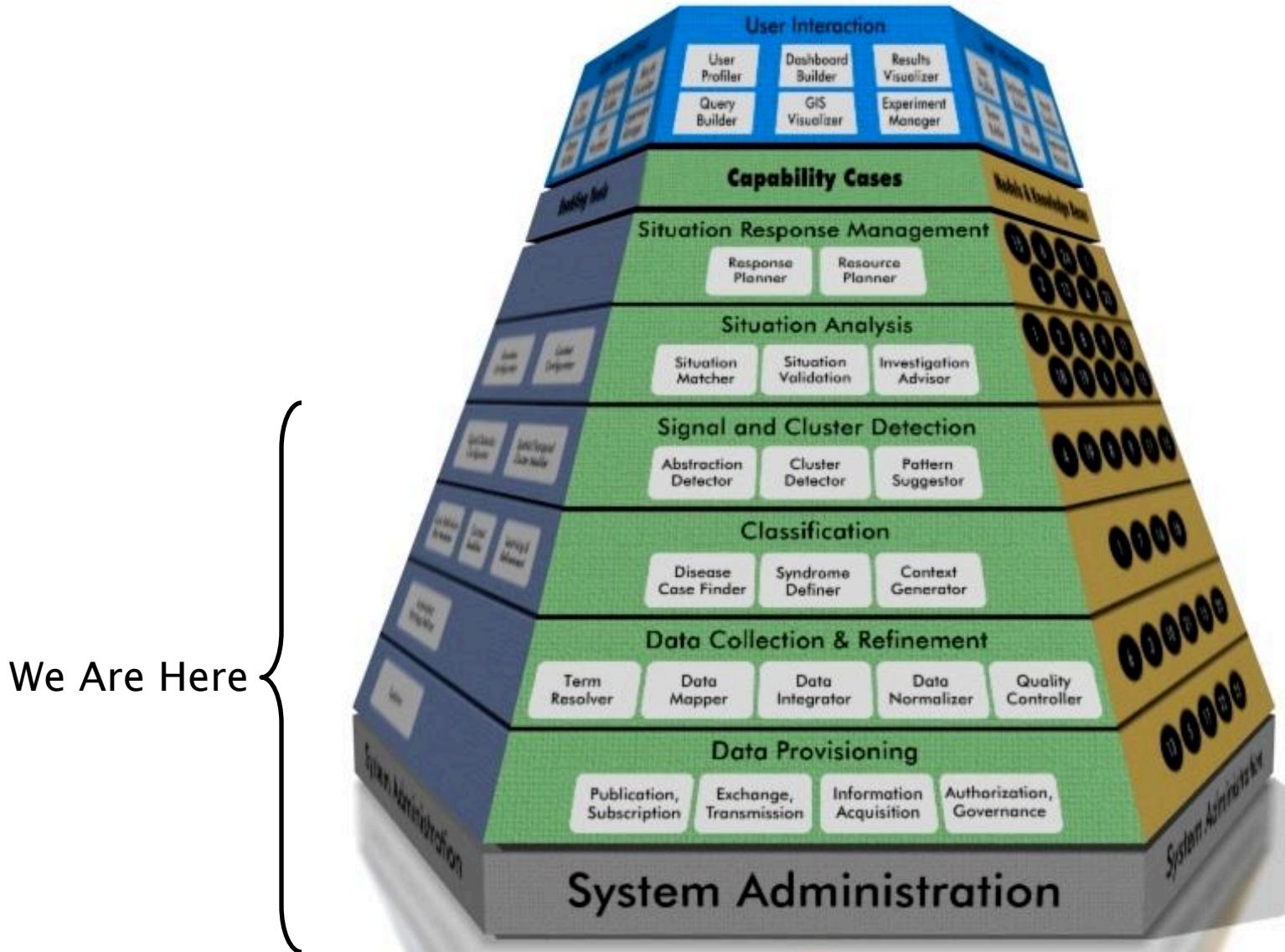
Plugin_Value="BP 159 / 78"

Choice="No adverse reaction"

Semantic Integration



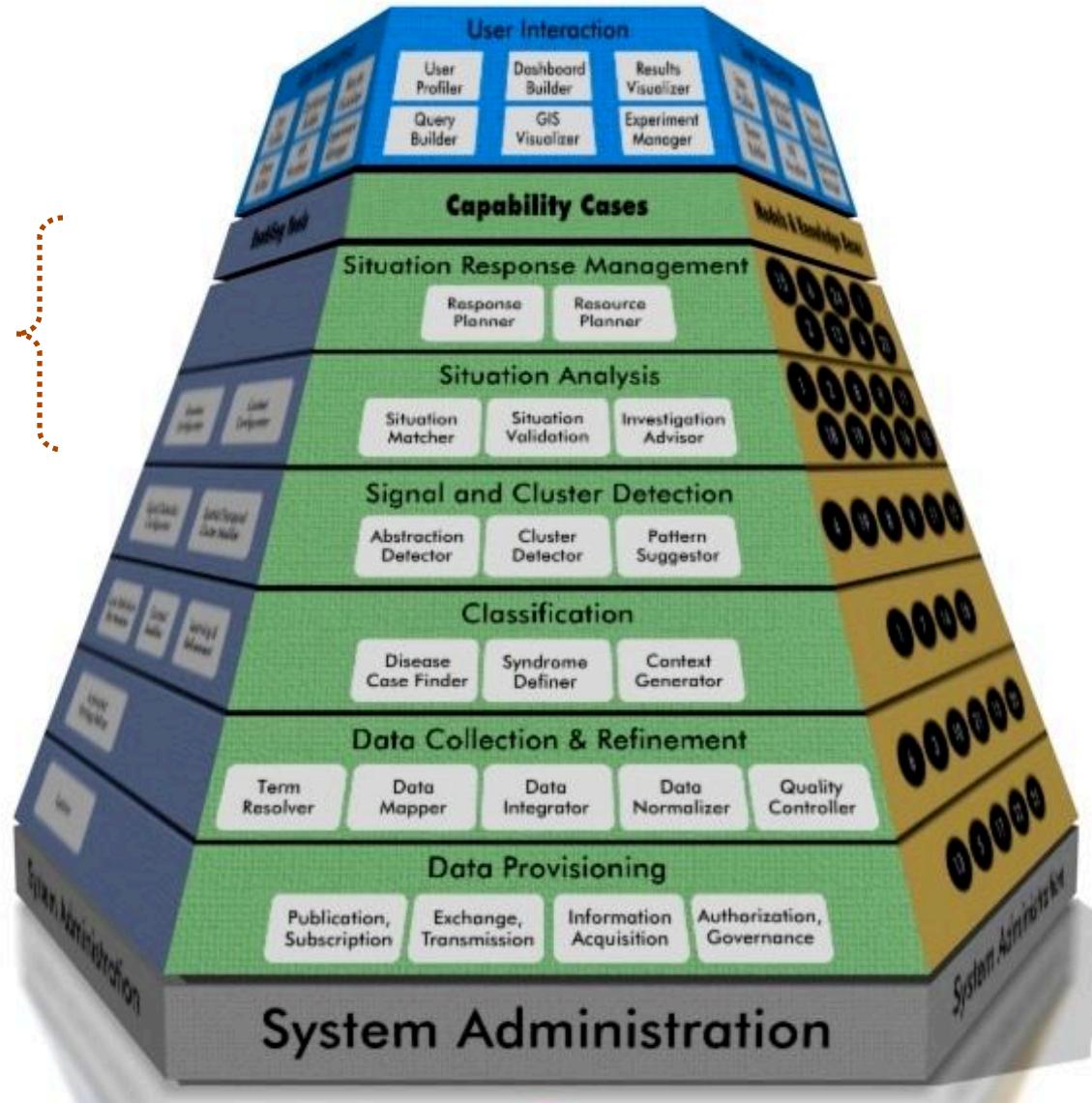
SAPPHIRE Implementation



SAPPHIRE Implementation

Proof of concept

We Are Here



Implementation Platform

- 1- TopBraid Composer as Ontology Management Tool
- 2- Jena from HP as API for Semantic Web
- 3- Eclipse Java Development Environment
- 4- Oracle Semantic Data Management (Started with 10gR2 on Windows, Currently 11gR1 on Linux)
- 5- Pellet/Jena OWL Micro Reasoner
- 6- Services Oriented Architecture
- 7- Microsoft SQL Server 2005 XML archive and Analysis Services
- 8- IBM Dual Xeon 2.8GH/3GB RAM Blade Server
- 9- EqualLogic iSCSI SAN (4 TB)
- 10- GB Ethernet LAN

This talk:

- Translational Bioinformatics and Information Integration Dilemma
- Case Study: Public Health Preparedness
- Our Vision
- Our Implementation
- Challenges



Challenges

- State of the frameworks
- Maturity of Tools
- Knowledge Engineering and Ontology Development
- Reasoning and Rules Support
- Scalability
- Performance

Academic and Industrial

- Scalable, High Performance RDF/OWL Repositories (Oracle, Franz)
- Scalable Semantic Application Programming Interface (Oracle, TopQuadrant, HP)
- Ontology based Business Intelligence and Data Mining (TopQuadrant)

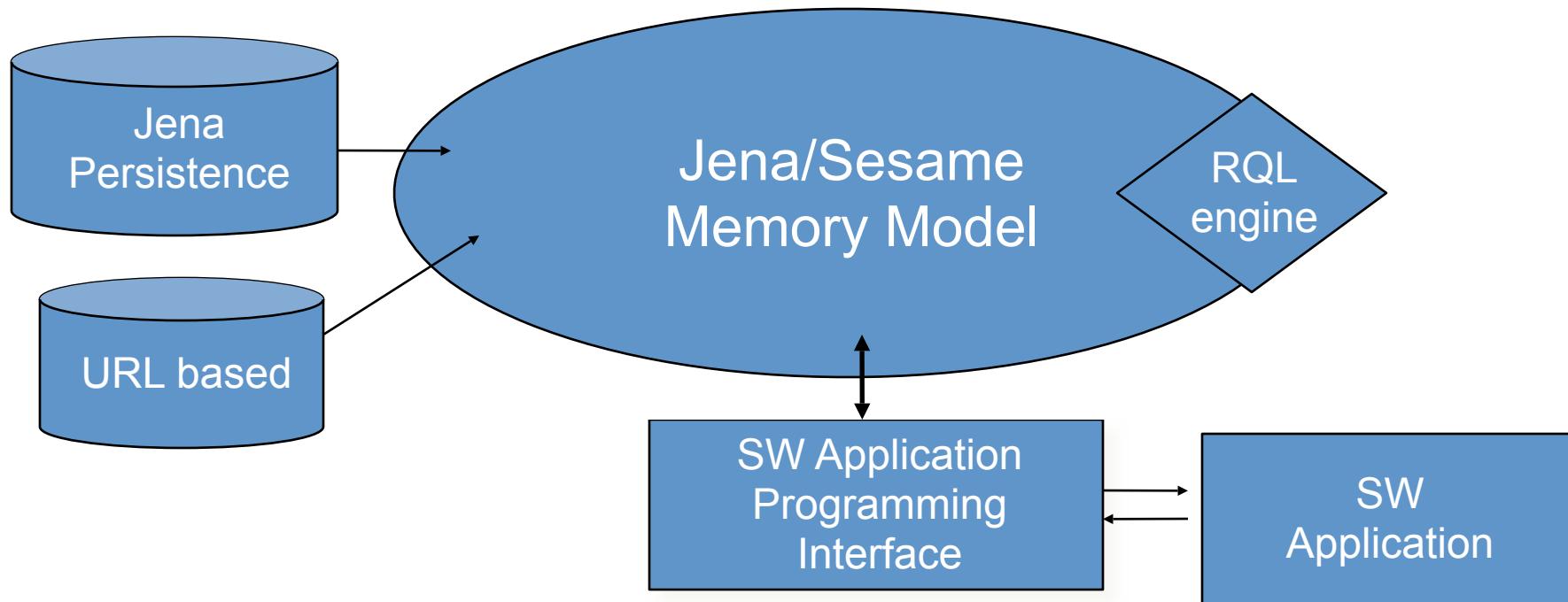
Scalable Semantic Application Programming

- Scalable SW application development interface for Oracle Semantic Data Management (SDM)
- Seamless integration of application development interfaces to Oracle SDM
- Without any intermediate or ‘in memory’ representation of semantic data

Method (conceptual)

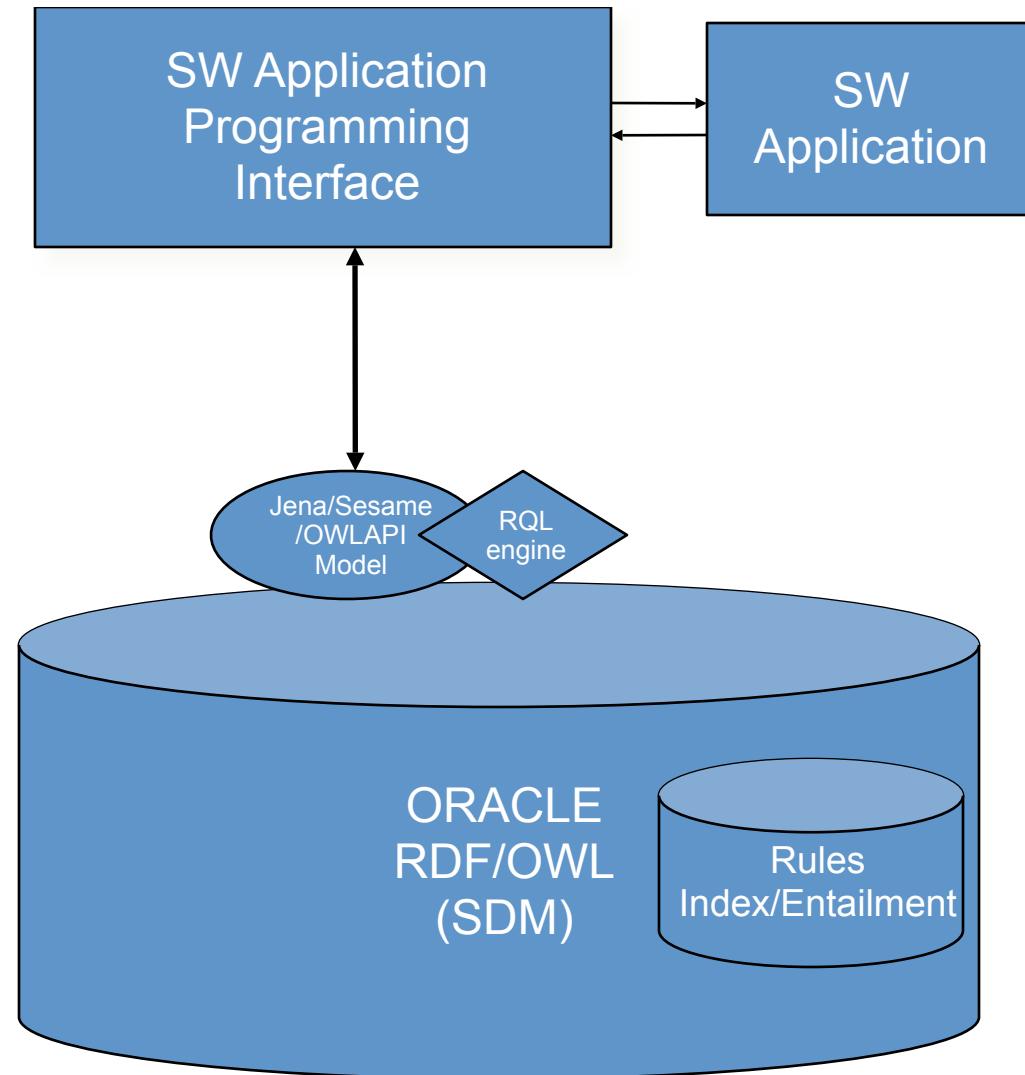
- Adopt existing SW application development frameworks (e.g. Jena)
- Extend by eradicating the intermediate representation of Semantic data in memory
- Support SPARQL in the applications side
- Enable invocation and use of integrated rules engines through API

Jena Conceptual Model As Is

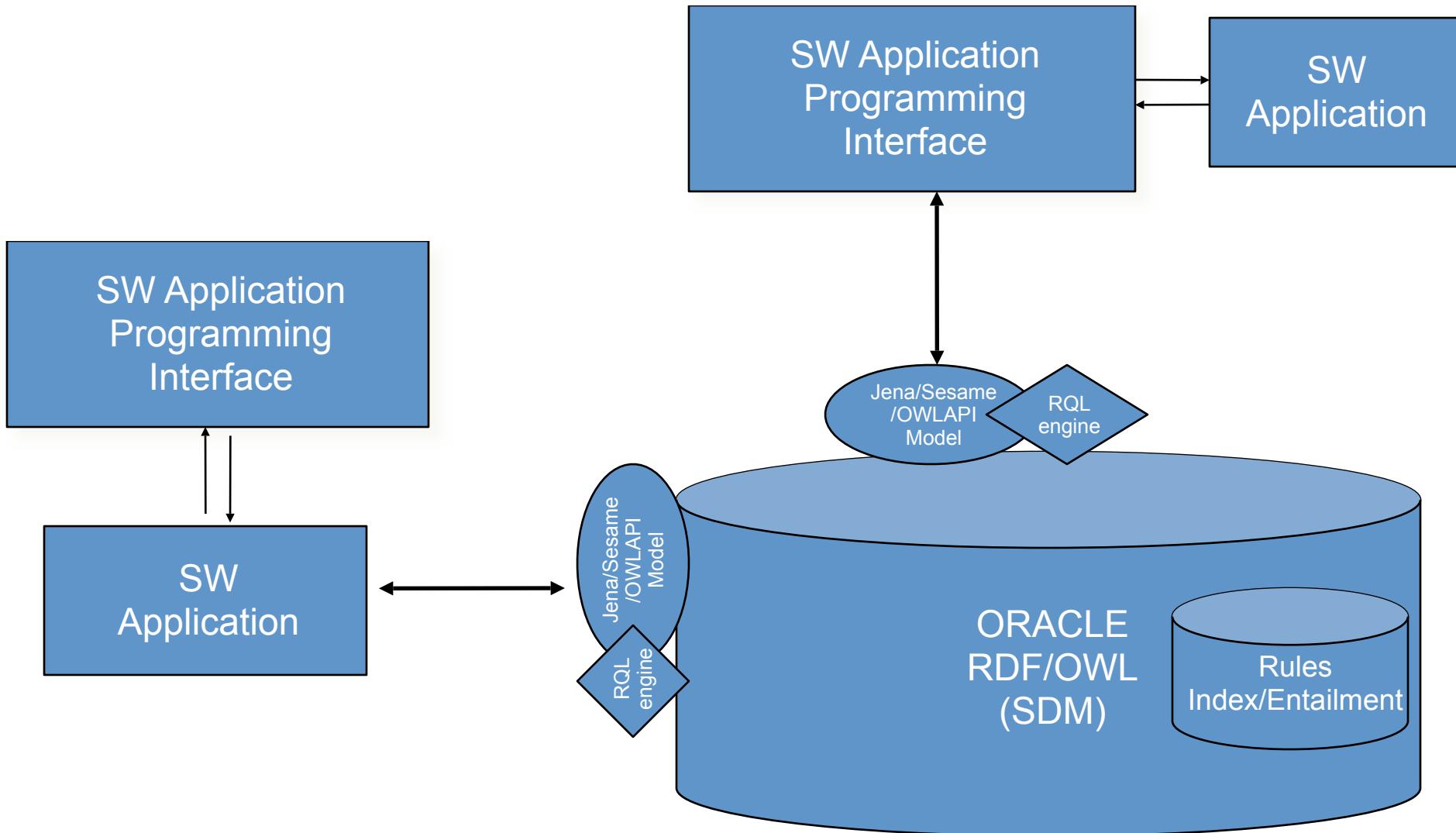


- 👎 Not scalable, depends on memory resources
- 👎 Does not support concurrent users and distributed applications
- 👎 Reasoning is not scalable and integrated with repositories

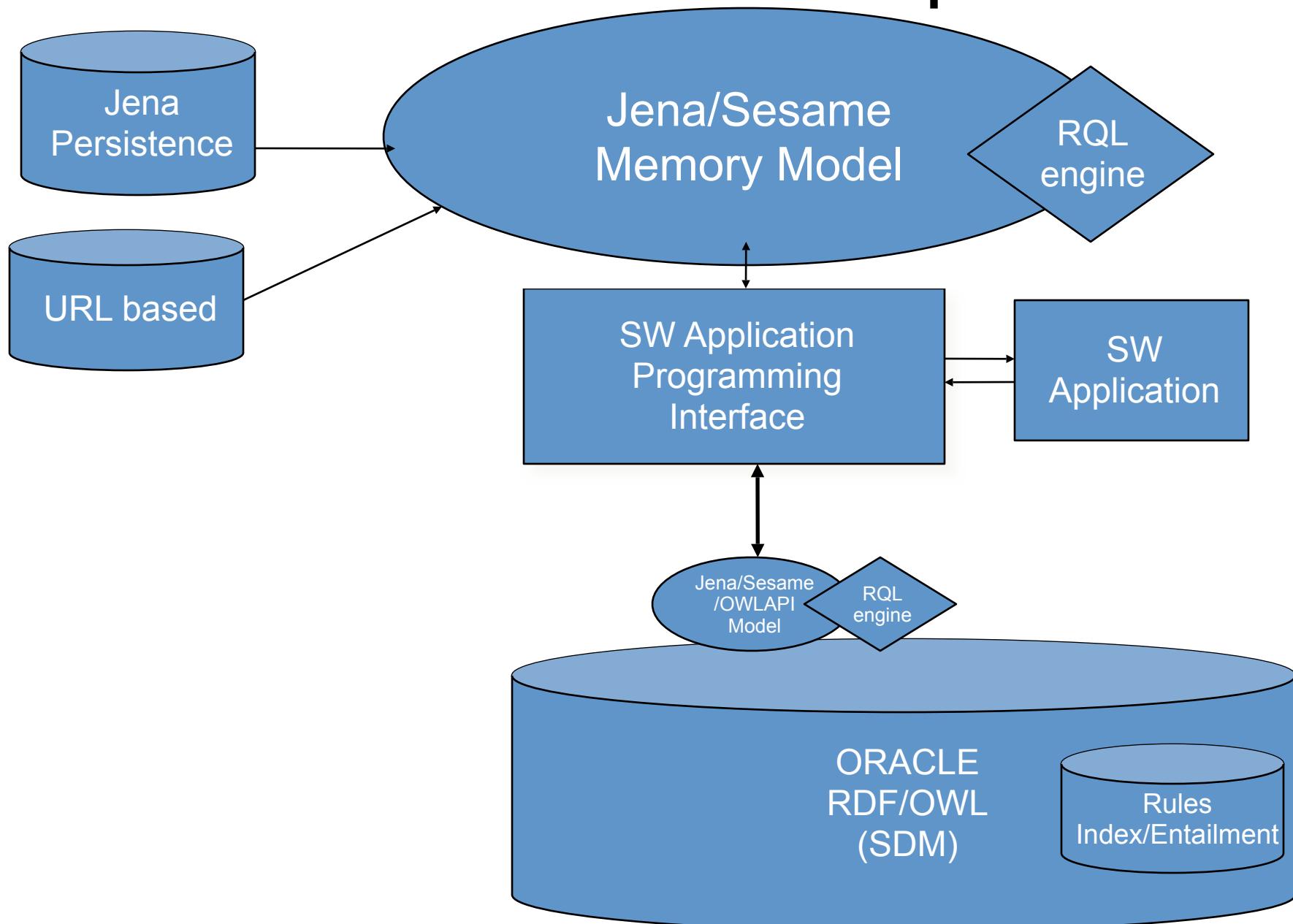
Oracle Jena Adapter



Oracle Jena Adapter



Oracle Jena Adapter



Pros and Cons

- ◀ Enables use of Oracle SDM for large scale implementations using Graph and Model objects
- ◀ Complete but indirect support of SPARQL
- ◀ Supports multi-user and distributed application environments
- ◀ Integrated Support of Oracle Reasoners (RDFS and OWLPrime) in the application side
- ◀ Robust performance through both programming interface as well as SPARQL querying

- ◀ Indirect support of SPARQL (not available outside of the API through SQL Developer for example)
- ◀ OntModel to be supported in future releases
- ◀ Rules reasoning not real time

Contact

- Parsa Mirhaji, MD Director

The Center for Biosecurity and Public Health
Informatics Research The School of Health
Information Sciences The University of Texas –
Health Sciences Center at Houston Office: (713)
500-3157 Fax: (713) 500-0370 Assistance
(Namiko Burleson): (713) 500-3938

- <http://www.phinformatics.org/ResearchProjects/SAPPHIRE/tabid/76/Default.aspx>