

Semantics for Healthcare and the Life Sciences

Vipul Kashyap
Partners HealthCare System


Eric Neumann
Teranode Corporation

Tonya Hongsermeier
Partners HealthCare System

vkashyap1@partners.org

Tutorial Presentation
International World Wide Web Conference (WWW 2006)
26th May, 2006
Edinburgh, UK

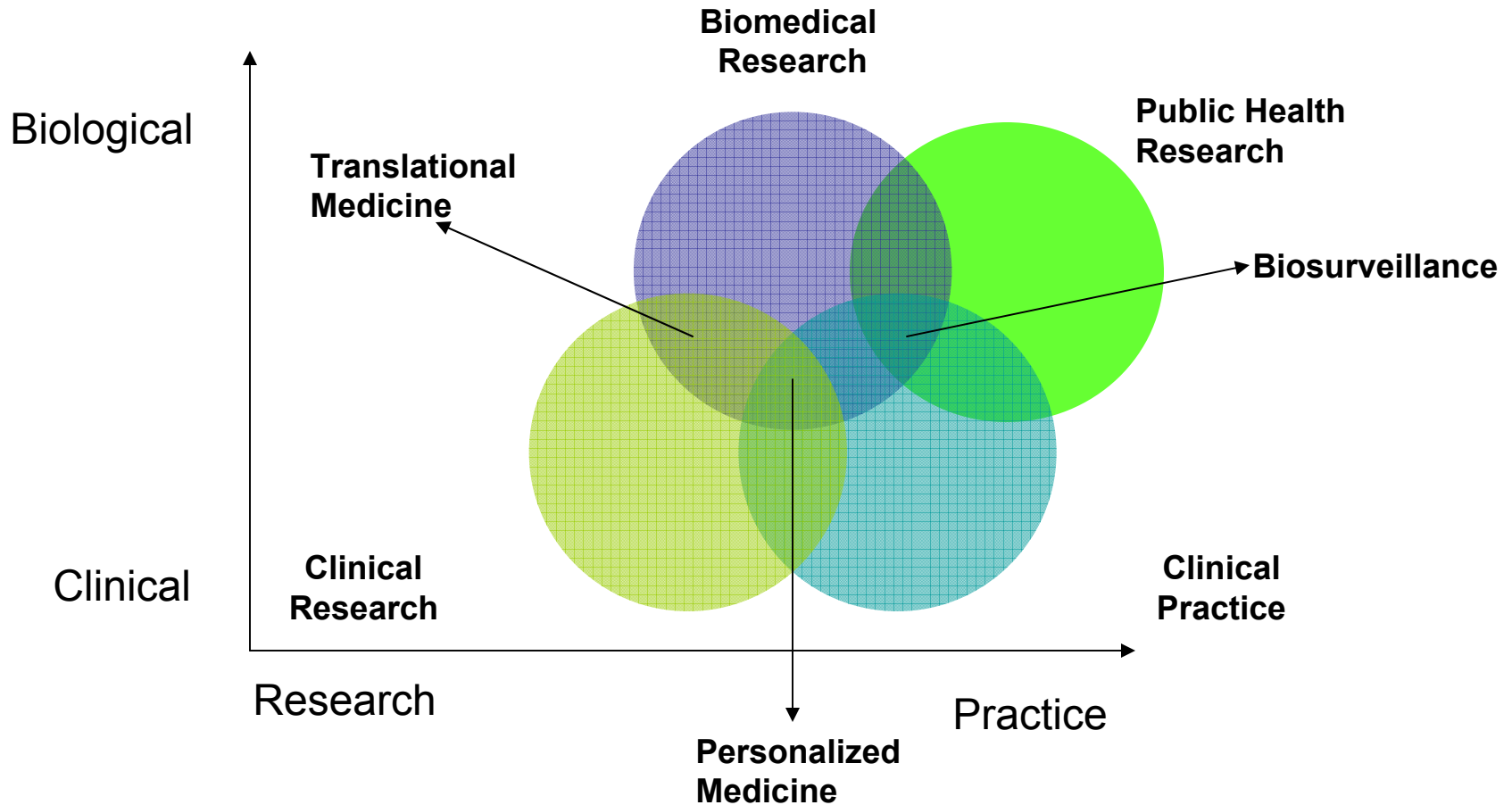
Outline

- The Bench ↔ Bedside Vision 
- Role of Data, Knowledge and Semantics in the HCLS Ecosystem
- Functional Requirements
 - Metadata-based Semantic Annotation
 - *Recombinant Data*: Information Aggregation and Integration
 - Ontology-driven Decision Support
 - Knowledge Update and Maintenance
- Conclusions

The Bench ↔ Bedside Vision

- Healthcare and Life Sciences: Framework
 - An Information/Knowledge Perspective
- Current Challenges
- The Healthcare Life Sciences (HCLS) Ecosystem
- HCLS Ecosystem: Business Drivers

Healthcare and Life Sciences: Framework



HCLS Framework: Biomedical Research

- Systems Biology/Physiology
 - Organism as an integrated an interacting network of genes, proteins and biochemical reactions
 - Human body as a system of interacting organs
- Molecular Cell Biology/Genomic and Proteomic Research
 - Gene Sequencing, Genotyping, Protein Structures
 - Cell Signaling and other Pathways
- Biomarker Research
 - Discovery of genes and gene products that can be used to measure disease progression or impacts of drugs on patients
- Pharmaco-genomics
 - Impact of genetic inheritance on effects of drugs on patients
- Drug Discovery and Translational Research
 - Use of preclinical research to identify promising drug candidates

HCLS Framework: Clinical Research

- Clinical Trials
 - Determination of efficacy, impact and safety of drugs for particular diseases
- Pharmaco-vigilance/ADE Surveillance
 - Monitoring of impacts of drugs on patients, especially safety and adverse event related information
- Patient Cohort Identification and Management
 - Identifying patient cohorts for drug trials is a challenging task
- Translational Research
 - Test theories emerging from pre-clinical experimentation on disease affected human subjects
- Development of EHRs/EMRs for both clinical research and practice
 - Currently EHRs/EMRs focussed on clinical workflow processes
 - Re-using that information for clinical research and trials is a challenging task

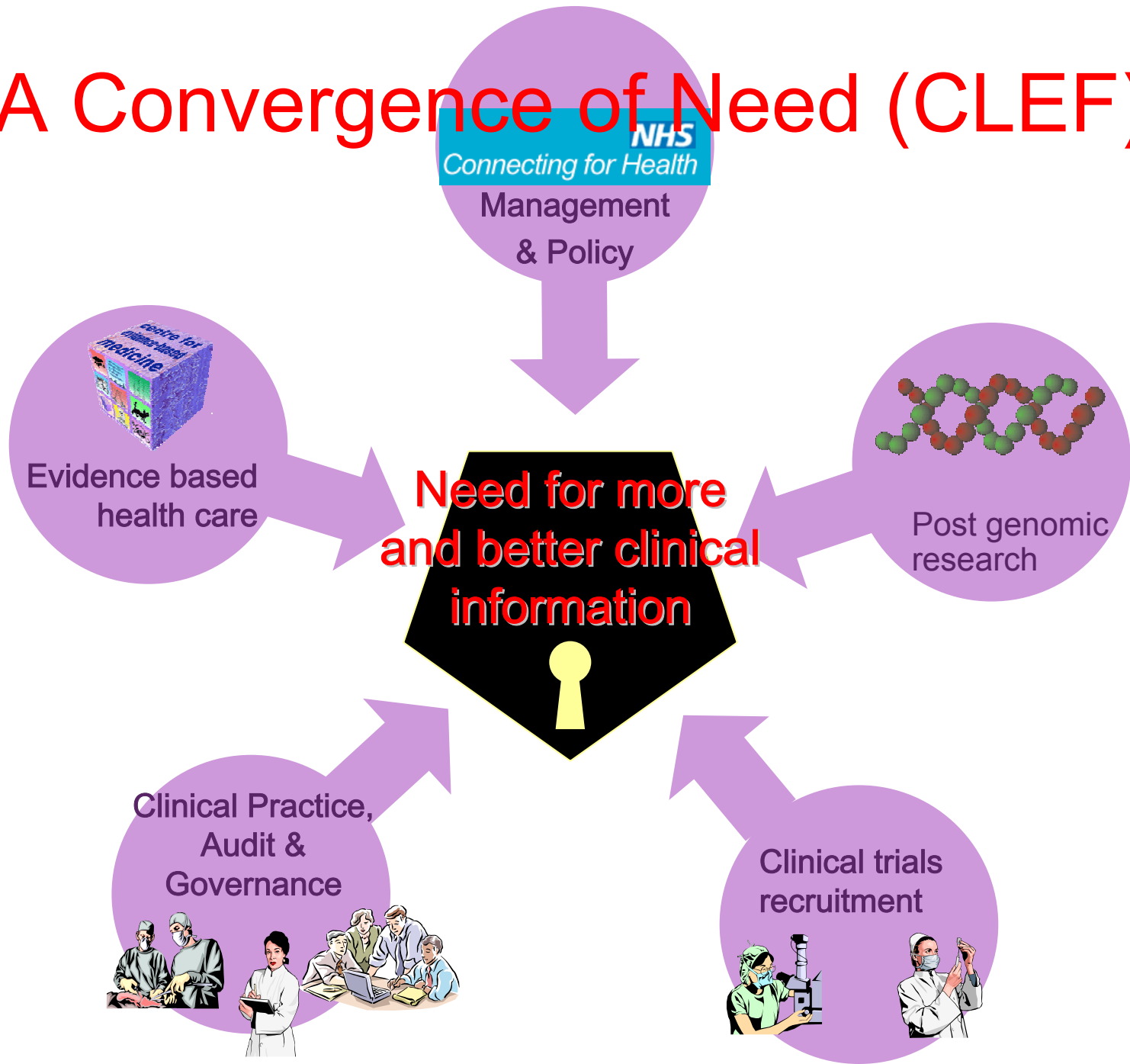
Clinical Practice

- Electronic Medical/Health Record
 - Integration of Structured and Unstructured Information
 - Design of EHRs/EMRs for both clinical research and practice
- Computerized Physician Order Entry
 - Computerized aids for submitting medication and lab orders
- Clinical Disease Support
 - Physician perspective:
 - Therapeutic Decision Support, Drug Drug Interactions
- Structured Clinical Documentation
 - Templated forms to aid structured observation capture and storage into the electronic medical record
- Enterprise Terminological Services
 - Standardization of definitions and codes for conditions, findings, observations, labs, therapies, diagnoses, etc.
- Disease Management
 - Portals containing information relevant to a particular disease condition, e.g., diabetes
- Personalized Medicine
 - Personalizing therapeutic recommendations based on genetic profile of patient

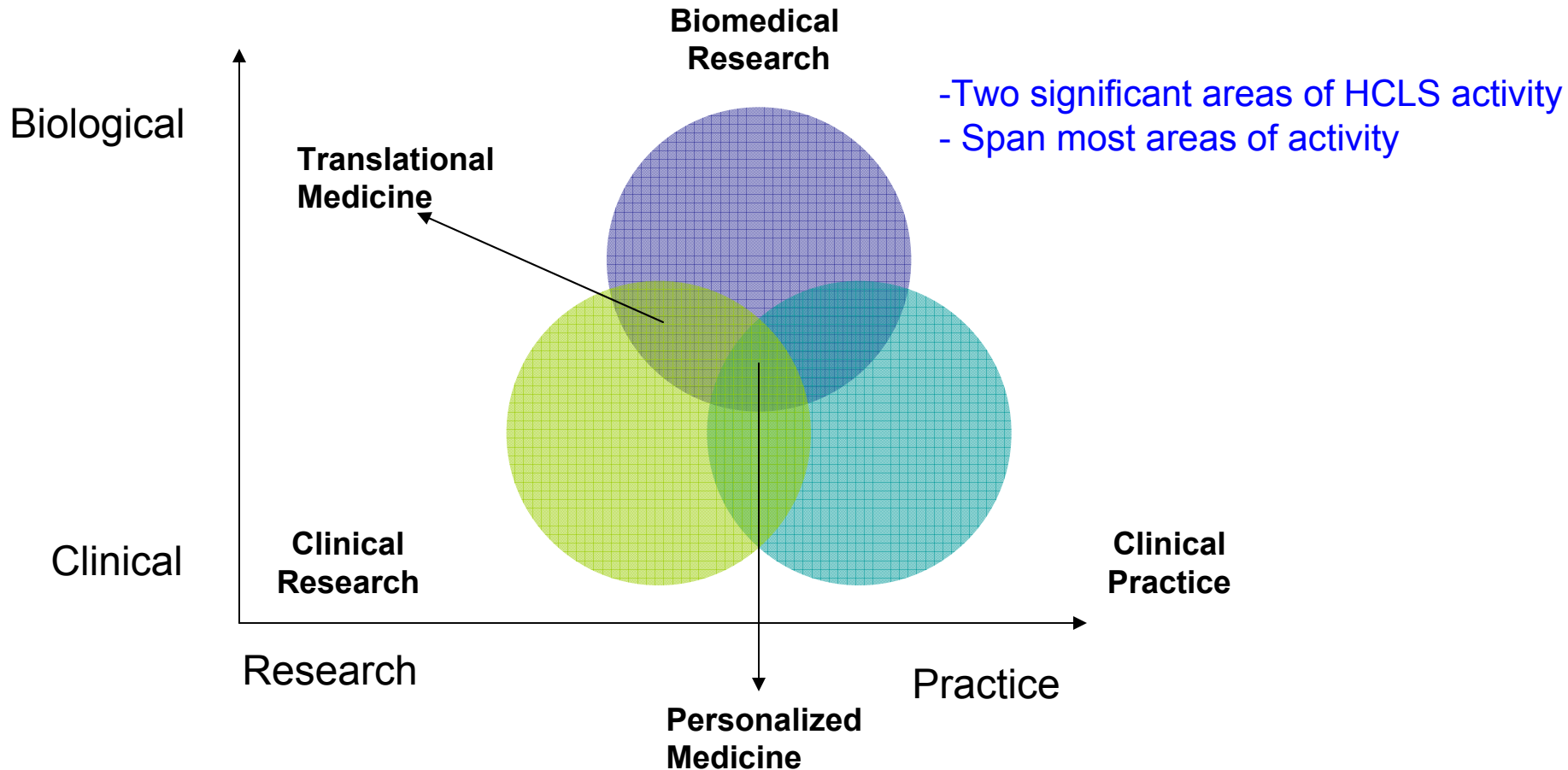
Public and Consumer Health

- Epidemiology/Bio-surveillance
 - Monitoring of disease occurrences for unusual patterns
 - Indicative of epidemics, terrorist attacks
- Bio-sensors
 - E.g., detection of cancer causing agents in ground water (ORNL)
- Consumer Health Portals
 - Health Information Prescription
 - Electronic Prescription
- Disease Management
 - Portals containing reminders and alerts for patients for upcoming physicals, labs, etc.
- Personalized Health Records
 - Presentation of patient health related information in a language understandable to the lay person
- Clinical Decision Support:
 - Patient Perspective – help in choosing a good doctor
 - Population perspective – help deploy appropriate resources in appropriate areas

A Convergence of Need (CLEF)



Translational Research and Personalized Medicine




Translational Research

- Improve communication between basic and clinical science so that more therapeutic insights may be derived from new scientific ideas - and vice versa.
- Testing of theories emerging from preclinical experimentation on disease-affected human subjects.
- Information obtained from preliminary human experimentation can be used to refine our understanding of the biological principles underpinning the heterogeneity of human disease and polymorphism(s).
- <http://www.translational-medicine.com/info/about>
- NIH Roadmap activity
 - <http://nihroadmap.nih.gov>

Personalized Medicine

- Propagation of insights from Genomic research into clinical practice
- Impact of new Molecular diagnostic tests hitting the market
 - How can they be incorporated into clinical care?
 - How does one update current clinical guidelines to incorporate the use of these tests
 - How can one enable novel clinical decision support?
- How can phenotypic characteristics and genomic markers be used to:
 - Stratify patient populations
 - “Personalize” clinical care
 - Genetic test results as risk factors
 - Therapeutic use of genomic markers

The Bench ↔ Bedside Vision

- Healthcare and Life Sciences: Framework
 - An Information/Knowledge Perspective
- Current Challenges 
- The Healthcare Life Sciences (HCLS) Ecosystem
- HCLS Ecosystem Business Drivers



Hygieia, G. Klimt

Healthcare and Life Sciences: Current Challenges

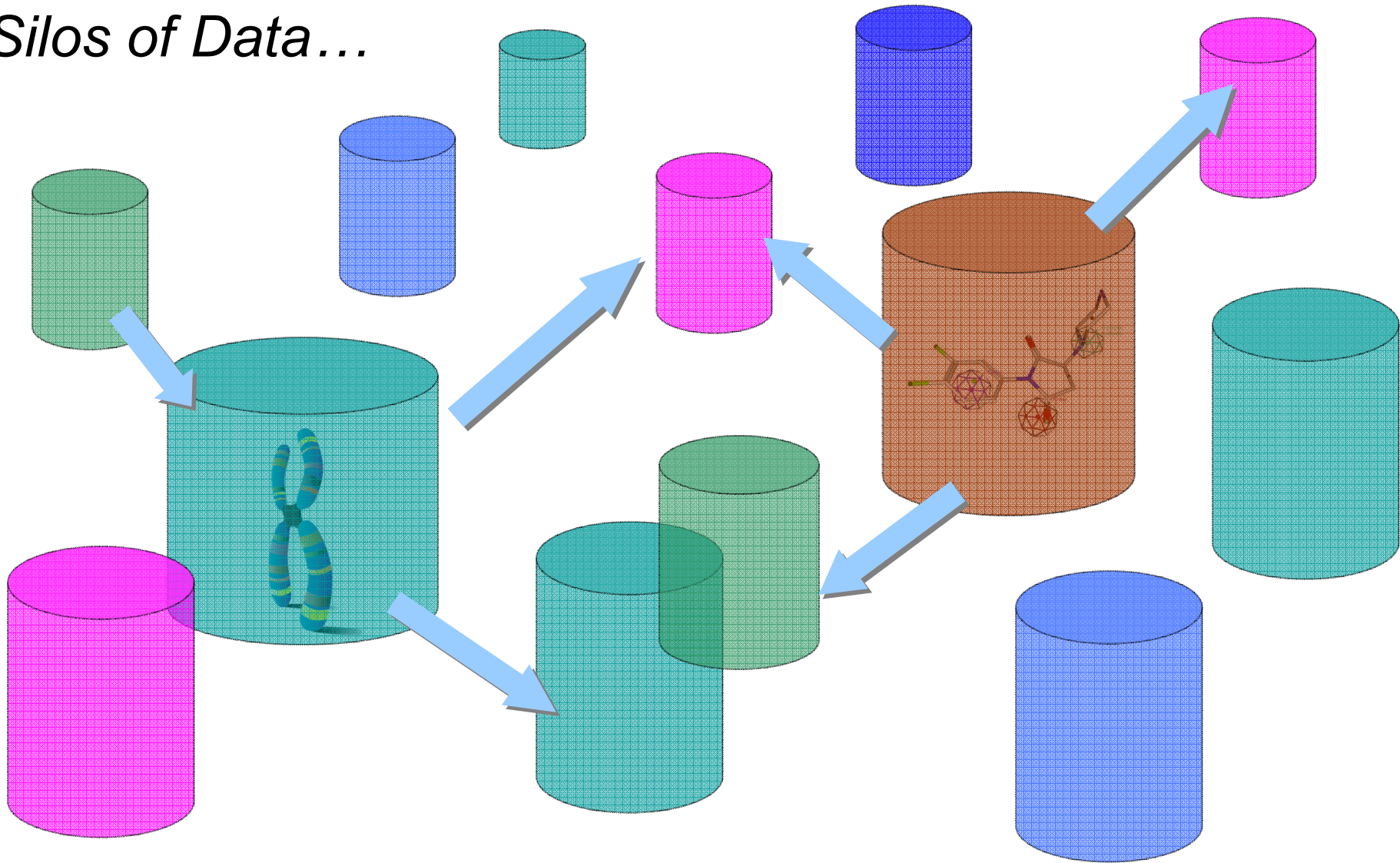
- *Health*
- *Practice*
- *Safety*
- *Prevention*
- *Knowledge*

Current Challenges: Biomedical Research

- Bridging between the genome and phenome
- Integrating information across medical specialties, across model organisms, between biological and clinical
- Lack of Data/Knowledge Provenance, Rich Annotations
- Lack of creation, management and inferencing with biomarker evidence.
- Drug Discovery Process
 - How to break away from current “Conveyor Belt Model” and adopt the “Translational Research Model” gaining and sharing insights throughout the process
 - Issues related to safety, efficacy and adverse event detection do not seem to have the attention of the biomedical research community

Distributed Nature of R&D

Silos of Data...



Current Challenges: Clinical Practice

- Medical literature doubling every 19 years
- 2 Million facts needed to practice
- Limited decision support functionality implemented by current vendor products
 - A typical drug order recommendation, accounts for, at best, Age, Weight, Height, Labs, Other Active Meds, Allergies, Diagnoses
- There are 3000+ molecular diagnostic tests on the market, genomics and personalized medicine will increase the speed of change of evidence exponentially

Current Challenges: Clinical Decision Support

- How do we supply meaningful decision support that both improves quality of care for patients and quality of life for clinicians (and self-managing patients)?
- How do we implement scalable decision support in presence of a large number of decision variables to be introduced by genomic knowledge?
- How do we affordably develop, acquire and maintain the knowledge bases required to deliver meaningful decision support?
- How do we adapt to ever changing clinical knowledge and incorporate new knowledge into clinical decision support knowledge bases?

Current Challenges:

Clinical Knowledge Management

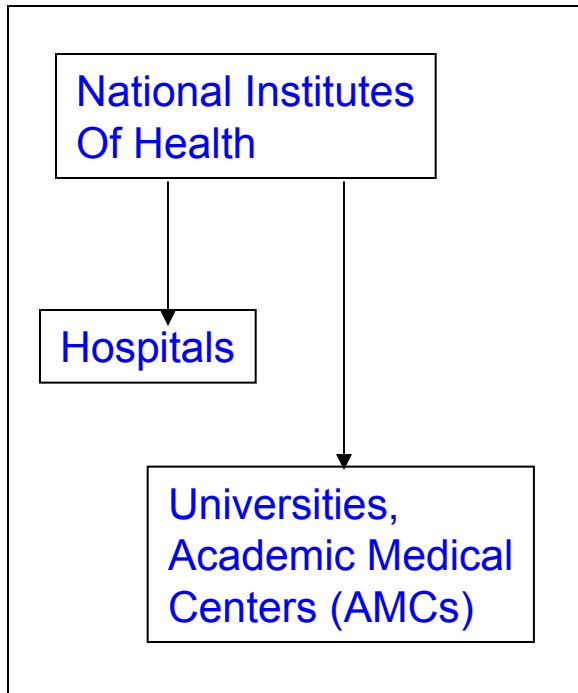
- The rate of change for contraindication definition today is very slow, yet it's a challenge for most EHRs to provide decision support for contraindication or indication management
- With the advent of molecular medicine, this rate of change could become more rapid, possibly daily
- Change in a contraindication definition can lead to changes in associated clinical decision support rules, order sets and templates and other related content areas.
- How does one create knowledge editing tools and software infrastructure that can be used by subject matter experts, knowledge engineers and informaticians to rapidly create and manage different types of clinical knowledge?

The Bench ↔ Bedside Vision

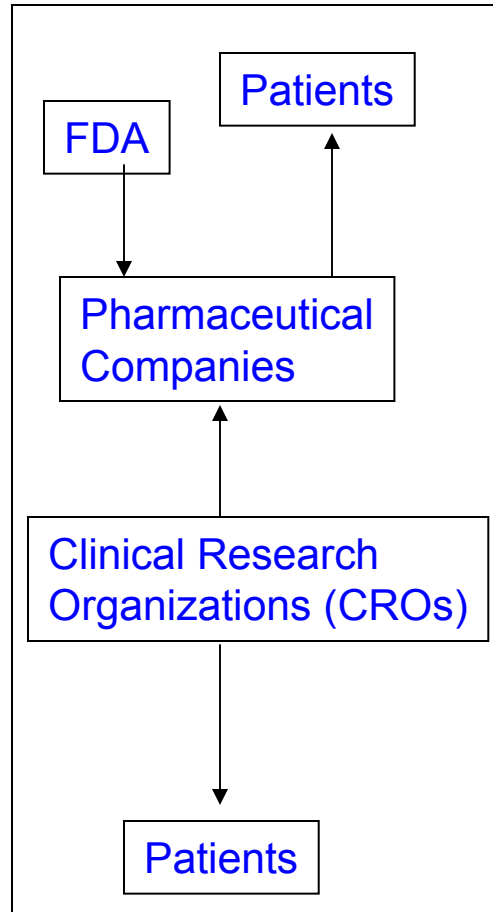
- Healthcare and Life Sciences: Framework
 - An Information/Knowledge Perspective
- Current Challenges
- The Healthcare Life Sciences (HCLS) Ecosystem
- HCLS Ecosystem Business Drivers

Ecosystem: Current State

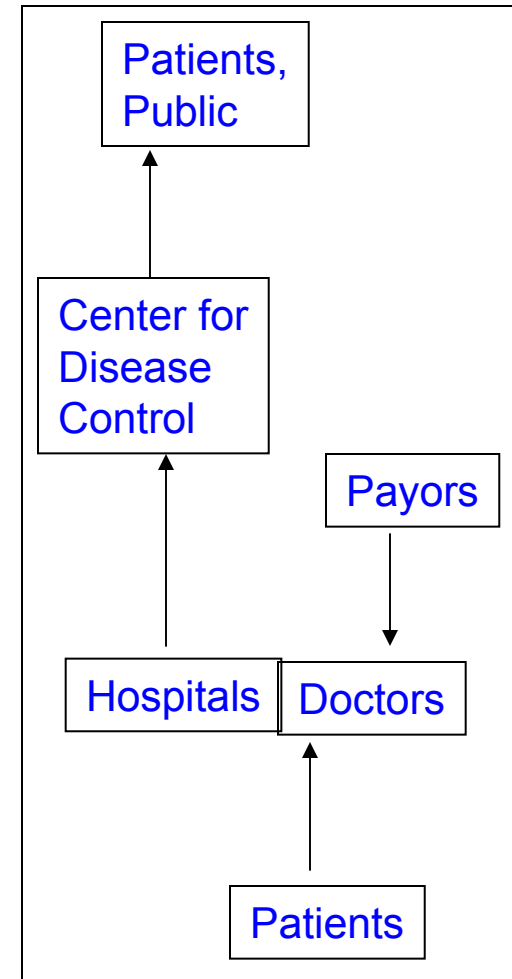
Characterized by silos with uncoordinated supply chains leading to inefficiencies in the system



Biomedical Research
Clinical Practice

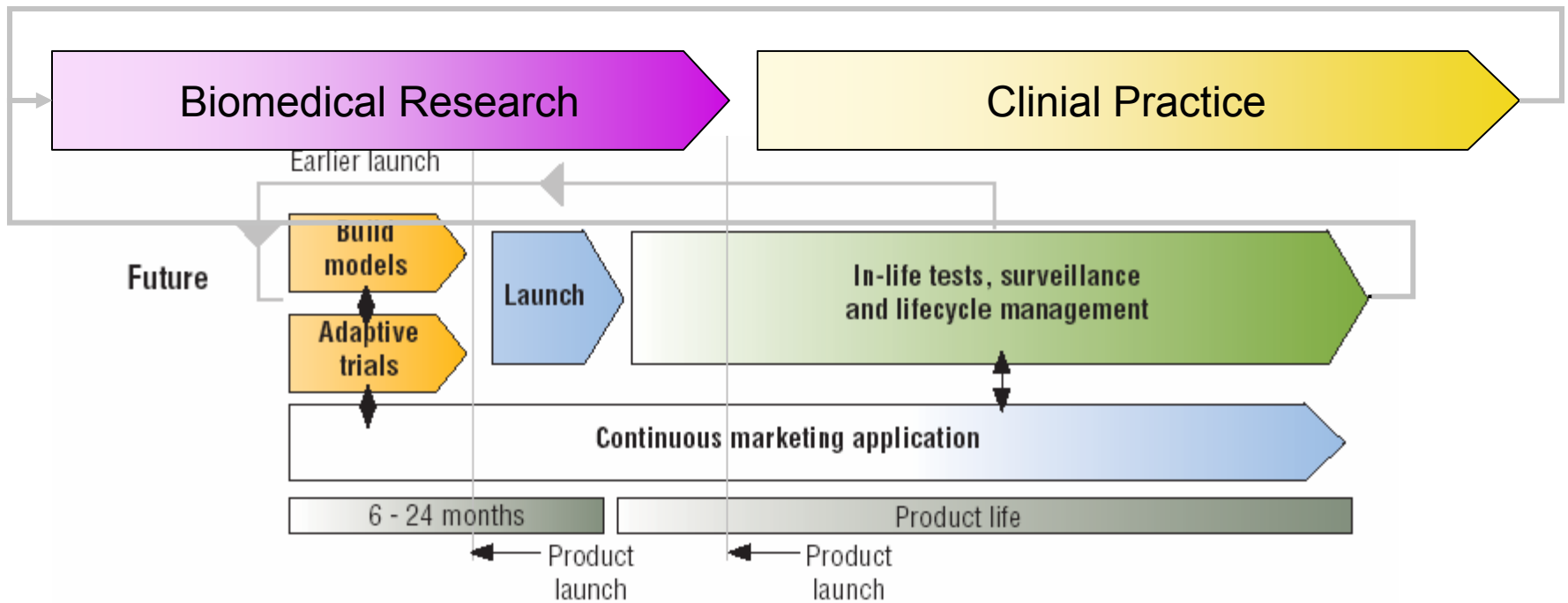


Clinical Trials/Research

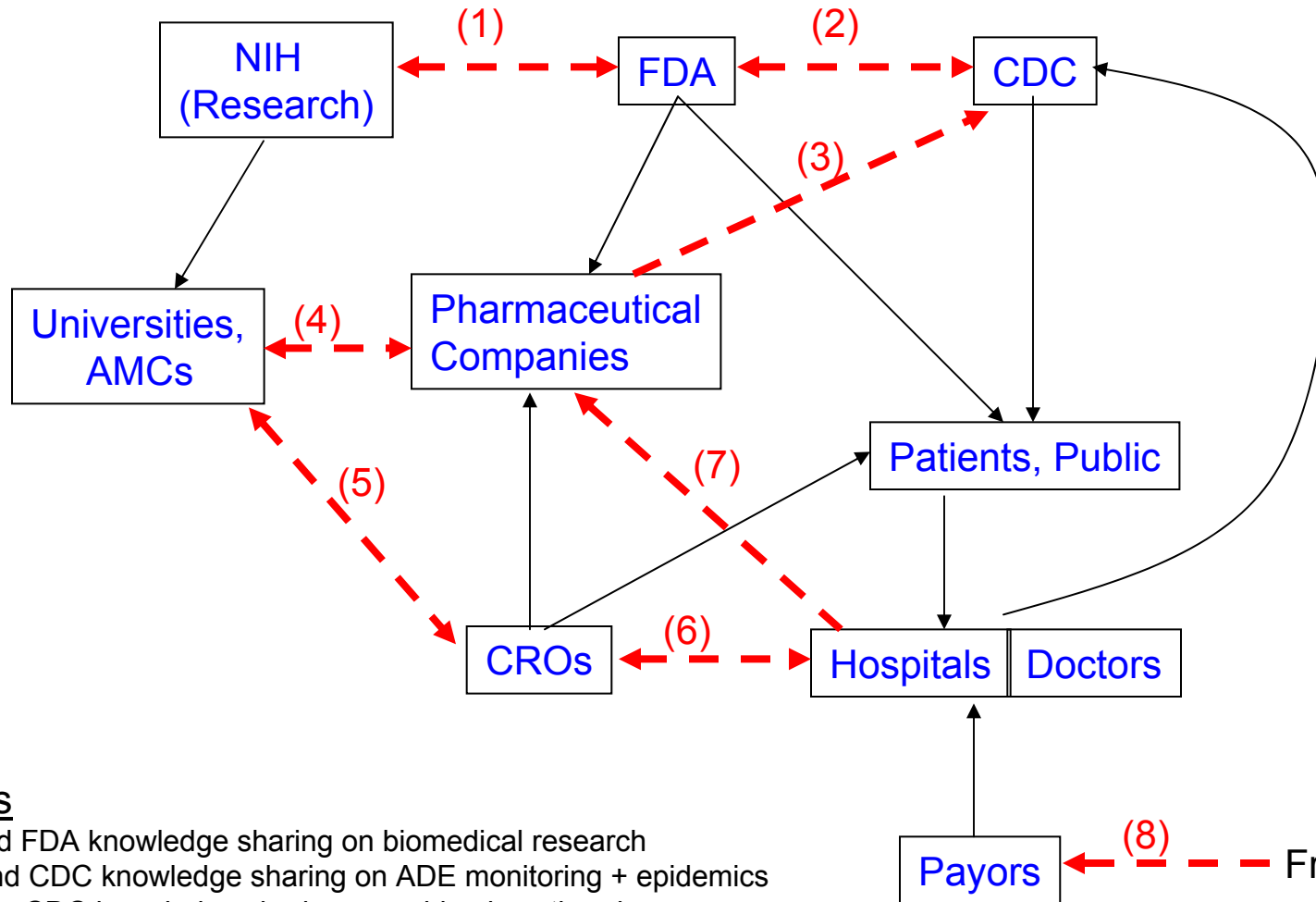


Clinical Practice

Ecosystem: Goal State



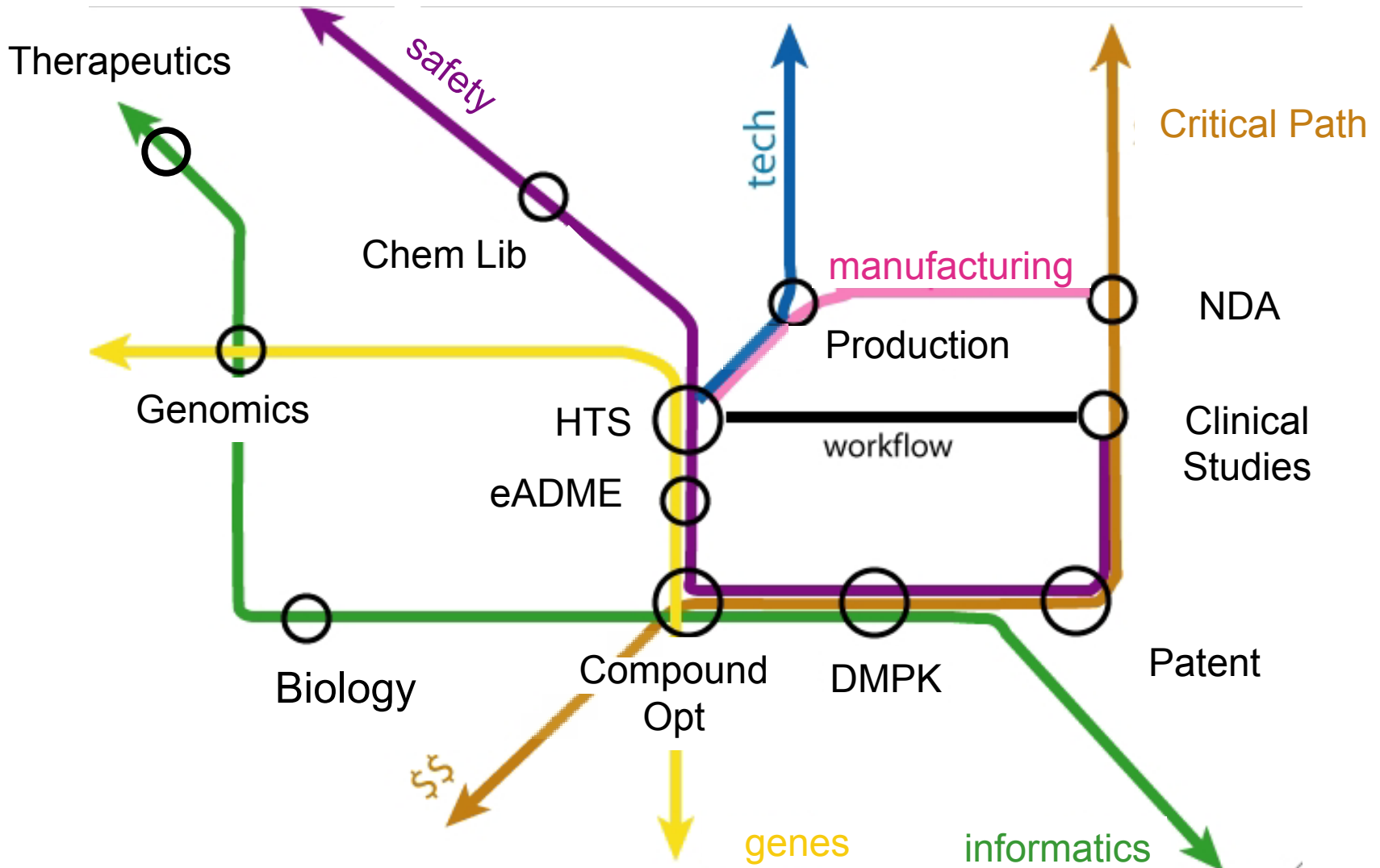
Ecosystem: Goal State



Synergies

- (1) NIH and FDA knowledge sharing on biomedical research
- (2) FDA and CDC knowledge sharing on ADE monitoring + epidemics
- (3) Pharma, CDC knowledge sharing on epidemic outbreaks
- (4) AMCs, Pharma knowledge sharing on clinical and biomedical research
- (5) Universities/AMCs, CROs knowledge sharing on clinical research and trials
- (6) CROs, Hospitals knowledge sharing on identifying patient cohorts
- (7) Pharma, Hospitals knowledge sharing on post market drug surveillance
- (8) Payors get information regarding new conditions, drug efficacies from FDA and CDC

Drug Discovery Application Space



Roadmap

- Need to support a “virtual” view of Goal Ecosystem State.
 - Alignment of key players unlikely
- Need to enable rapid and precision information and knowledge sharing across all players in the Ecosystem.
- Ability to characterize information flows and enable “semantic” standards to achieve this
- Potential role for semantic web technologies
- Governmental Initiatives
 - NIH Roadmap Initiative
 - FDA Critical Path

IMPACT: Likely to leverage efficiencies of operation and scale across the HCLS Ecosystem.

The Bench ↔ Bedside Vision

- Healthcare and Life Sciences: Framework
 - An Information/Knowledge Perspective
- Current Challenges
- The Healthcare Life Sciences (HCLS) Ecosystem
- HCLS Ecosystem Business Drivers

Ecosystem:

Economic, Business and Social Drivers

- Biomedical Research
 - Strong public sector funding for long term research initiatives
- Drug Discovery + Clinical Trials
 - Costly, lengthy, siloized process
 - Lack of economic drivers for certain disease areas
 - Increased spending has corresponded to lesser number of submissions to FDA.
 - Need for data and guidelines to drive adoption of new diagnostic tests in the market and support from payors
- Healthcare
 - Delay in Innovation Adoption
 - Consequent impacts on Patient Safety and Quality of Care
 - Patient Safety: Impact of Medical Errors
 - Pay for Performance
 - Economic and Patient Safety Impacts of Rapid Knowledge Change
 - Electronic Health Record: National Initiative (NHIN, NHS)

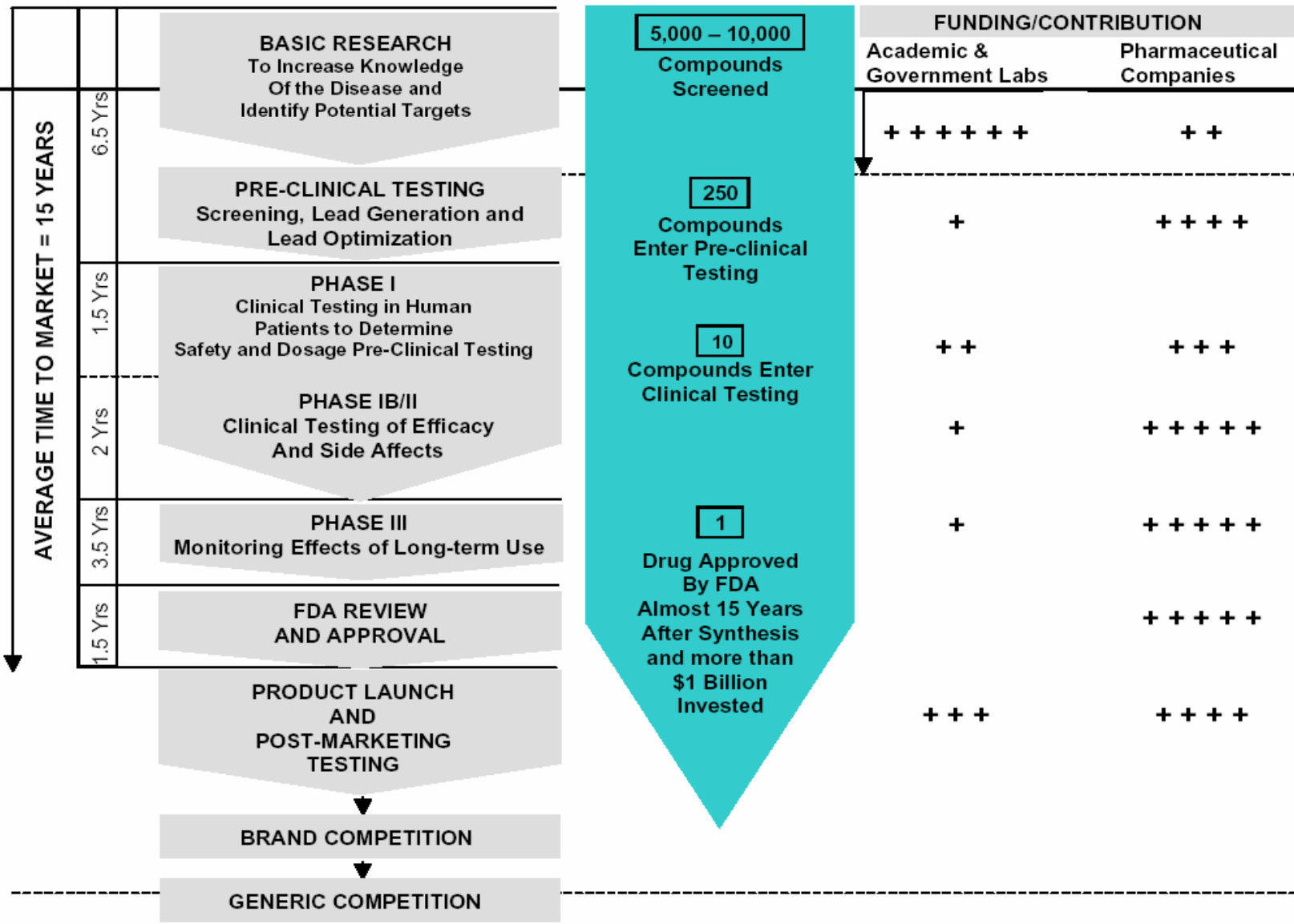
Biomedical Research: Socio-Economic Drivers

- NIH Funding Initiatives
 - Program Announcements (PAs) that seek to initiate new programs of research in various disease and translational areas
 - Request for Applications (RFAs) touching on various fields of biomedical and clinical research typically within the context of a program
 - SBIR/STTR: Small Business, Technology Transfer grants
 - Types of Grants
 - R01: PI Lead Research Grant
 - R03: Small Research Grant
 - R21: Exploratory Research Grant
 - R34: Clinical Trial Planning Grant
 - R56: High Priority, Short Term Project Grant
- Other Funding Agencies: AHRQ, DoD, DoE, Other Foundations
- Sample Funding Opportunities
 - “Innovations in Biomedical Computational Science and Technology”
 - “Translational Research in the Social Neuroscience of Mental Health”

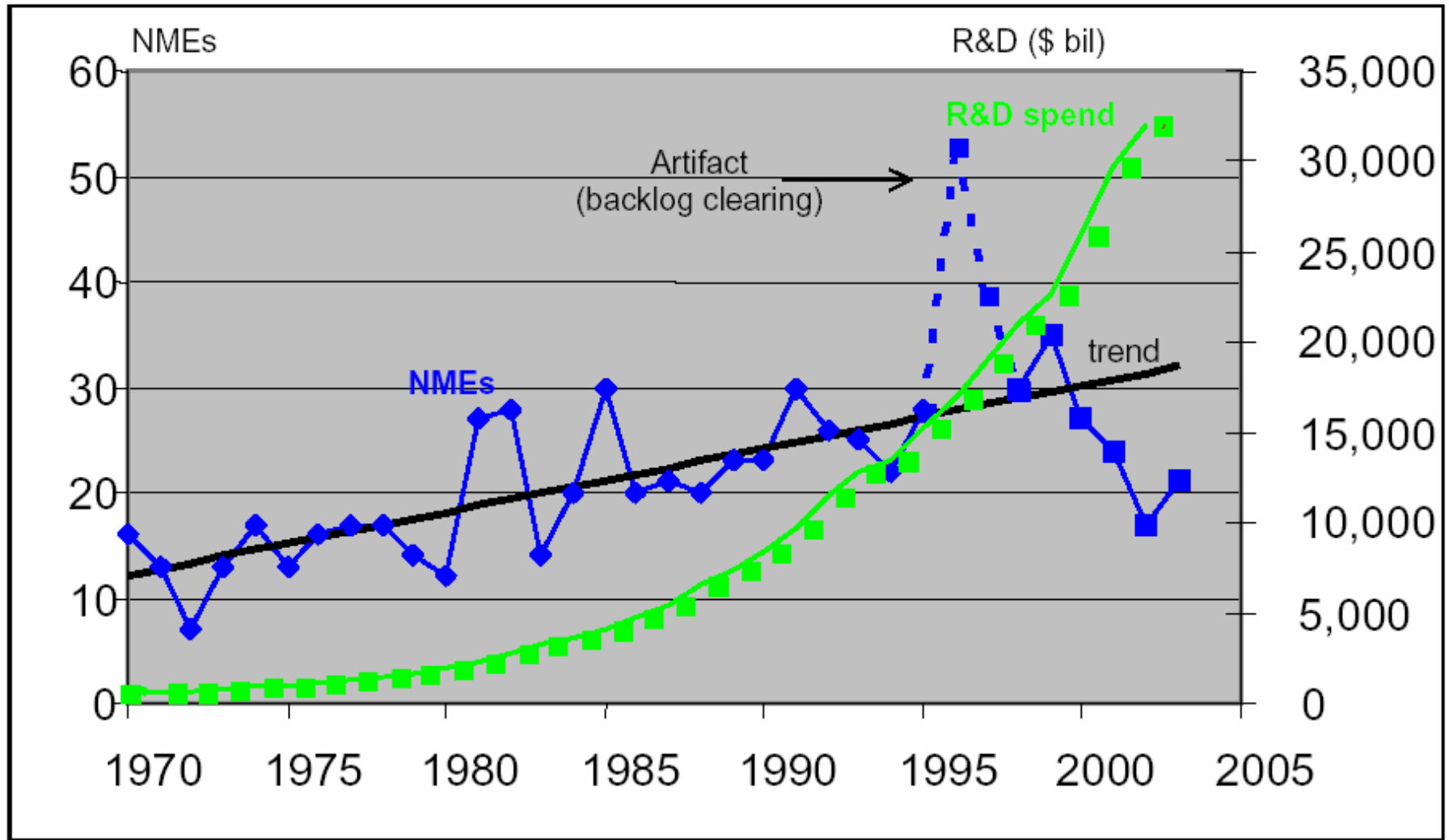
Drug Discovery: Economic Drivers

- Need to find new uses for existing therapies, e.g., antibiotics
- Insufficient economic drivers for certain disease areas
- Post-market surveillance is weak
- Higher spending on Biomedical Research doesn't translate into larger number of drug and biological product submissions
- Increase in investment required for a successful drug launch
- Need for better data and guidelines for using molecular diagnostic tests, e.g., Roche P450 test

Business Drivers: How Medicines make it to the Market

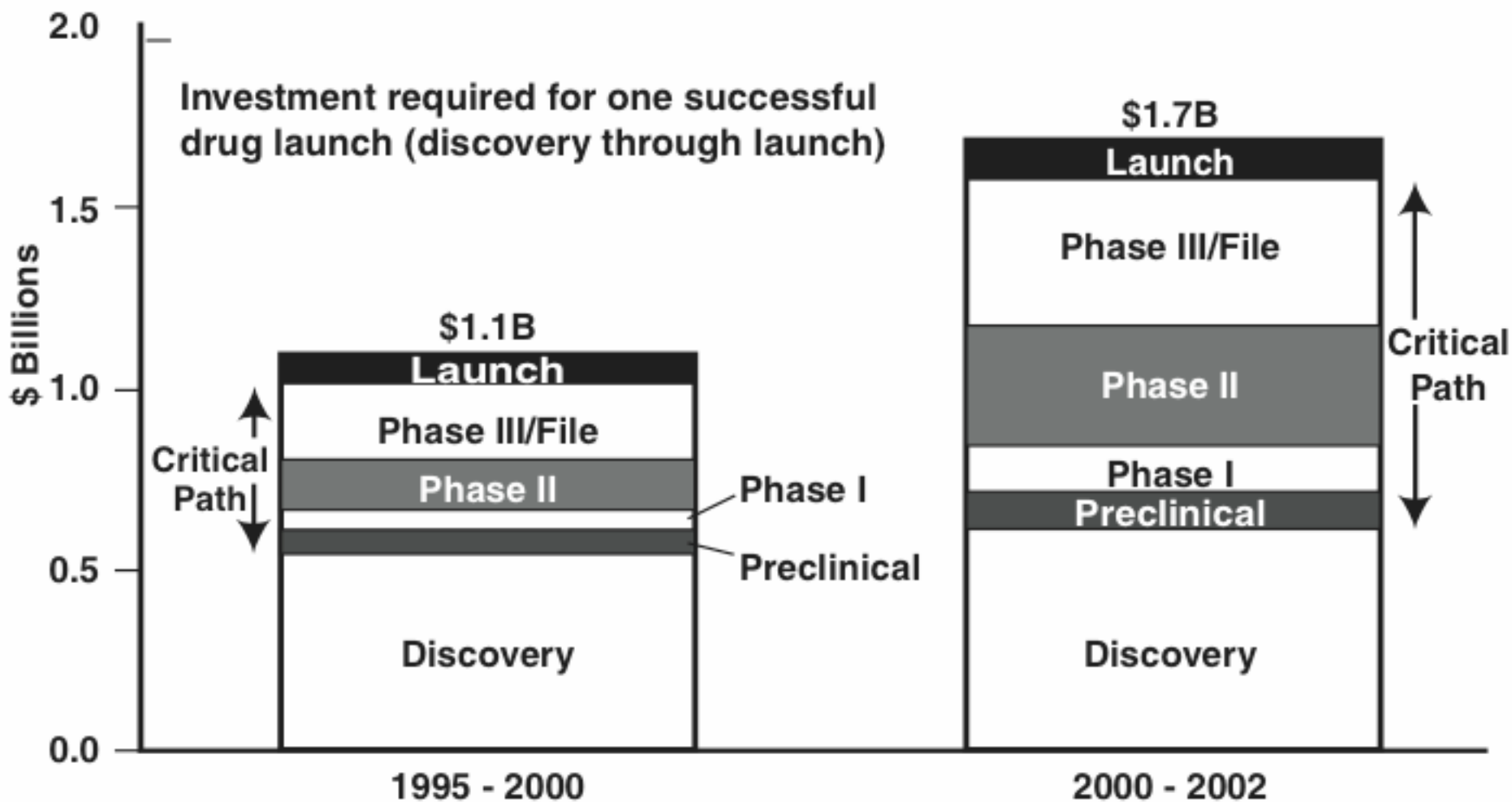


The Pharma R&D Productivity Conundrum



Source: FDA and PhRMA

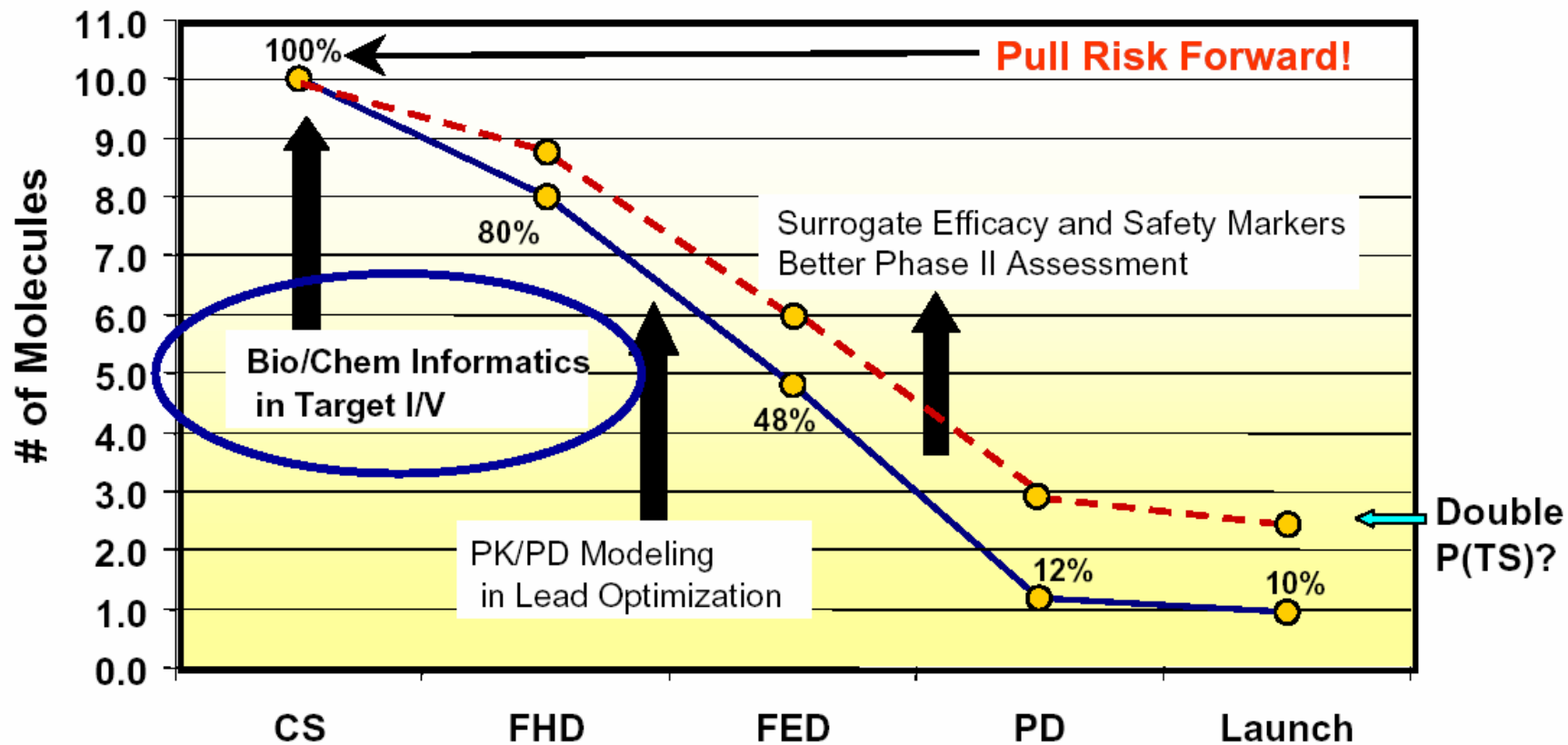
Business Drivers: Drug Discovery



SOURCE: Windhover's In Vivo: The Business & Medicine Report, Bain drug economics model, 2003

from *Innovation or Stagnation*, FDA Report March 2004

Reducing the Development Attrition Curve



Drug Discovery: Business Drivers

The Cytochrome P450 test...
new data to drive drug choice and dose
When do you order this test?
How do you use the test result?



Leading the News: Roche Test Promises to Tailor Drugs to Patients --- Precise Genetic Approach Could Mean Major Changes In Development, Treatment

June 25, 2003

Roche Holding AG is launching the first gene test able to predict how a person will react to a large range of commonly prescribed medicines, one of the biggest forays yet into tailoring drugs to a patient's genetic makeup.

The test is part of an emerging approach to treatment that health experts expect could lead to big changes in the way drugs are developed, marketed and prescribed. For all of the advances in medicine, doctors today determine the best medicine and dose for an ailing patient largely by trial and error. The fast-growing field of "personalized" medicine hopes to remove such risks and alter the pharmaceutical industry's more one-size-fits-all approach in making and selling drugs.



Subject: Genotyping for Cytochrome P450 Polymorphisms to Determine Drug-Metabolizer Status

Policy #: LAB.00013 **Current Effective Date:** 04/28/2005

Status: New **Last Review Date:** 04/28/2005

Description/Scope

Genotyping for cytochrome P450 polymorphisms is a genetic test that is designed to predict a patient's response to a drug, based on the patient's ability to metabolize the drug. It has been proposed that this genetic test be performed prior to the initiation of a drug to permit dosage adjustments based on genetic factors. This policy addresses this laboratory test globally, rather than focusing on specific drugs or clinical situations.

Policy Statement

Investigational/Not Medically Necessary:

Genotyping for cytochrome P450 polymorphisms to determine drug-metabolizer status is considered **investigational/not medically necessary**, including but not limited to, patients initiating therapy with warfarin, phenytoin, antidepressants or antipsychotics.

Rationale

Predicting therapeutic failures or severe adverse drug reactions by genotyping for important polymorphisms in key drug-metabolizing enzymes, receptors, transporters, etc, has the potential to optimize drug choice and/or dose earlier for more effective therapy, avoid serious adverse effects and decrease medical costs. While genotyping for the CYP450 enzymes would only need to be done once per patient and the results could be used to consider other drugs metabolized by the same enzymes, whether or not genotyping is clinically useful would need to be determined for each drug. Even drugs of the same class may variably rely on specific CYP450 enzymes. For example, the plasma level of the selective serotonin reuptake inhibitor (SSRI) fluoxetine is significantly affected by polymorphisms whereas the SSRI sertraline appears to be marginally affected. Additionally, different approaches to genotyping may result in a variety of tests to be validated, each directed toward a subset of enzymes or drugs.

Clinical utility studies of genotyping for well-established brand name and generic drugs are in their infancy. A literature search did not identify any published controlled studies that demonstrated that therapy directed by the results of genotyping resulted in improved patient management. Appropriate outcomes for evaluation might include adverse events, days hospitalized and timed to a clinically significant, predefined change in condition using an appropriate and validated measure. While the potential of pharmacogenomics is intriguing for many clinical applications, it is not yet clear which are most likely to yield clinical benefit in the near future. As this field evolves and matures, and if pre-prescription testing is shown to be of clinical utility for specific drugs, it will be important to establish

- Payors won't pay for this test
- No data on it's value
- Worse, no knowledge base on how to use the test result
- No titration algorithms
- No substitution algorithms
- Test could be as ubiquitous as serum creatinine level for renal function
- Must plan for this in the early discovery process, not after it's on the market

Clinical Practice: Business Drivers

- As much as a 17 year innovation adoption curve from discovery into accepted standards of practice
- Patient safety is a big issue:
 - Even if a standard is accepted, patients have a 50:50 chance of receiving appropriate care, a 5-10% probability of incurring a preventable, anticipatable adverse event
- Healthcare inflation is an issue -- past market utilization management measures have not succeeded
- Economic and social costs of chronic diseases such as diabetes
- Quality measures drive reimbursement

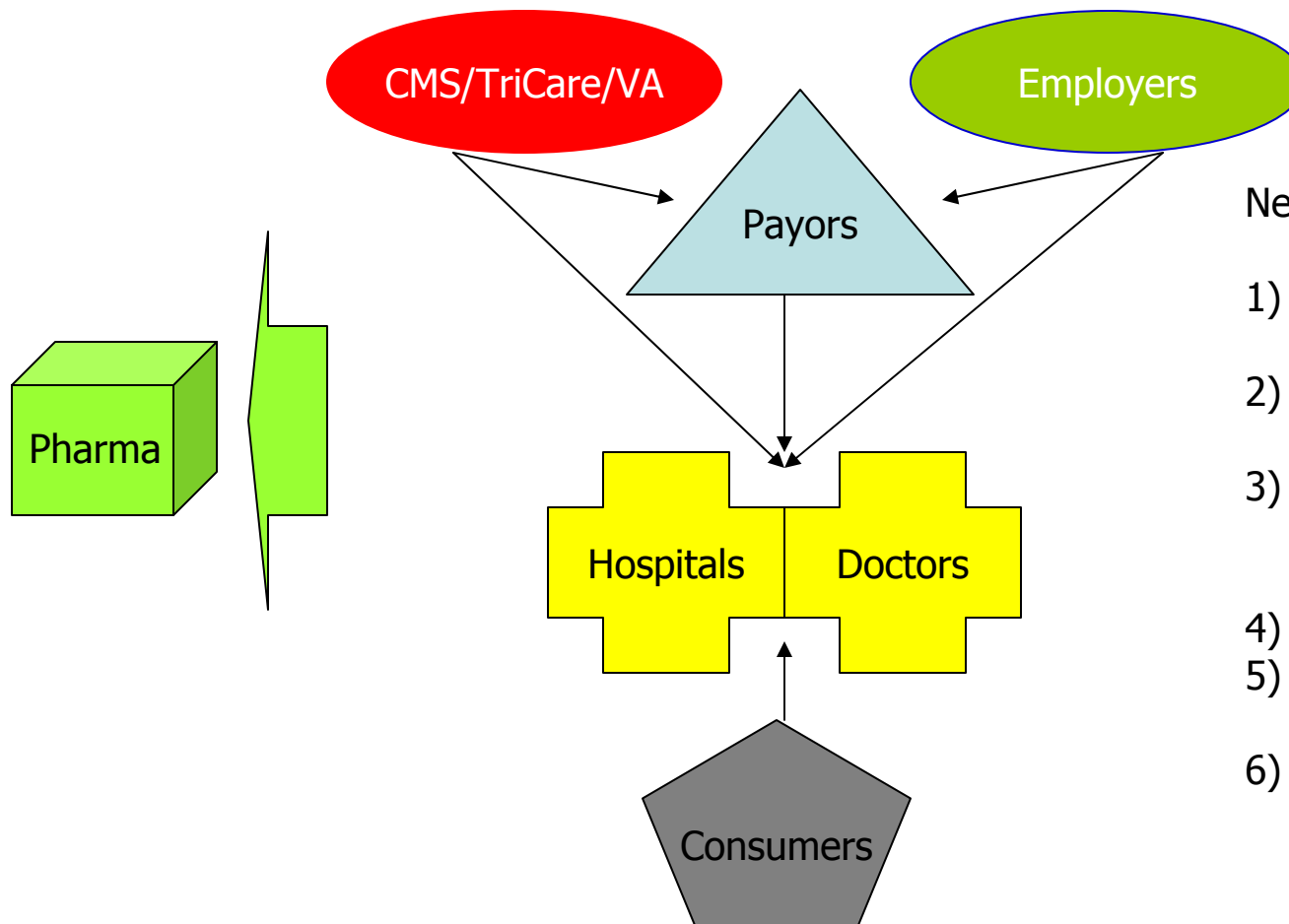
Clinical Practice:

Patient Safety is an issue!



- Medical Errors kill between 44,000 and 98,000 people each year
- 7.3% of hospital admissions incur preventable medication errors
 - 66% of these were not intercepted
 - 25% resulted in patient harm
 - 360 preventable Adverse Drug Events for a hospital with 20,000 annual admissions, almost 1 ADE per day

Clinical Decision Support: Business Drivers



New Healthcare Business
Climate in the US:

- 1) Give consumers greater financial responsibility
- 2) Reduce Hospitalizations and cost per hospitalization
- 3) Increase quality performance – value based purchasing
- 4) Reduce drug expenditures
- 5) Reduce imaging study expenditures
- 6) Increase wellness/reduced sick days

Clinical Practice: Business Drivers

Today's Globe Opinion Magazine Education Science NECN Special reports Obituaries

[HOME](#) > [NEWS](#) > [LOCAL](#) > [MASS.](#)

State posts data on doctors

The Boston Globe

Website provides surgery statistics

By Liz Kowalczyk, Globe Staff | February 1, 2006

Before b Today's Globe Opinion Magazine Education Science NECN Special reports Obituaries

answer t

[HOME](#) > [NEWS](#) > [LOCAL](#) > [MASS.](#)

The stat the publi performi including

Plan would tie copayments to doctors' rankings

The Boston Globe

Research certain d experien figures to consum

Lower fee for using 'best value' physicians

By Jeffrey Krasner, Globe Staff | January 27, 2006

But phys linking q on the s

The agency that oversees health insurance for 144,000 state workers wants to launch a program to control runaway healthcare expenses: a ranking of doctors' quality and efficiency that would be tied to lower copayments for patients who seek care from higher-rated doctors.

The web informati hospitals' surgery, endarter year that operati patient v

The Group Insurance Commission this morning will examine proposals it requested from Harvard Pilgrim Healthcare, Tufts Health Plan, Health New England, Fallon Community Health Plan, and other insurance companies. A key feature of each proposal is variable copayments. For example, under the Harvard Pilgrim and Tufts scenarios patients would pay \$15 to visit a doctor who is rated highly and \$25 to see one with a lower rating.

ARTICLE TOOLS

 [PRINTER FRIENDLY](#)

 [SINGLE PAGE](#)

 [EMAIL TO A FRIEND](#)

 [MASS. RSS FEED](#)

 [MOST EMAILED](#)

MORE:

[Globe City/Region stories](#)
[Latest local news](#)

Dolores L. Mitchell, executive director of the commission, said ranking doctors would be an important step toward making medical care more transparent and keeping premiums down.

"We're saying to our members, these specialty doctors and primary-care physicians represent the best value, and we'll reward you if you use them," Mitchell said.

But many of the state's physicians oppose the idea, saying there is no widely accepted measure of quality care or efficiency. They also warn about the risks of adopting a new program with such a large group of beneficiaries without first conducting pilot studies.

- Employers and consumers are now paying for performance

- They will not wait for the data to be right or fair because they believe if they wait, it never will be

- Instead, they are using the process of rewarding performance to force the healthcare providers to "make the data right"

- Defined Contribution lays even greater purchasing responsibility at the door of the consumer

- Pay for performance and defined contribution will increase EHR adoption

Business Drivers: EMR

- Pay for performance will provide motivation for capturing and storing accurate data
 - Will “force” the providers to make the data “right”
 - EMRs will provide a cost-effective tool for this
- Need for computing quality measures that will drive reimbursement
- Provides the substrate for implementing decision support
- Socio-governmental driver: Federal Government initiatives and incentives
 - NHIN effort in the US
 - NHS modernization in the UK

Diabetes: Economic and Social Impacts

- Epidemic, associated with obesity
- Estimated avg \$21,000/year per diabetic employee in absenteeism, disability and medical costs (study of 6 employers with 375,000 employees)
- Creation of Quality measures that drive reimbursement
 - Maintain HbA1c <7 (diet, oral agents and/or insulin)
 - If Renal Disease and no contraindication, should be on ACE inhibitor or ARB
 - If lipid disorder and no contraindication, should be on a statin like lipitor

When Knowledge Changes...

Feb. 1, 2006 5:37pm

CBS NEWS

SEARCH CBSNews.com The Web

[Home](#) | [U.S.](#) | [World](#) | [Politics](#) | [SciTech](#) | [Health](#) | [Entertainment](#) | [Business](#) | [Opinion](#) | [Strange News](#) | [Sports](#)
[The Early Show](#) | [CBS Evening News](#) | [48 Hours](#) | [60 Minutes](#) | [CBS Sunday Morning](#) | [Face The Nation](#) | [Up To Th](#)

HEALTHWATCH

- [Health Main Page](#) >
- [Interactives](#) >
- [Early Show: Health News](#) >
- [Early Show: Shape Up](#) >
- [Early Show: Emily Senay](#) >
- [Audible.com Downloads](#) >
- [HealthWatch Video](#) >
- [CBS Cares](#) >



[Q&A: Understand Risks From Pain Relievers](#)

To assist patients with the recall of one of the three COX-2 drugs yanked from the market, Merck offers the following information

[E-MAIL THIS STORY](#) [PRINTABLE VERSION](#)

Merck Yanks Vioxx From Shelves

NEW YORK, Sept. 30, 2004



(CBS/AP)

(CBS/AP) Vioxx, the blockbuster arthritis drug heavily promoted on TV and taken by tens of millions of people, was pulled from the market by its maker Thursday after a study found it doubled the risk of heart attacks and strokes.

Experts advised patients to immediately stop taking Vioxx and talk to their doctors about alternatives.

"Given the availability of alternative therapies, and the questions raised by the data, we

[1](#) [2](#) [3](#)

[PREVIOUS IMAGE](#)

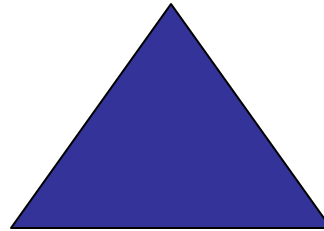
[NEXT IMAGE](#)

How quickly can you change the content of your rules, order sets, templates, and reports?

CDS + Knowledge Acquisition: Business Drivers

Clinical Standardization

*Standards of Practice,
Role/Venue Requirements
Billing/Regulatory Requirements*



Improvisation

Patient Preferences

User Personalization

*End-user role, workflow preferences
Learning and User-defined*

Outline

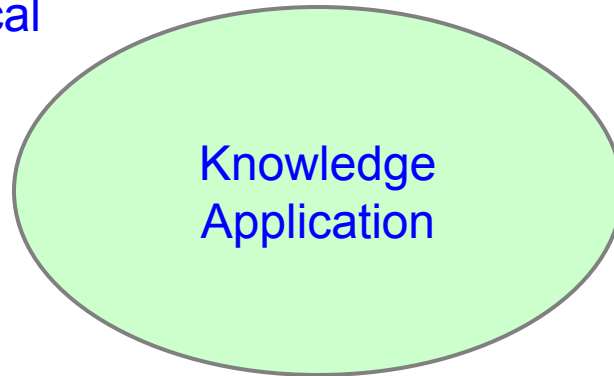
- The Bench ↔ Bedside Vision
- Role of Data, Knowledge and Semantics in the HCLS Ecosystem
- Functional Requirements
 - Metadata-based Semantic Annotation
 - *Recombinant Data*: Information Aggregation and Integration
 - Ontology-driven Decision Support
 - Knowledge Update and Maintenance
- Conclusions

Role of Data, Knowledge and Semantics in the HCLS Ecosystem

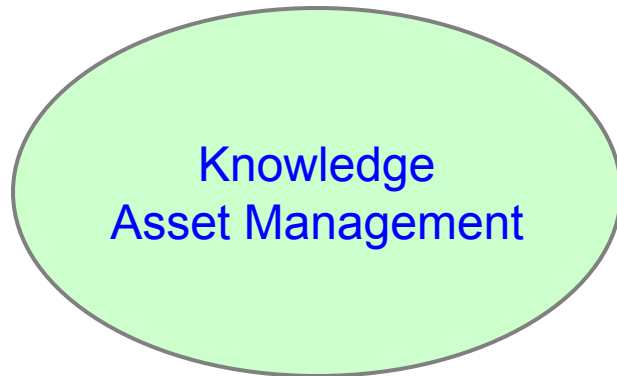
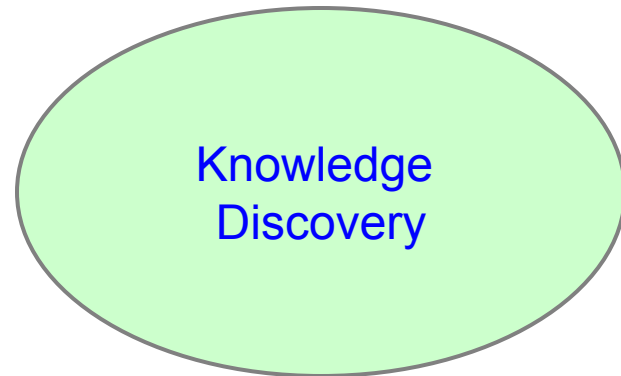
- Knowledge Feedback Loop
- Bench ↔ Bedside Vision:
 - Knowledge Dependent Architecture
 - Use Case Flows
 - Personalized Medicine
 - Drug Discovery and Development
- The Data Semantics Continuum
- The Knowledge Semantics Continuum

HCLS: The Knowledge Feedback Loop

E.g., Application of Clinical
Decision Support Rules

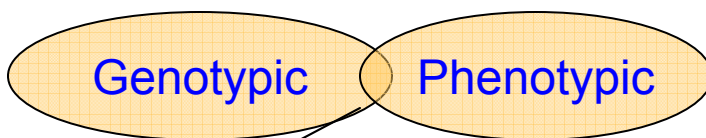
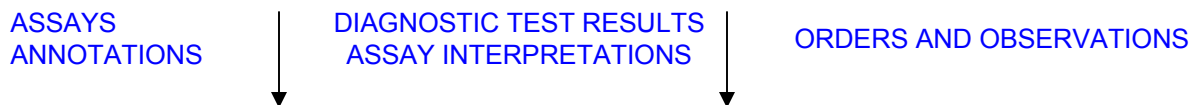
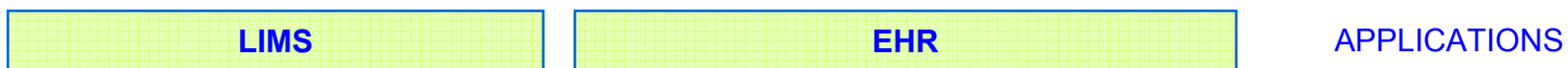
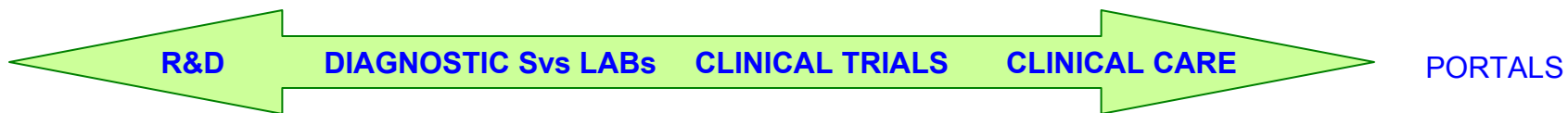


E.g., Analysis of clinical
Care transactions for
New or updated Rules



E.g., Creation and Maintenance
of Clinical Decision Support Rules

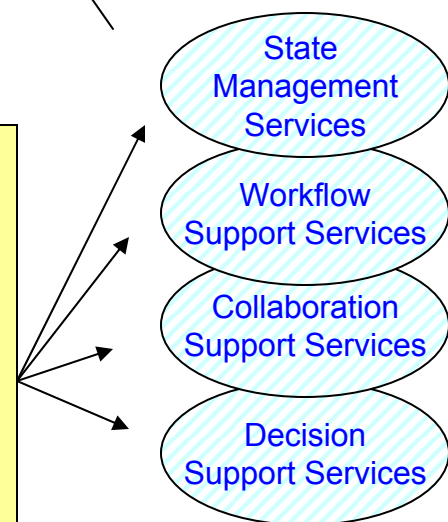
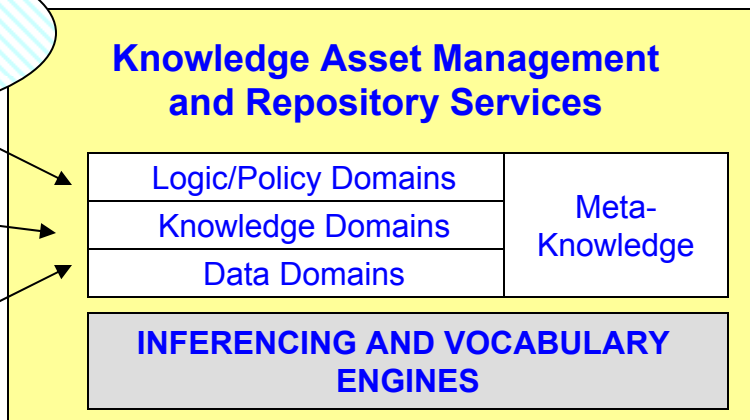
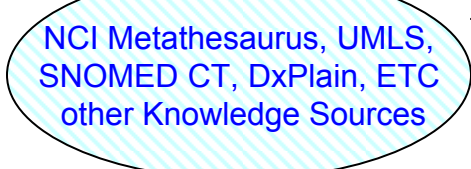
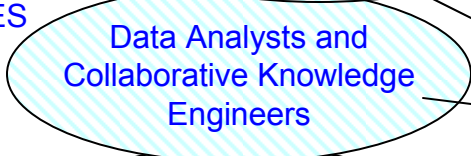
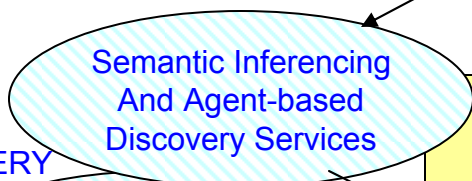
Bench ↔ Bedside: Knowledge Dependent Architecture



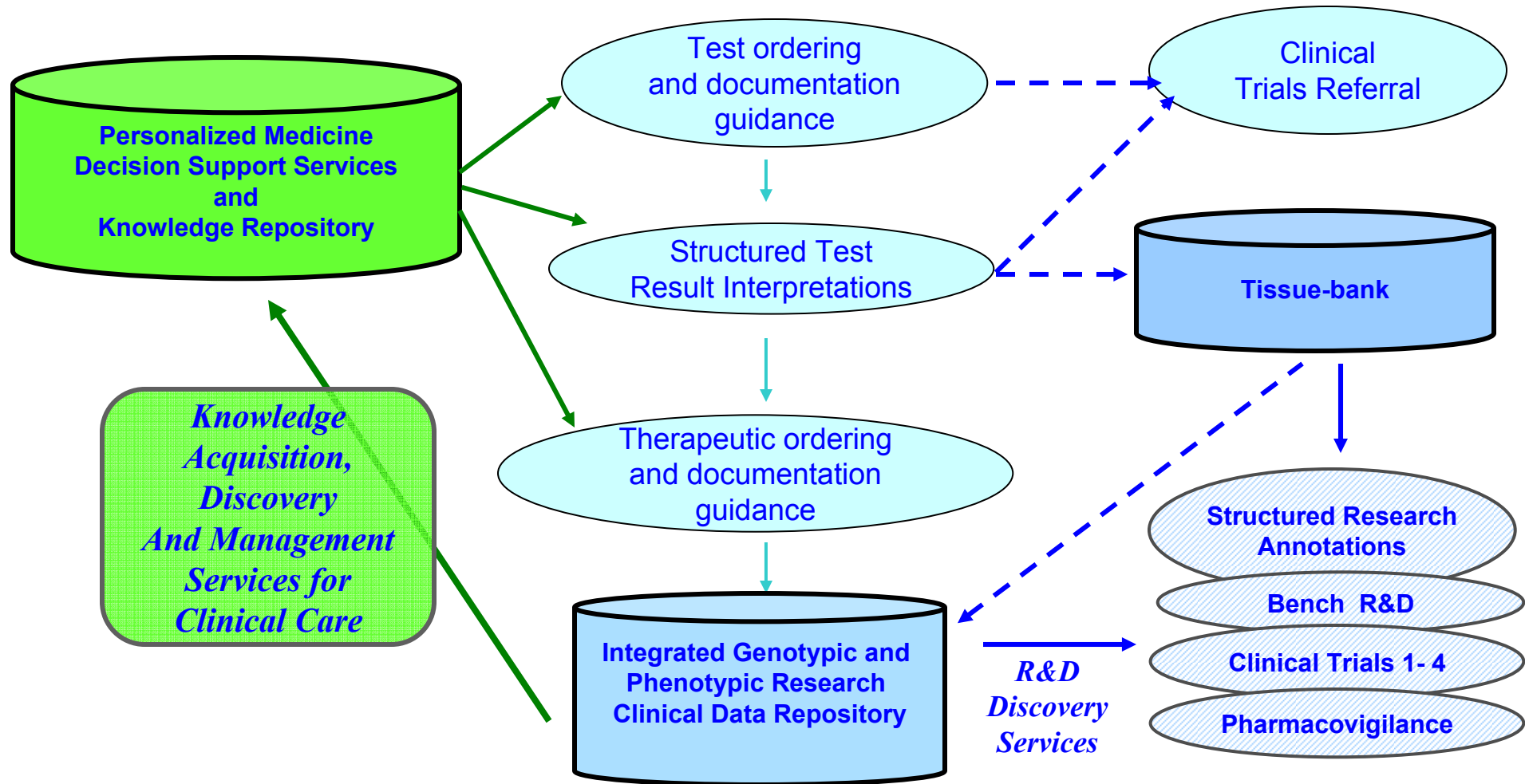
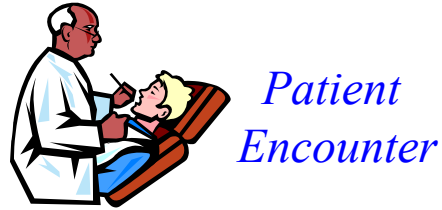
KNOWLEDGE and WORKFLOW DELIVERY SERVICES FOR ALL PORTAL ROLES

DATA REPOSITORIES AND SERVICES

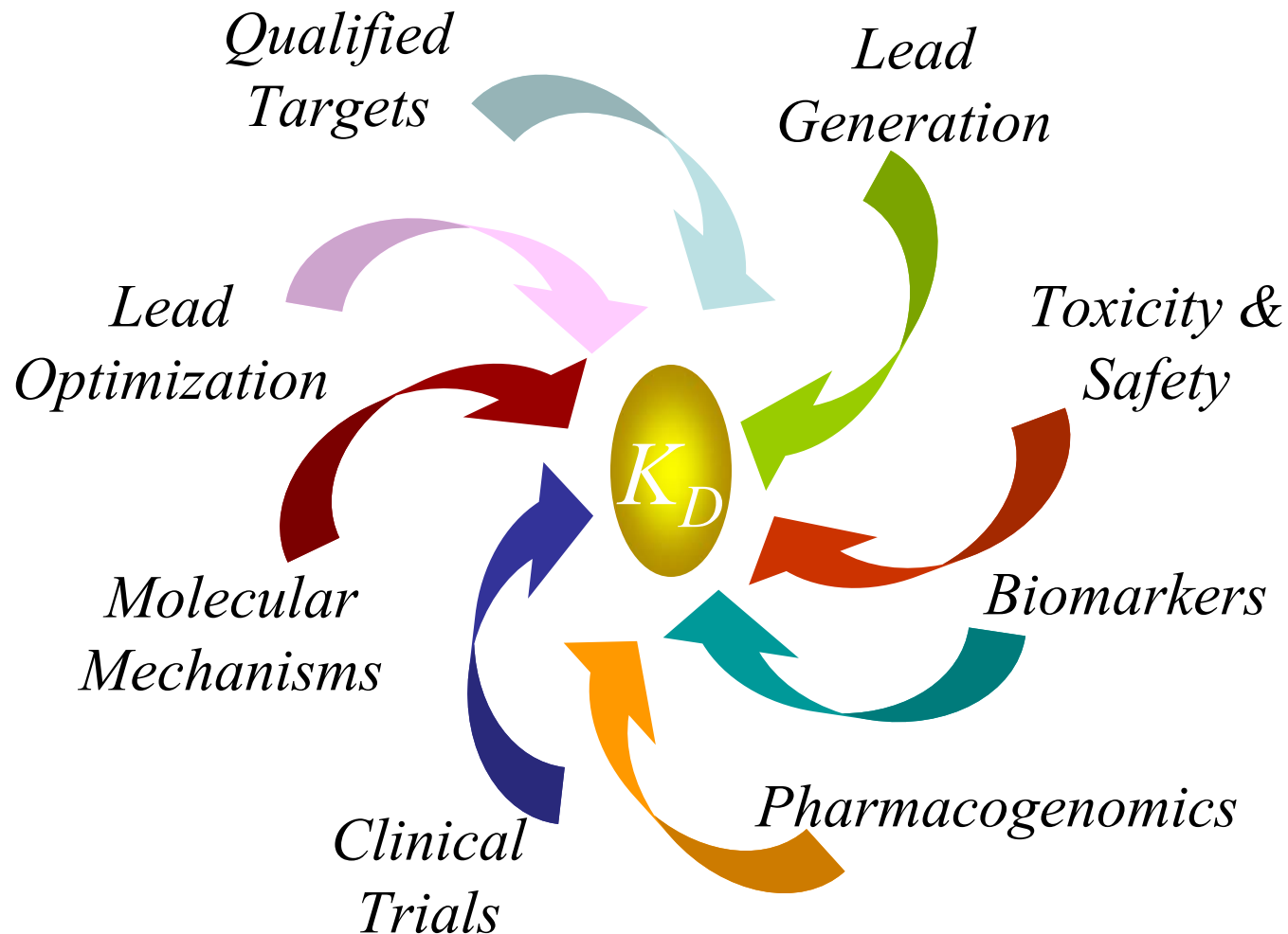
KNOWLEDGE ACQUISITION AND DISCOVERY SERVICES



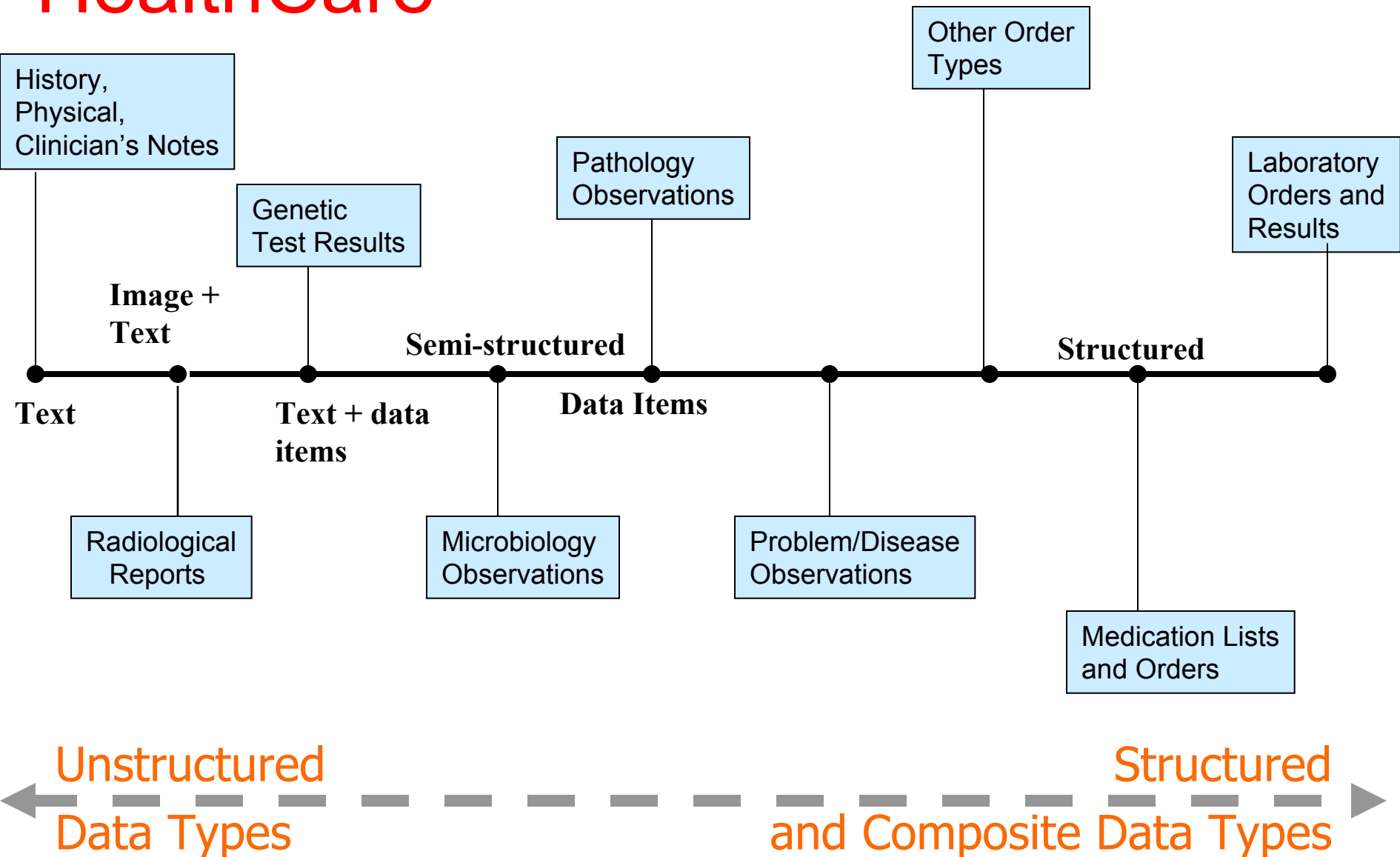
Bench ↔ Bedside: Use Case Flow based on Shared Semantics



Use Case Flow: Drug Discovery and Development



The Data Semantics Continuum: HealthCare



Healthcare Data: Clinician Notes

HISTORY OF PRESENT ILLNESS: <X> had a fall on September, 15, bruised and cut her face and lip and required some sutures. A review of her laboratory tests reveals an **albumin, creatinine ratio of 139** the test done in 06/04. She is currently on **75 mg of captopril** and plans are to retest microalbumin. Again, revisit in three months. Nocturia one to two times nightly. She does have occasional paresthesias in her upper extremities.

MEDICATIONS: **500 mg of metformin** one in the morning, one at noon and 1000 mg in the evening and Avandia 4 mg a day. In addition, to her **captopril** she is on **Norvasc**.

REVIEW OF SYSTEMS: No **chest pain**, shortness of breath, or **cardiovascular symptoms** at this time. She has not had her eyes examined for the past year and will make an appointment at the **BI-Deaconess Medical Center**.

PHYSICAL EXAMINATION: No change. Her **weight is 223 pounds**. **Blood pressure 122/80**, **pulse 75** regular. Blood sugar this morning 121.

PLAN: Revisit in three months, continue medication.

<Y>, M.D.

Dictated By: <Y>

eScripton document:1-5763294 UF

DD: 10/05/04

DT: 10/05/04

DV: 10/05/04"

Mostly unstructured data divided into broad sections

Opportunity: Identification and tagging using standardized concepts

Healthcare Data: Radiological Report

Exam Number: random number

Report Status: Final

Type: CT SCAN ABDOMEN W/O CONTRAST

Date/Time: date:time

Ordering Provider: TURCHIN, ALEXANDER M.D.

Report Below from Associated Order A07439985: CT SCAN PELVIS W/O CONTRAST

HISTORY: Please evaluate whether the right adrenal mass and the pancreatic
REPORT

HISTORY: 5 cm right adrenal mass. Pancreatic mass.

TECHNIQUE: CT scan of the abdomen and pelvis was performed following the
administration of oral contrast. Intravenous contrast was not administered.

PRIOR STUDIES: CT scan abdomen and pelvis dated <Date>.

FINDINGS:

LUNG BASES: Subsegmental atelectasis is seen in the right lung base and slight
atelectatic changes are noted in the left base. Small area of high
attenuating material is seen in the medial aspect of the right lower lobe
which may be related to prior aspiration. This finding is unchanged.

ABDOMEN: Evaluation of the solid abdominal viscera is limited without intravenous
contrast material. The liver, gallbladder, spleen, and left adrenal gland are
unremarkable. There is a 4.4 x 2.7 cm cystic mass in the pancreatic body,
essentially unchanged from the prior examination. The pancreatic duct is
normal in caliber.

PELVIS: Heavy vascular calcifications are seen in the abdominal
aorta and its major branches.

OSSEOUS STRUCTURES: Left hip enthesopathy is noted. There are no suspicious
lytic or blastic lesions. Degenerative changes are seen throughout the
thoracolumbar spine.

IMPRESSION:

1. 4.6 x 2.9 cm cystic mass within the pancreatic body. Differential diagnostic
considerations include cystic neoplasms such as serous microcystic adenoma
or mucinous macrocystic adenoma. Further evaluation with an MRI examination
is recommended.
2. Stable right adrenal mass which may also be evaluated during the time of the
pancreatic MRI.
3. Several indeterminate left renal cystic lesions.

END OF IMPRESSION:

eScription document:5-7347531 ABBWH

RADIOLOGISTS:

SIGNATURES: GIRSHMAN, JEFFREY, MD(R)

MORTELE, KOENRAAD, MD(T)

Still primarily unstructured data.
However sections are more finely
defined (e.g., Findings – Lung Bases,
Abdomen, Osseous Structures, etc.)
and presence of “facts”, e.g., type of
report.

Opportunity:
Grounding of facts in standardized
Ontologies
Tagging of sections and fragments of
unstructured text using standardized
concepts

Healthcare Data: Pathology Observations

Accession Number: randomnumber
Report Status: Final
Type: CytologySpecimen
Type: FINE NEEDLE ASPIRATION, Right THYROID
Procedure Date: date:time
Ordering Provider: BRIAN W KIM M.D.
CASE: randomnumber
PATIENT: patientname
Cytotechnologist: Janet A Cronin, CT(ASCP)
Pathologist: Xiaohua Qian, M.D., Ph.D.
FINE NEEDLE ASPIRATION, RIGHT THYROID
FINAL CYTOLOGIC INTERPRETATION
INTERPRETATION:
NO MALIGNANT CELLS IDENTIFIED.
DIAGNOSIS:
Benign-appearing follicular cells and colloid.
Cyst lining cells.
The findings are consistent with a benign thyroid nodule.
CLINICAL DATA
GROSS DESCRIPTION
50cc colorless.
MATERIALS
Total slides: 2
By his/her signature below, the senior physician certifies that he/she personally conducted a microscopic examination of the described specimen(s) and rendered or confirmed the diagnosis (es) related thereto.
Final Diagnosis by Xiaohua Qian M.D., Ph.D.,
Electronically signed on 4/25/2006

Very Structured in Nature
A collection of data items and
Some free text.

Opportunity:
Link “properties” and “values” to standardized
Namespaces and concepts
Tagging for fragments of unstructured text.

Healthcare Data: Microbiology Observations

Specimen: randomnumber
Collected date:time
Received date:time
Ordering Provider:
Specimen Group: NOSE/NASOPHARYNX
Specimen Type: NARES FOR MRSA MRSA CULTURE
Reported: date:time
NO METHICILLIN-RESISTANT STAPHYLOCOCCUS AUREUS ISOLATED

Mostly structured items

Opportunity: Link “properties” and values to standardized namespaces and contexts.

Healthcare Data: Genetic Test Results






















```
<?xml version="1.0" encoding="UTF-8"?>
<geneticTestOrder xmlns="http://www.partners.org/genetics"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.partners.org/genetics geneticTestOrder_1_0.xsd">
  <order>
    <indicationCode code="DA-95200" codeSystemName="SNOMED" text="Sensorineural hearing
loss"/>
    <indicationClassCode code="Constitutive" />
    <testCategoryCode code="Diagnostic" />
  </order>

  <fulfillment>
    <fulfillmentRecord referenceSystemId="PowerPath" recordId="BL-05-N00748" />
    <testReference referenceSystemId="HPCGG GVAD" testId="35" testVersionId="3" />
    <effectiveTime>20051025</effectiveTime>
    <overallResult>Positive</overallResult>
    <variants count="2">
      <sequenceVariant>
        <dnaChange level="Intragenic" type="Deletion">
          <name>35delG</name>
        </dnaChange>
        <aminoAcidChange type="Frame Shift"/>
        <variantReference referenceSystemId="HPCGG GVAD" variantId="45873" />
        <genomicLocation>
          <chromosome>13</chromosome>
          <chromosomeRegion>13q12</chromosomeRegion>
          <gene>GJB2</gene>
          <geneRegion>Exon 1</geneRegion>
        </genomicLocation>
        <interpretationCode code="Pathogenic" />
        <allelicStateCode code="Homozygous" />
        <significanceCode code="Non-incidental" />
      </sequenceVariant>
    </variants>
  </fulfillment>
</geneticTestOrder>
```

Mostly structured items

Opportunity: Link “properties” and values to standardized namespaces and contexts.

Healthcare Data: Problem List

| Select Desktop Pt Chart: P | |
|---|-------------------------|
| Active Problems (21) Inactive Prob | |
| Add New Favorites Re-order Problems | |
| Problem Description | |
|  | HYPERCHOLESTEROLEMIA |
|  | CORONARY ARTERY DISEASE |
|  | THYROID NODULES |
|  | ARTHRITIS |
|  | DEPRESSION |
|  | HX HYPERCALCEMIA |
|  | DERMATITIS |
|  | HX HYPERURICEMIA |
|  | HX PYELONEPHRITIS |
|  | DIZZINESS/ S/P ENT E... |
|  | COLON POLYPS |
|  | BREAST CYST |
|  | Total knee replacement |
|  | Meningioma |
|  | Hypertension |
|  | Breast cancer |
|  | Pulmonary embolism |
|  | * Pancreatic mass |
|  | Hemoptysis |
|  | Obesity |
|  | Chronic renal dysfun... |

Structured and in some cases “coded” data

Opportunity: Link to standardized ontologies such as Snomed
Create complex definitions of concepts

Healthcare Data: Medication List

| Active Meds (21) | | Inactive Meds (31) | | Active Non-Meds (1) | | Inactive Non-Meds (0) | | Discharge (15) | | All | Practice | My |
|---|--------|---------------------|-----------|---|-------------------------------------|-----------------------|---------------|----------------|----------|-----------------|----------|----|
| Enter new medication... | | Add New | Favorites | Renew | Discontinue | Verified | Pharmacy Info | | | | | |
| Patient's Payor | | | | | | | | | | Select Payor... | | |
| Allergies: IV Contrast - Anaphylaxis | | | | | | | | | | | | |
| <input type="checkbox"/> | Co-pay | Medication | Rx | Sig | Special Instructions | Dispense | Rfl | Start Date | End Date | Orig. Date | | |
| <input type="checkbox"/> | | Allopurinol | Rx | 100MG TABLET take 1 Tablet(s) PO BID | | 90 TABL... | 3 | 12/13/02 | | 12/13/02 | | |
| <input type="checkbox"/> | | Arimidex | Rx | 1 MG (1MG TABLET take 1) PO QD | | 90 Tabl... | 3 | 03/20/02 | | 03/20/02 | | |
| <input type="checkbox"/> | | Aspirin | | 81 MG (81MG TABLET take 1) PO QD | | Tabl... | | 01/23/06 | | 01/23/06 | | |
| <input type="checkbox"/> | | Calcitriol | Rx | 0.75 MCG (0.25MCG CAPSULE take 3) PO QD | x 90 days | 90 Caps... | 3 | 09/28/05 | 12/26/05 | 09/28/05 | | |
| <input type="checkbox"/> | | * Calcitriol | | PO | | | | 04/21/05 | | 04/21/05 | | |
| <input type="checkbox"/> | | Cozaar | Rx | 100 MG (100MG TABLET take 1) PO QD | x 90 days | 90 Tabl... | 3 | 03/16/06 | 06/14/06 | 09/21/04 | | |
| <input type="checkbox"/> | | Ecasa | Rx | 81 MG PO QD | | 60 | 3 | 12/28/01 | | 09/02/98 | | |
| <input type="checkbox"/> | | Kayexalate | Rx | 15 GM (15G/60ML ORAL SUSP ML) PO QD | | 1 Mont... | 1 | 02/16/06 | | 02/16/06 | | |
| <input type="checkbox"/> | | Lasix | Rx | 80 MG (80MG TABLET take 1) PO BID | x 30 days | 60 Tabl... | 3 | 01/17/06 | 02/16/06 | 11/07/01 | | |
| <input type="checkbox"/> | | Lipitor | Rx | 80MG TABLET PO QD | x 90 days | 90 Tabl... | 3 | 02/02/04 | 05/02/04 | 02/02/04 | | |
| <input type="checkbox"/> | | Lisinopril | | 80 MG PO QD | | 60 | 3 | 12/28/01 | | 09/20/00 | | |
| <input type="checkbox"/> | | Neurontin | Rx | 600 MG PO TID | | 60 | 3 | 12/13/02 | | 10/13/99 | | |
| <input type="checkbox"/> | | Nitroglycerin 1/150 | Rx | 1 TAB SL x3 q5 min PRN chest pain | if no relief after 3 doses call 911 | 100 | 3 | 08/02/02 | | 07/24/02 | | |

Structured data consisting of properties and values

Opportunity: Link to standardized drug classification ontologies and definitions of normality and abnormality

Healthcare Data: Medication Orders

Select Desktop Pt Chart: Medications Oncology Custom Reports Admin Sign Results ? Resource Popup

Allergies
IV Contrast - Anaphylaxis

ALLOPURINOL 100MG TABLET
Sig: 1 Tablet(s) PO BID

Basic Variable Alternate

| Dose | Strength & Form, Take | Frequency |
|----------------------------------|------------------------------|-----------|
| <input checked="" type="radio"/> | 100MG TABLET take 1 Tablet | BID |
| <input type="radio"/> | 50 MG 100MG TABLET take 0.5 | QD |
| <input type="radio"/> | 100 MG 100MG TABLET take 1 | QD |
| <input type="radio"/> | 150 MG 300MG TABLET take 0.5 | QD |
| <input type="radio"/> | 200 MG 100MG TABLET take 2 | QD |
| <input type="radio"/> | 300 MG 300MG TABLET take 1 | QD |

PRN: Patient Educated
Duration: day(s) No Substitutes
Dispense: 90 Tablet(s) Expire
Refills: 3
Start Date: 12/13/2002 Orig. Date: 12/13/2002
End Date:

Special Instructions

Comments (This will not print on prescription)

Add to My Practice Favorites as:

Rx Print/Fax no Rx

History Discontinue Link to Problems Ok Renew OK & Sign Cancel

Structured and Composite Data containing multiple data items (properties, values)

Opportunity: Link to standardized information models and ontologies for interoperability and decision support

Healthcare Data: Lab Orders

BICS Terminal Emulator - OMABICS12.partners.org

ViewOrders PtLookup Feedback Help Goodbye

OETEST.WILHEMINA 76F 16356792 Adm: 12/22/03 Room: 17A-501

COMMON LAB ORDERS

| | | |
|--------------------------------|----------------------|------------------------|
| \$66 [X]A CBC | \$15 [X]H BUN | \$53 []IR Amylase |
| \$33 [X]B DIFF | \$17 [X]J Creatinine | \$53 []IS Lipase |
| \$73 []IC PT | \$16 []K K+ | \$17 []IU Bilirubin |
| \$73 []ID PTT | \$15 []L Glucose | \$7 []IV HCT |
| \$33 []IE U/A + Sed | \$107 []M Blood Gas | \$87 []IW Urine C+S |
| \$225 []IF Comp Metabolic(12) | \$35 []N Magnesium | \$241 []IX Type+Hold |
| \$128 []IG Basic Metabolic(7) | \$17 []P Calcium | \$113 []IY BC x 2 |
| | \$96 []Q CK + MB | \$15 []IZ F'stick Glu |

<NEXT TEST STAT> NO

BUN \$15
CRE \$17 TOTAL= \$131

0 Other Lab Test Names (separated by commas)

[]

These labs will be drawn on times listed below.

T Collection Time: NEXT AVAILABLE

I Instructions:

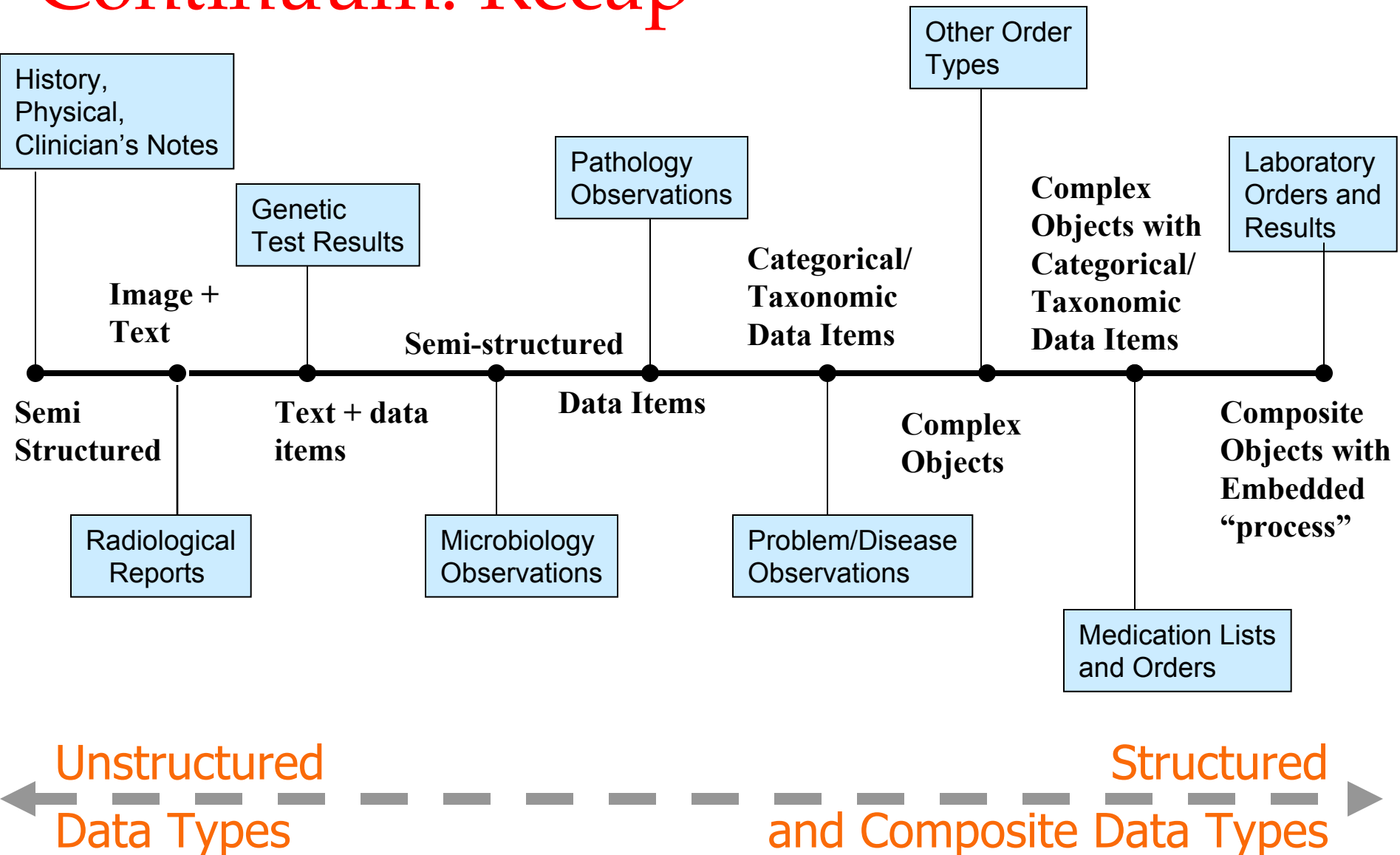
< Ok > <Cancel> <mMulti-day labs> <D/C-Change> <Microbiology>

Type the letter of test you wish to order. Type the letter again to deselect.
Press Alt + red letter of the desired operation. Alt-0: accept. Esc:cancel

Structured and Composite Data containing process information

- Lab panel contains individual data items and panels.
 - "Process" related information .. Perform a particular lab before another
- Opportunity: Link to standardized ontologies and information models
Link to descriptions of guidelines and processes

The Healthcare Data Semantics Continuum: Recap



Putting it all together!

Guided Data Interpretation

Guided Observation Capture

Guided Ordering

Patient Demographics

Smart Form, Next Gen

SmartView | Graphs | Patient View

No Filter | DM | CAD | CHF

Problems Add New >

DM-Related

- Diabetes Mellitus Type 2 07/05/2005
- CAD 12/07/2004
- PVD 06/30/2001

Other

- CHF 01/14/2005

Medications Renew Add New >

Antihyperglycemics

- Glipizide 5 MG (5MG TABLET take 1) PO BID 07/05/2005

Aspirin/Antiplatelet

- Aspirin (ACETYSALICYLIC ACID) 650 MG (650MG TABLET take 1) PO QD PRN 03/03/2005

ACE/ARB

- Lisinopril 10 MG (10MG TABLET take 1) PO QD 06/07/2005

Lipid-lowering

- Zocor (SIMVASTATIN) 20 MG (20MG TABLET take 1) PO QHS 06/07/2005

Beta-blockers

- [Reason no Rx...](#)

Other

- Tylenol (ACETAMINOPHEN) 650 MG (325MG CAPSULES take 2) PO Q4H PRN pain 07/04/2005
- Digoxin 0.25MG Tablet QD. Dispense 30 Tablets, Refills: 3 01/14/2005

Allergies Add New >

- Diphospholipitron Oleobipufone - Itching 07/05/2005
- ATENOLOL - Bronchospasm or Wheezing 12/07/2004
- METFORMIN - Rash 01/14/2005

Vitals Add New >

| | 1/23/2006 | 9/23/2004 | 6/22/2004 | 3/21/ |
|--------------|-----------|-----------|-----------|-------|
| BP (<130/85) | | 120/80 | 130/80 | 125/7 |
| Weight | 185 | 184 | 178 | 180 |
| Height | | 5'10" | 5'10" | 5'10" |

Note

Carry Forward... | Template...

Chief Complaint
Chest Pain

History of Present Illness

[Angina Template]

The patient is a 63 year-old with a history of coronary artery disease, poorly controlled diabetes and high blood pressure who was diagnosed with a myocardial infarction, poor... Stable angina since last visit? yes no

Angina occurs with: Choose...

Angina Class: Choose...

Symptoms

- Dyspnea on Exertion
- Orthopnea
- Paroxysmal Nocturnal Dyspnea
- Pedal Edema
- Palpitations
- Lightheadedness

Comments

Past Medical History

- Diabetes, type II, diagnosed 15 years ago on insulin
- Hypertension, diagnosed 15 years ago on lisinopril
- Chronic Kidney Disease, diagnosed 15 years ago on lisinopril
- Coronary Artery Disease, diagnosed 15 years ago, had a catheterization done then showed diffuse disease and managed with CABG
- Hypercholesterolemia
- Hypothyroidism
- Nephrolithiasis
- Obesity

Past Surgical History

- C-section, age 21

Medications

- Aspirin, 325 mg daily for treatment of CAD
- Metoprolol 25 mg twice daily for HTN and CAD
- Benazepril 40 mg daily for HTN and proteinuria
- Amlodipine 10 mg daily for HTN
- Atorvastatin 20 mg daily for high cholesterol
- Famotidine 20 mg twice daily for heartburn
- Nitroglycerine 0.3mg tabs to be used prn chest pain
- Synthroid 175 mcg a day for hypothyroidism
- Tylenol 650 mg 2-3 times a day for back pain
- Insulin regular: 35 units with meals and NPH 35 units at bedtime for diabetes

Allergies

Sulfa drugs cause a rash and penicillins cause hives.

Family History

The patient's mother died from post-partum bleeding after delivery of the patient's youngest sibling. Her father was a smoker and developed lung cancer having died at age 72. 2 brothers and 3 sisters have Diabetes and hypertension. Her one older brother has had multiple MIs and is on dialysis for renal failure. She has 3 children all alive and well.

Social History

Orders/Assessment/Plan

Highlights

- HgA1c is too high (9.0 on 01/15/2005)
- Statement of Assessment for each Treatment Goal

Glycemia

HbA1c is too high (9.0 on 01/15/2005, goal: HbA1c < 7)

- Start Metformin 500MG (500MG Tab take 1) PO BID
- Start Metformin PO...
- Start Metformin...
- Start an Insulin...
- Adjust Metformin...
- Discontinue Digoxin
- Order a test
- Refer for teaching
- Refer to a specialist
- Schedule Follow-up
- Print Patient Education Materials
- Add New Medication
- Add New Order

Comment:

Treatment Goal Name

Statement of Assessment (result, date)

- Order 1
- Order 2
- Add New Medication
- Add New Order

Comment:

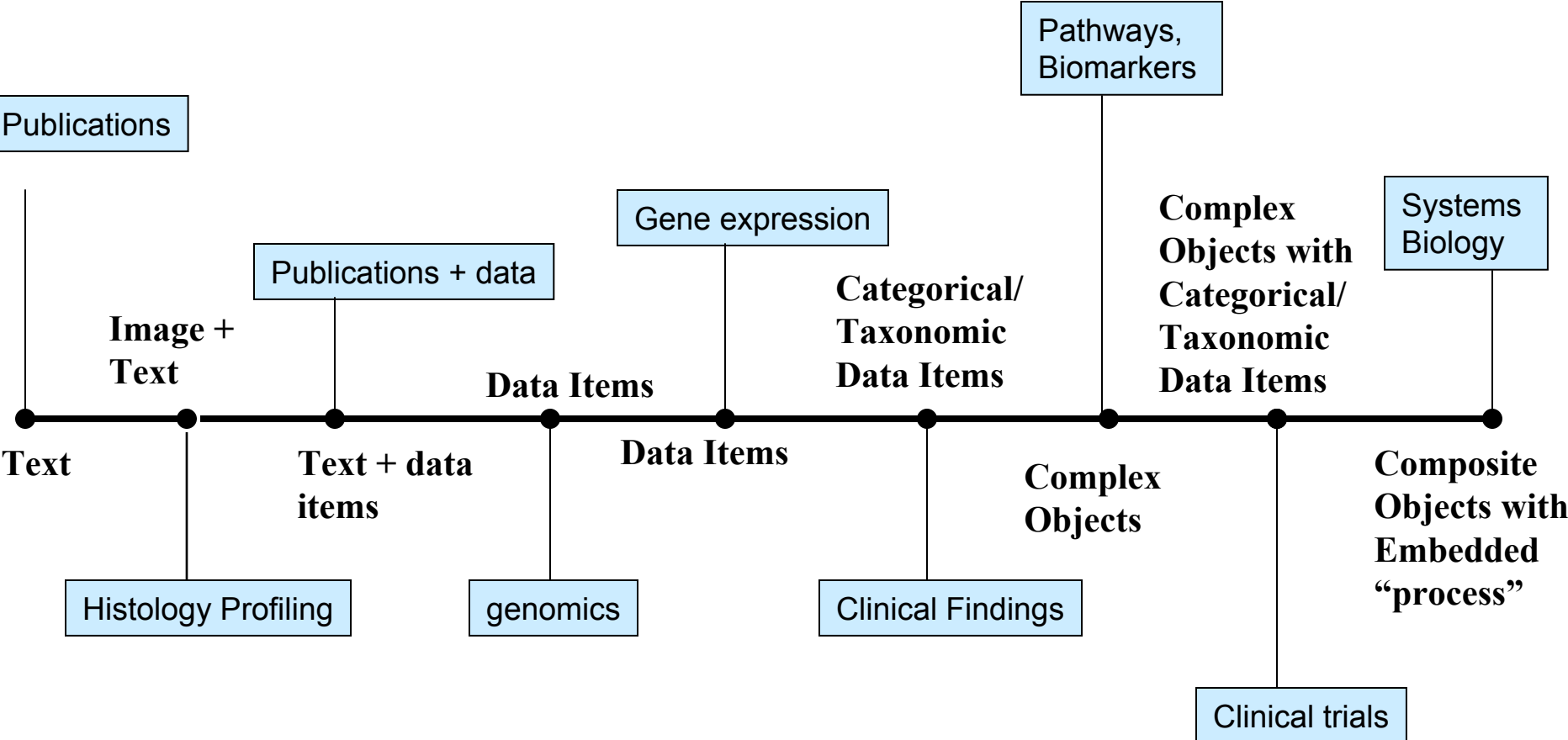
Treatment Goal Name

Statement of Assessment (result, date)

Opportunities for Semantics in HealthCare

- Enhanced interoperability via:
 - Semantic Tagging
 - Grounding of concepts in Standardized Vocabularies
 - Complex Definitions
- Semantics-based Observation Capture
- Semantics-based Clinical Decision Support
 - Guided Data Interpretation
 - Guided Ordering
- Semantics-based Knowledge Management

Data Semantics in the Life Sciences



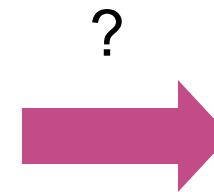
Unstructured
Data Types

Structured
and Complex Data Types

Pharma Data: Losing Connectedness in Tables

*Fast Uptake and ease of use,
but loose binding to entities and terms*

| | D | E | F | G | H |
|----|----------------|--------------------|--------------|------------|--------------|
| 3 | | | | | |
| 4 | Sample_Tissue_ | Sample_General_Pat | CASP2 | TNFRS | BAD |
| 5 | | | | | |
| 6 | BREAST | MALIGNANT | 0.272629684 | 0.13815558 | 0.172021004 |
| 7 | BREAST | NORMAL | 0.127580378 | 0.04861223 | 0.049566081 |
| 8 | COLON | MALIGNANT | -0.044756676 | -0.0857321 | -0.11424752 |
| 9 | COLON | NORMAL | -0.152073618 | -0.2025134 | 0.054543655 |
| 10 | KIDNEY | MALIGNANT | -0.067732844 | 0.130794 | 0.096568566 |
| 11 | KIDNEY | NORMAL | -0.063332286 | -0.1209417 | -0.024562741 |
| 12 | LIVER | MALIGNANT | -0.31962625 | -0.3550275 | -0.41515775 |



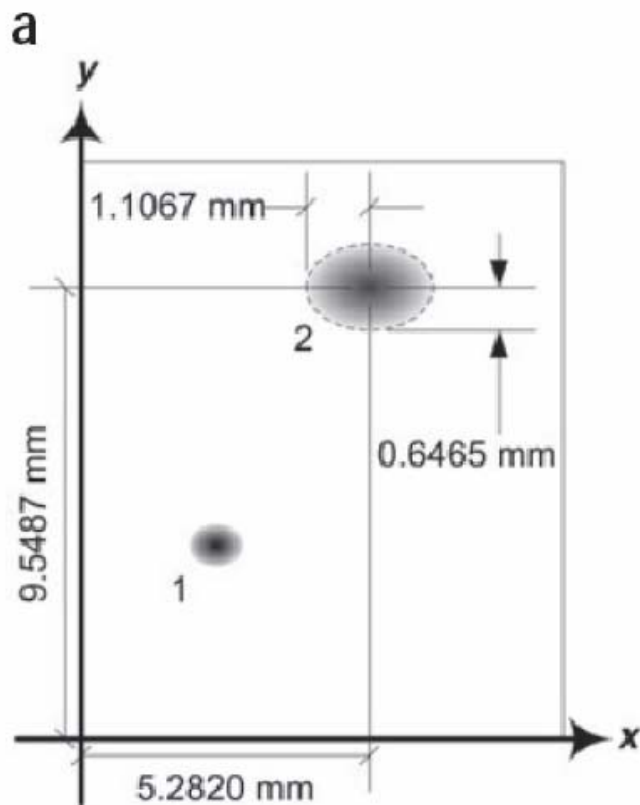
Genes

Tissues

- Querying Databases is not sufficient
- Data needs to include the Context of Local Scientists
- Concepts and Vocabulary need to be associated

Information ↔ Knowledge

Pharma Data



b

```
<spot>
  <spot_num>2</spot_num>
  <coord_x>5.2820</coord_x>
  <coord_y>9.5487</coord_y>
  <dia_x>1.1067</dia_x>
  <dia_y>0.6465</dia_y>
</spot>
```

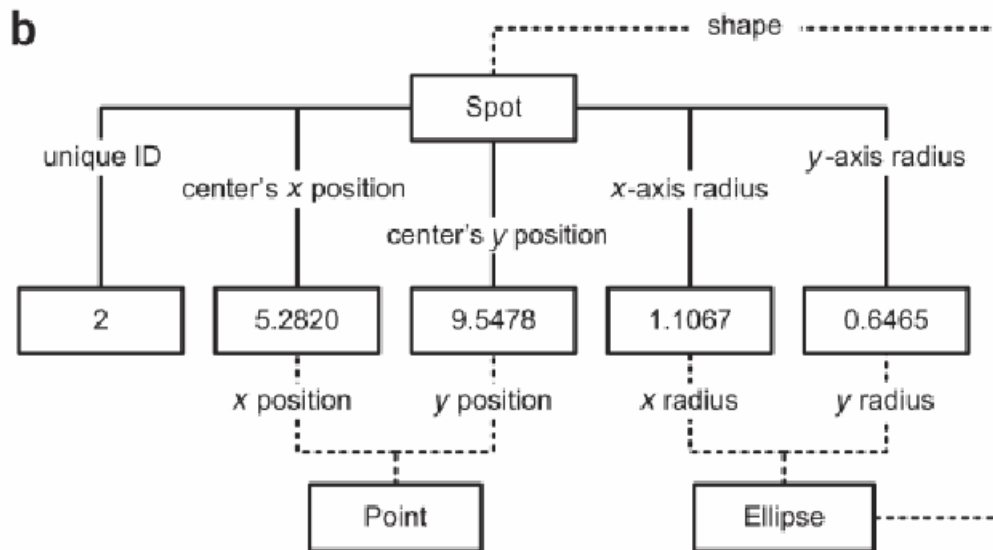
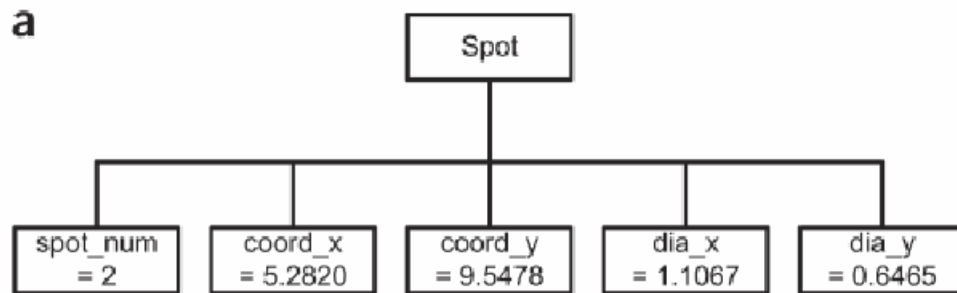
c

```
<spot>
  <spot_label>2</spot_label>
  <spot_location>
    <spot_position x="5.2820"
      y="9.5487"/>
    <spot_area width="1.1067"
      height="0.6465"
      type="ellipse"/>
  </spot_location>
</spot>
```

Annotated Gel
Markup Language
(AGML)

Human Proteome
Markup Language
(HUP-ML)

Pharma Data: Lack of Semantics

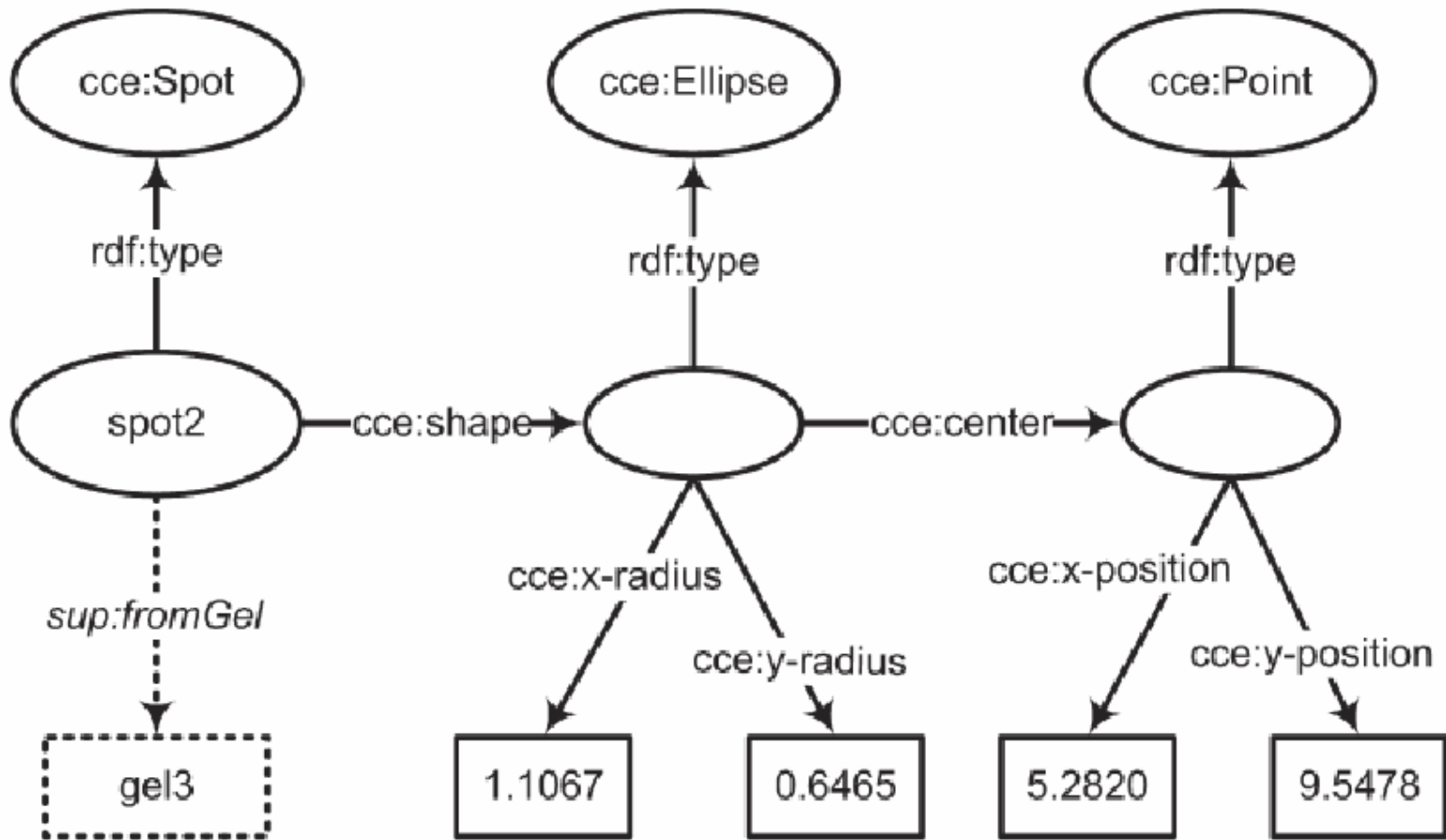


Missing description of semantic *relationships* between nested holders – required to invoke appropriate algorithms (shown as dotted lines in diagram b).

This is essentially a communication problem between the information and the *application* of information.

(Wang, Gorlitsky & Almedia, Nature, Sep. 2005)

Pharma Data: Opportunities for Semantics



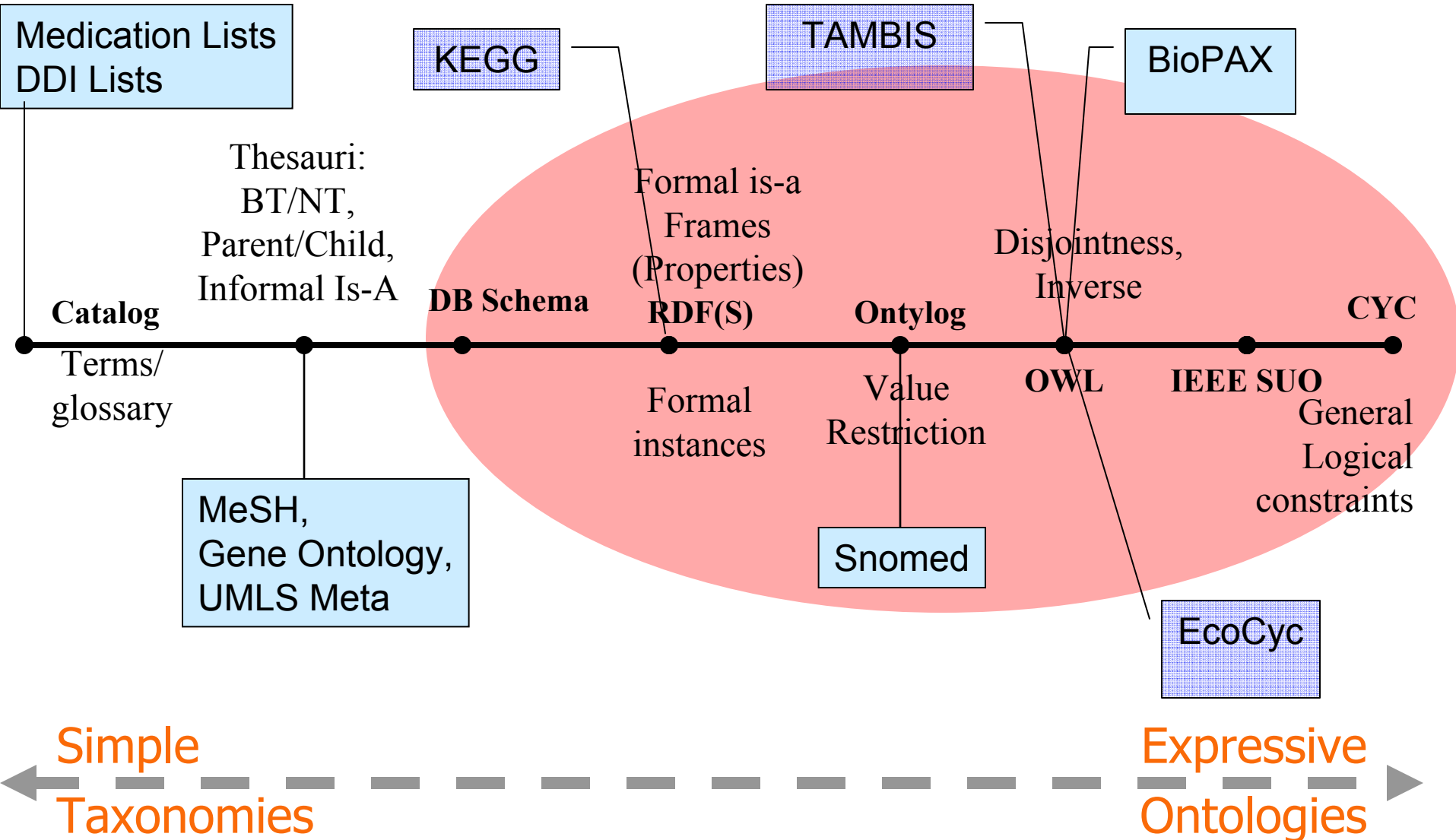
As a graph, the RDF model is oblivious to both syntax and semantics, but can be serialized into XML or N3 (NTriples).

(Wang, Gorlitsky & Almedia, Nature, Sep. 2005)

Predictiveness

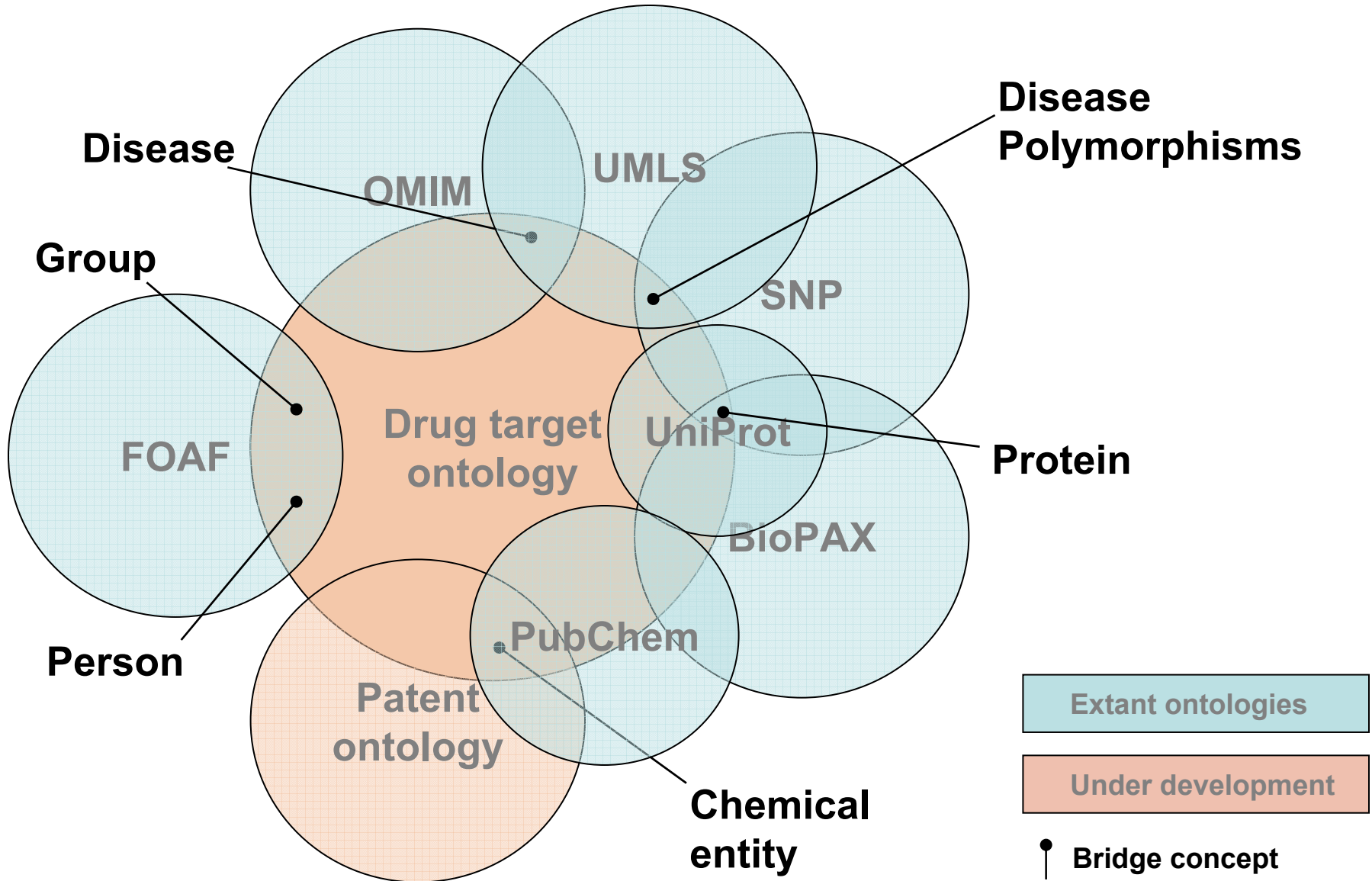
- Knowledge of Target Mechanisms
- Knowledge of Toxicity
- Knowledge of Patient-Drug Profiles

The Knowledge Semantics Continuum

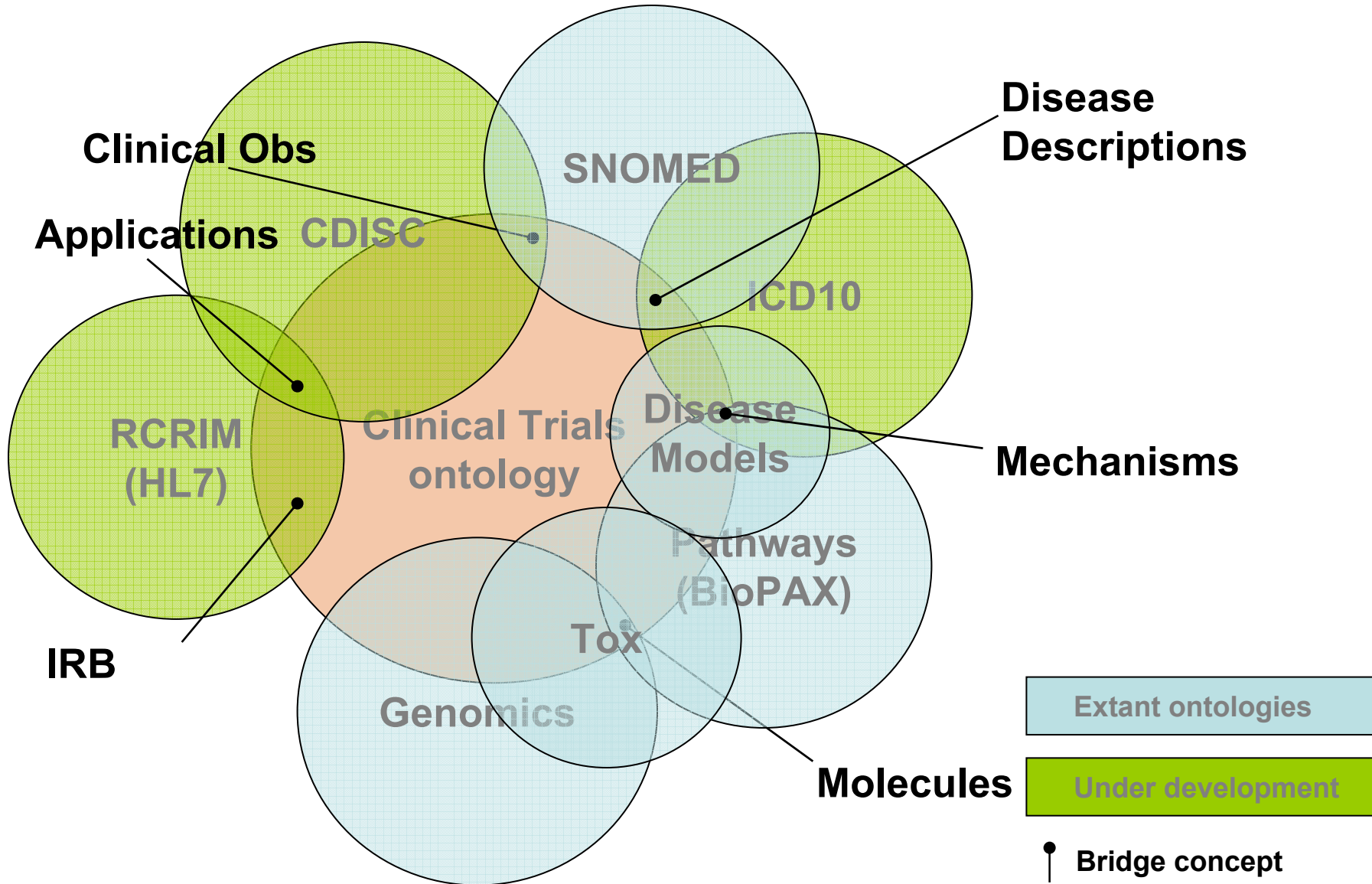


Ontology Dimensions based on McGuinness and Finin

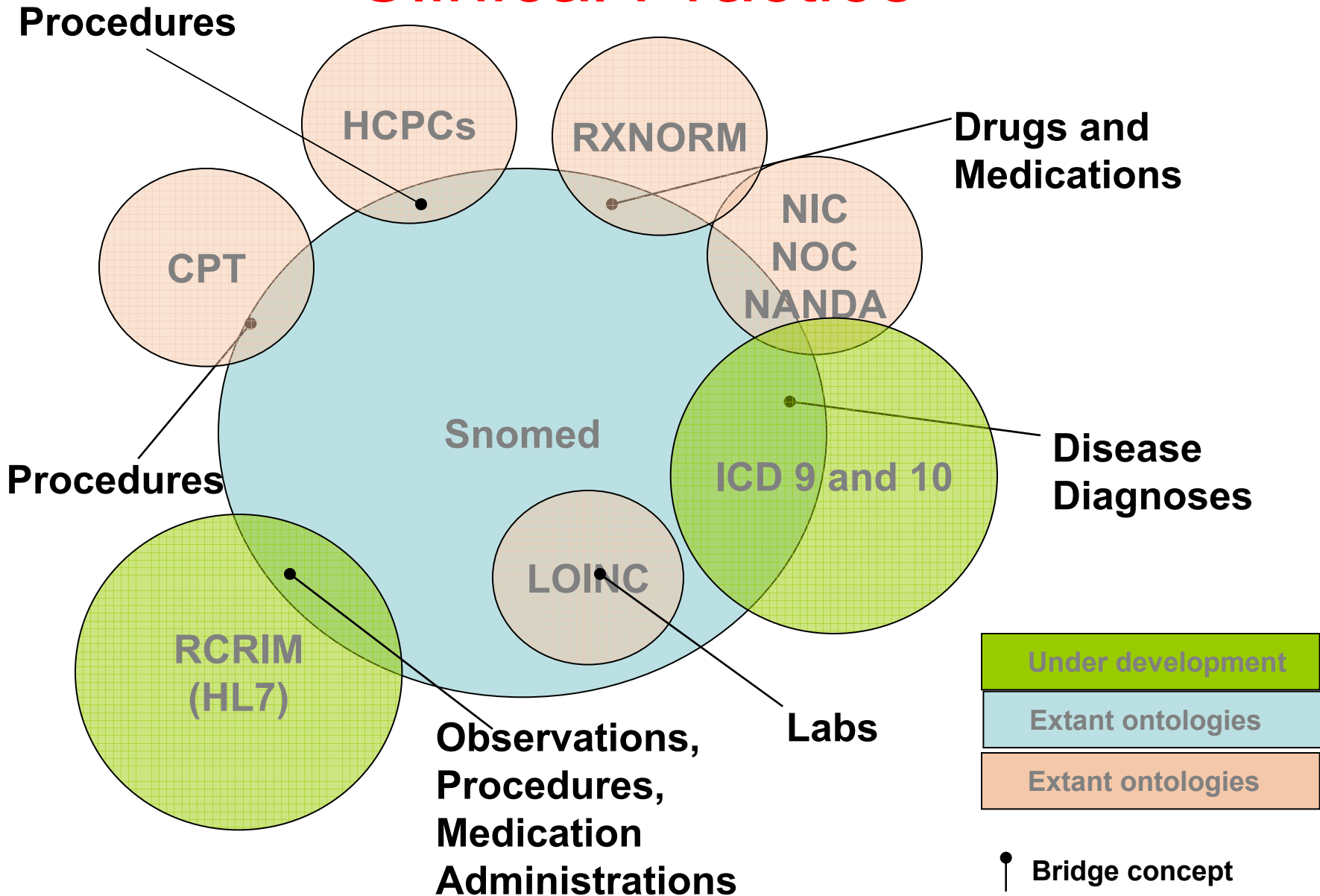
Knowledge Brokering: Life Sciences



Knowledge Brokering: Clinical Research



Knowledge Brokering: Clinical Practice



Opportunities for Semantics: Health Care

Data: 3000 Ambulatory Notes for Ortho Practice

NLP techniques + mining

Knowledge: Back Pain associated with:

- X History of Present Illness
- X Review of Systems
- X Physical Exam findings
- X Assessment possibilities
- X Diagnostic Test Possibilities (e.g., MRI)
- X Initial Therapy Possibilities

Elicited or Inferred
Clinician Preference

Rule/Policy: If Pain is > 6 weeks
And focal neuro findings,
May/Should Perform MRI

MRI is ordered only if clinical and admin/policy conditions are met!



Opportunities for Semantics: Biomedical/Clinical Research

Hypothesis Exploration based on Semantic Reasoning

Fact: CML is caused by tyrosine kinase secreted by fused ABL-BCR gene

Fact: Gleevec inhibits tyrosine kinase

Inference: Gleevec may reverse CML

Learning Clinical Guidelines based on Semantic Reasoning + Data Mining

Fact: Confusion correlates with increased risk of Falls in age>65

Fact: Sedatives/Hypnotics increase risk of confusion even in appropriate doses for elderly

Rule/Knowledge: If patient>65 on sedatives/hypnotics, institute falls precautions protocol

Opportunities for Semantics: Clinical Practice

- Clear definitions of Decision Support Categories
- IF patient has Diabetes and Renal Disease
AND no contraindication to ACEi or ARB
THEN prescribe ACEi or ARB
- Define “Contraindication to ACEi or ARB”
 - Allergy to ACEi or ARB
 - Cough symptom on adverse reaction list
 - Hyperkalemia on problem list or high K test result
 - Pregnancy (Needs further definition)
 - Could lead to complex and nested definitions
 - Post-coordinated approaches
 - Patient refuses or failed the drug
 - New Molecular diagnostic test

Outline

- The Bench ↔ Bedside Vision
- Role of Data, Knowledge and Semantics in the HCLS Ecosystem
- Functional Requirements
 - Metadata-based Semantic Annotation
 - *Recombinant Data*: Information Aggregation and Integration
 - Ontology-driven Decision Support
 - Knowledge Update and Maintenance
- Conclusions


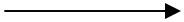
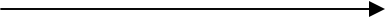
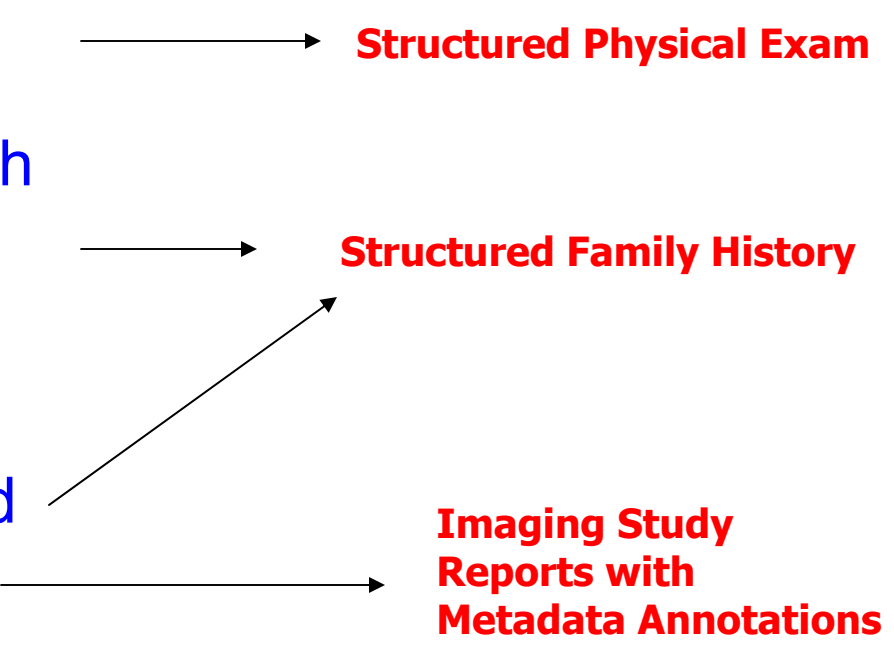
Functional Requirements



- Use Cases
- Metadata-based Semantic Annotation
- *Recombinant Data*: Information Integration and Aggregation
 - Discovery Decision Support
- Ontology-driven Decision Support
- Knowledge Update and Maintenance

Use Case: Personalized Medicine

Dr. Genomus Meets Basketball Player who fainted at Practice

- Clinical exam reveals abnormal heart sounds  **Structured Physical Exam**
 - Family History: Father with sudden death at 40,  **Structured Family History**
 - 2 younger brothers apparently normal
 - Ultrasound ordered based on clinical exam reveals cardiomyopathy  **Imaging Study Reports with Metadata Annotations**
- 

Use Case: Actionable Decision Support

LMR Smart Form - Microsoft Internet Explorer provided by Partners HealthCare System

Address: http://is.partners.org/prototype/dtaylor/marchdmsf/smartform3.htm

LMR Smart Form

Patient: GUBERNATH, JANETTA DOB: 08/25/1947 Age: 58
MRN: 0000001 (MGH) Sex: F Tel: (H) 617-555-1212

Conditions to include: DM CAD ARI

Summary | Graphs | Note Preview | Patient-friendly

CC: Carry forward all note content from: Most recent note 3/18/05

HPI: Lab Results

| Order | 12/20/04 | 11/10/04 | 9/14/04 | 6/4/03 |
|-------------------------------|----------|----------|---------|--------|
| Glucose (mg/dL) | 175 | 185 | - | 145 |
| A1C (4.4%-6.4%) | 8.1 | 8.2 | - | 8.3 |
| Total Chol (<200 mg/dL) | 210 | - | 240 | 190 |
| HDL (>45 mg/dL) | 59 | - | 60 | 53 |
| LDL (<100 mg/dL) | 112 | - | 115 | 118 |
| Trig (<200 mg/dL) | 125 | - | 125 | 125 |
| SGOT (U/L) | 10 | 7 | - | - |
| BUN (mg/dL) | 12 | 11 | - | - |
| Cr (mg/dL) | 1.2 | 1.1 | - | - |
| ualb / Cr ratio (mg Ald/g Cr) | - | - | - | - |

Orders / Assessment / Plan

Needs Attention

- A1C high (8.1 on 12/20/04)
- LDL high (112 on 12/20/04)
- Total Chol high (210 on 12/20/04)
- ualb / Cr overdue (Last 11/9/03)
- Foot exam overdue (Last 4/14/03)
- Eye exam overdue (Last 10/6/03)
- Need current BP
- BMI high (34.0 today)
- Home glucose monitoring not documented

A1C

A1C high (8.1 on 12/20/04)

- ◆ Adjust glycemic therapy
- Refer to CDE
- Have patient report AM FBG after first 3-5 days
- Patient ed: What is insulin?
- Patient ed: Giving an insulin injection

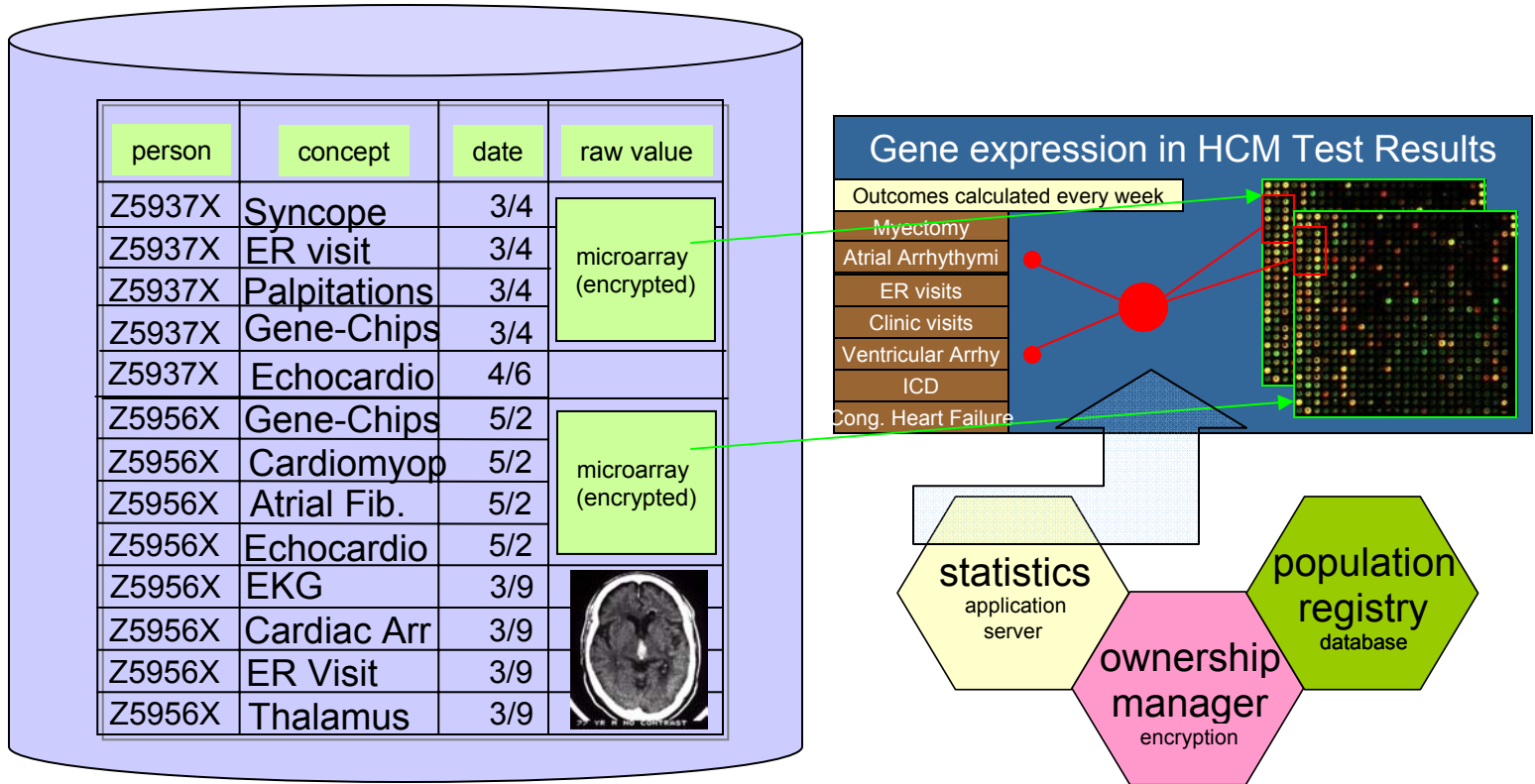
Cancel Complete

Echo triggers guidance to screen for possible mutations:

- MYH7, MYBPC3, TNN2, TNNI3, TPM1, ACTC, MYL2, MYL3

Use Case: Connecting Genotypic and Phenotypic Data

Connecting Dx, Rx, Outcomes and Prognosis Data to Genotypic Data for Cardiomyopathy



Use Case: Drug Discovery

ApoA1 ...

- ... is produced by **the Liver**
- ... is expressed less in **Atherosclerotic Liver**
- ... is correlated with **DKK1**
- ... is cited regarding **Tangier's disease**
- ... has Tx Reg elements like **HNFR1**

Functional Requirements

- Use Cases
- Metadata-based Semantic Annotation
- *Recombinant Data*: Information Integration and Aggregation
 - Discovery Decision Support
- Ontology-driven Decision Support
- Knowledge Update and Maintenance

*The first step of any biomedical activity
(research, practice, knowledge gathering)
should be on the computer!*

Metadata-based Semantic Annotations: Connotea

Connotea: Bookmarks with search terms cox-2 and inhibitors - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Refresh Mail Print Address Book

Address <http://www.connotea.org/search?q=cox-2+inhibitors> Go Links

Google Connotea Search PageRank 7524 blocked Check AutoLink AutoFill Options Connotea

logged in as **vipul_kashyap** **Logout**

Connotea Search All cox-2 inhibitors Search My Library | Registration

Home | Latest News | About This Site | Site Guide | FAQ | Community Pages | Popular Links | Recent Activity | Contact Us

Tags used on these bookmarks:

- therapy
- Medical Sciences
- Pharmacology
- COX-2 inhibitors
- viral
- figure
- Pregnancy
- HIV
- network
- COX inhibitors
- Amino Acids - Protei...
- Sonva's Database Que...
- rheumatoid arthritis
- mammalian protein ki...
- US GOV
- oligohydramnios
- Gene Bank -Genomics ...
- Drug resistance
- active rheumatoid ar...
- immunology
- molecular research
- fetus
- Capsid Proteins
- Zoonoses-vector jump...
- Economics and or Tra...
- Ecological Sciences
- ulcers
- 2005
- Diagnosis and screen...
- review

Bookmarks with search terms cox-2 and inhibitors **EXPORT LIST** **RSS** **?**

Note: Your search term matches the global tag [COX-2 inhibitors](#).

Number of bookmarks per page: **10** | 25 | 50 | 100

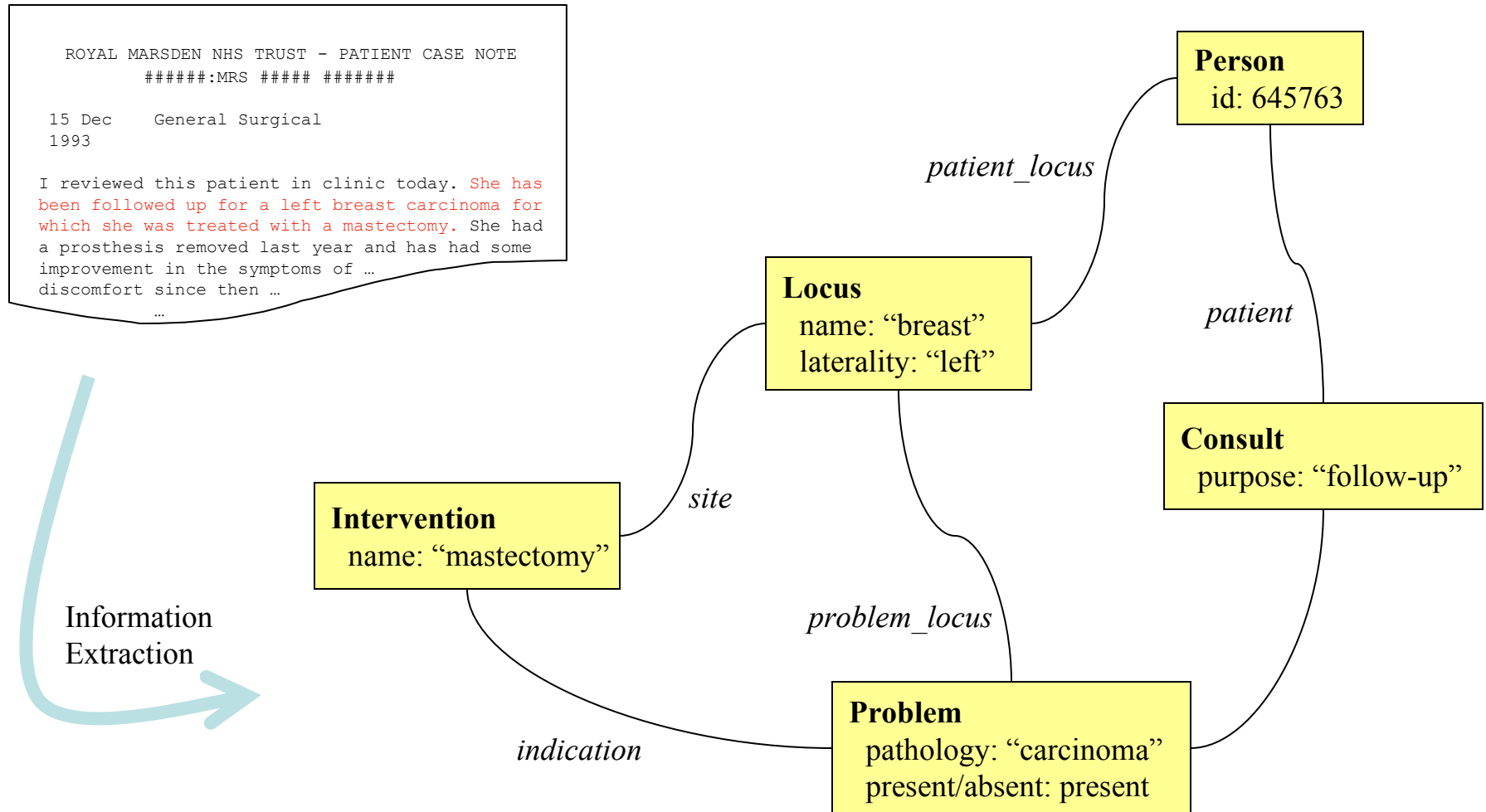
copy
[Replication-selective virotherapy for cancer: Biological principles, risk management and future directions](#) (info)
David Kirn, Robert Martuza, and James Zwiebel
Nat Med **7** (7), 781-7 (Jul 2001)
[doi:10.1038/89901](#)
Posted by **madhu** to [virotherapy](#) on [Mon Jan 30 2006](#) at 15:23 UTC

copy
[COX-2 inhibitors and metabolism of essential fatty acids](#). (info)
Undurti N Das
Med Sci Monit **11** (7), RA233-7 (Jul 2005)
[PMID: 15990700](#)
Posted by **ingerida** to [review 2005 metabolism PRINT-D F: Das U N COX-2 inhibitors pathway essential fatty acids network figure](#) on [Thu Sep 01 2005](#) at 09:25 UTC

copy
[Is the use of COX-2 inhibitors in gastroenterology cost-effective?](#) (info)
David Graham and Francis KL Chan
Nat Clin Pract Gastroenterol Hepatol **1** (2), 60-1 (Dec 2004)
[doi:10.1038/ncpgasthep0043](#)
Posted by **NatureClinicalPractice** to [ulcers therapy](#) on [Thu Dec 08 2005](#) at 18:08 UTC

copy
[Genome Biology | Full text | Attacking pathogens through their hosts](#) (info)
Genome Biology **7** (1), 201 (2006)
[doi:10.1186/gb-2006-7-1-201](#)
.. "This is partly because the disease burden of any one pathogen is unlikely to reach sufficient levels for pharmaceutical companies to justify the enormous cost of developing a new drug, which (although hotly debated) is estimated to be between \$0.5 billion and \$1.7 billion [4-6]. Although legislation over the past 20 years in the USA, especially the 'Orphan Drug Act', is designed to reduce such barriers, the difficulty of developing new antimicrobial drugs remains, and it is compounded by the fact that many infectious diseases requiring treatment occur in developing countries, which cannot cope with the costs of new drugs." .. "The concept of attacking the microbe by altering or augmenting a host-cell function or process is not new. The use of interferon α (IFN α) in combination with ribavirin in the treatment of hepatitis C virus infection is successful in 50% of infected

Metadata-based Semantic Annotations: Clinical E-Science Framework



Metadata-based Semantic Annotations (HubMed)

HubMed: cox-2 inhibitors - Microsoft Internet Explorer

Address <http://www.hubmed.org/search.cgi?q=cox-2+inhibitors&sort=relevance>

Search: **cox-2 inhibitors** Results 1-20 of 3293

sort by: date/relevance 0 Next 20

All

1 **Cox-2 inhibitors.**
Brown E
Physician Exec. 1999 Jan-Feb ; 25(1): 74-6

[Abstract](#) [FullText](#) · [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) · [Tag](#)

tags (space-separated)

annotation

2 **COX-2 inhibitors.**
Brooks PM, Day RO
Med J Aust. 2000 Oct 16; 173(8): 433-6

[Abstract](#) [FullText](#) · [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) · [Tag](#)

3 **COX-2 inhibitors.**
Becker RC
Tex Heart Inst J. 2005; 32(3): 380-3

[Abstract](#) [FullText](#) · [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) · [Tag](#) · [Review](#)

4 **COX-2 inhibitors.**
Hawkey CJ
Lancet. 1999 Jan 23; 353(9149): 307-14

[Abstract](#) [FullText](#) · [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) · [Tag](#)

5 **4,5-Diaryloxazole inhibitors of cyclooxygenase-2 (COX-2).**
Talley JJ, Bertenshaw SR, Brown DL, Carter JS, Graneto MJ, Koboldt CM, Masferrer JL, Norman BH, Rogier DJ, Zwwifel BS, Seibert K
Med Res Rev. 1999 May ; 19(3): 199-208

<http://www.hubmed.org/tags/edit/10387277>

start Kingston (F:) http://innovation.sw... https://wg.eicc.co.uk... HubMed: cox-2 inhib... Final Tutorial Present... Internet 9:47 AM

Metadata-based Semantic Annotations: PubMed

HubMed Tag Storage: demo - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://www.hubmed.org/tags/users/demo>

HUBMED TAGS (BETA!)

Rat CD8+ FOXP3+ T suppressor cells mediate tolerance to allogeneic heart transplants, inducing PIR-B in APC and rendering the graft invulnerable to rejection. 2004
demo: pir-b

Tolerization of dendritic cells by T(S) cells: the crucial role of inhibitory receptors ILT3 and ILT4. 2002
demo: ilt4

Generation and function of antigen-specific suppressor and regulatory T cells.
demo: ilt4 specificity treg

Manipulation of immune regulation in systemic lupus erythematosus. 2005
demo: test

Induction of allopeptide-specific human CD4+CD25+ regulatory T cells ex vivo. 2003
demo: test

Targeted CTLA-4 engagement induces CD4+CD25+CTLA-4high T regulatory cells with target (allo)antigen specificity. 2004
demo: test

The role of TCR specificity in naturally arising CD25+ CD4+ regulatory T cell biology. 2005
demo: test

CD4+ regulatory T cell responses induced by T cell vaccination in patients with multiple sclerosis. 2006
demo: test

Identification of a CD4+CD25+ T cell subset committed in vivo to suppress antigen-specific T cell responses without additional stimulation. 2004
demo: cd134 treg

home
search
login
register
tag cloud
recommendations
->hubmed

start Kingston (F:) <http://innovation.sw...> <https://wg.eicc.co.uk...> HubMed Tag Storage... Final Tutorial Present... Internet 9:50 AM

Metadata-based Semantic Annotations: Active Semantic EMR

AHC Acct No: 222222 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites

Address C:\jex\final.xml

30606 Phone: 706-208-9700 Fax: 706-208-0806

Athens Heart Center
Shyam "Sham" Prabhakar
333Dogs Drive, Apt 604, ATHENS, GA 30606
SSN: 222-22-2222 MR #: 222222 Sex: M DOB: 05/07/1970 Age: 33

Visit on 07/29/2005

Referred doctor from Practice Ontology

Office Visit Note - Complete History & Physical

Other Physicians: David Almand, M.D. E Emergency Medicine (770) 922-3023 Timothy Gibson, M.D. E Family Practice 706-227-2027 Alfredo Alarcon, E (404) 256-5212

Problem List:

- 3. Backache unspecified E
- 4. Hypercholesterolemia E
- 5. Chest Pain E
 - A. Hypertension E
 - B. Shortness of Breath.
- 6. Dilated cardiomyopathy gthghjg
- 7. Abnormal ECG
- 9. Acute Glomerulonephritis with other specified Pathological Lesion in Kidney E
- 10. Something
- 11. Chest Pain E

Chief Complaint: Follow up of abdominal aortic aneurysm, angina, aortic stenosis, aortic valve replacement, dental clearance, and atrial fibrillation status post abnormal stress test. Cardiac clearance for aneurysm removal. Follow up of recent hospitalization at BJC - Commerce for atrial fibrillation.

History of Present Illness: Mr. Prabhakar is a 35 year old patient of Dr David Almand, Dr Timothy Gibson, and Dr Alfredo Alarcon. He was admitted to Ty Cobb Memorial by Dr. Alfredo Alarcon for bradycardia. He was found to have complete heart block. He was treated with mitral valve replacement. And he did not respond well. The patient was then transferred to Emory. He was admitted to Emory by Dr. Timothy Gibson for angina. He was found to have atrial fibrillation and complete heart block. He was treated with cholecystectomy and he responded well. The patient was then transferred to St. Mary's Hospital. He is here today for follow up management of arrhythmia, atrial fibrillation, and ICD function. Since his last visit no new problems have developed. He is taking his medications as prescribed. There appear to be possibly some side effects related to the medications. Overall, he believes that his arrhythmia, atrial fibrillation, and ICD function is poorly controlled. He is here today for follow up management of cardiomegaly and coronary artery disease. Since his last visit no new problems have developed. He is taking his medications as prescribed. There appear to be possibly some side effects related to the medications. Overall, he believes that his cardiomegaly and coronary artery disease is stable. The chest pain is associated with itchy. He reports that his chest pain is aggravated by bending. The chest pain is relieved by belching. The patient reports this morning with her spell. Dizziness is associated with palpitations. He states that this symptom is aggravated by activity, bending, and high blood pressure. The dizziness is relieved by sitting down. He states that the palpitations are aggravated by position changes.

ICD9 codes from Diagnosis Procedure Ontology

Lexical annotation

Current Medications **Medications After Visit**

Done My Computer

Metadata-based Semantic Annotations: Active Semantic EMR

AHC Acct No: 222222 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address C:\lex\final.xml

| Current Medications | Medication |
|--|---|
| Intropin injection 40mg/ml, 1 inj qd I E | Intropin injection 40mg/ml, 1 inj qd I E |
| Tasmar tablets 200mg, 1 tab qd I E | Advil 100mg/5ml, 1 susp qd F E |
| | Tasmar tablets 200mg, 1 tab qd I E |
| | Tylenol extended release 325mg, 1 tab er qd A F E |

Pharmacy: Carson's Commerce Drug Company Phone: 706-754-4128 Phone: 706-335-3111

Allergies: AMPICILLIN, IVP DYE, PENICILLIN, TYLENOL

Past Medical History: No past trauma.

Family History: Mr. Prabhakar has a positive family history of coronary artery disease. The patient's **crap2** is deceased.

Psychosocial: Mr. Prabhakar resides in a **apartment** home. He is lives with her daughter to hkkd. He has excellent social support.

Life History:

Review of Systems:

General Patient reports **daily chills** associated with **chest pain**. These symptoms **have not** been worked up by his primary care physician. Patient reports recent unintentional weight **gain**. This problem **has not** been worked up by his primary care physician. Patient reports recent onset of **severe trouble falling asleep**. This problem **has** been working up by his primary care physician.

HEENT The patient reports **migraine** headache.

Genitourinary Patient denies dysuria. Patient complains of **recent onset** of hematuria. Patient complains of **chronic presence** of hesitancy. Patient denies dribbling. Patient complains of **recent onset** of burning with urination.

Hematologic Patient reports history of blood transfusion as a result of **anemia**. Patient reports he **did** experience a transfusion reaction.

Skin Patient reports **frequent of mild pruritis** associated with **weakness**. This problem **has been worked up by the patient's dermatologist**. Patient reports a **walnut-sized keratosis** that is located **over the entire body**.

Psychiatric Patient admits to a history of **panic attacks** that is currently **managed by common mental health**. His symptoms are felt to be **not under control**.

Vital Signs: Height: 170 Weight: 280 lbs BP: 120/80 Pulse: 80 Respirations: 20

Physical Examination:

General The patient **appears the stated age**.

Done My Computer

Formulation Recommendation
Using Insurance ontology

Drug Interaction using
Drug Ontology

Drug
Allergy

Metadata-based Semantic Annotations: Active Semantic EMR

AHC Acct No: 222222 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites

Address C:\lex\final.xml

Explore: Drug *Tasmar*

| Current Medications | Medications After Visit |
|--|---|
| Intropin injection 40mg/ml, 1 inj qd I E Tasmar tablets 200mg, 1 tab qd I E | Intropin injection 40mg/ml, 1 inj qd I E Advil 100mg/5ml, 1 susp qd F E Tasmar tablets 200mg, 1 tab qd I E Tylenol extended release 650mg, 1 tab er qd A F E |

Pharmacy: Carson's Commerce Drug Company
Phone: 706-754-4128
Phone: 706-335-3111

Allergies: AMPICILLIN, IVP DYE, PENICILLIN, TYLENOL

Past Medical History: No past trauma.

Family History: Mr. Prabhakar has a positive family history of coronary artery disease. The patient's **crap2** is deceased.

Psychosocial: Mr. Prabhakar resides in a **apartment** home. He lives with her daughter to hkk. He has excellent social support.

Life History:

Review of Systems:

General Patient reports **daily chills** associated with **chest pain**. These symptoms **have not** been worked up by his primary care physician. Patient reports recent unintentional weight **gain**. This problem **has not** been worked up by his primary care physician. Patient reports recent onset of **severe trouble falling asleep**. This problem **has** been working up by his primary care physician.

HEENT The patient reports **migraine** headache.

Genitourinary Patient denies dysuria. Patient complains of **recent onset** of hematuria. Patient complains of **chronic presence** of hesitancy. Patient denies dribbling. Patient complains of **recent onset** of burning with urination.

Hematologic Patient reports history of blood transfusion as a result of **anemia**. Patient reports he **did** experience a transfusion reaction.

Skin Patient reports **frequent** of mild pruritis associated with **weakness**. This problem **has been worked up** by the patient's dermatologist. Patient reports a **walnut-sized keratosis** that is located **over the entire body**.

Psychiatric Patient admits to a history of **panic attacks** that is currently **managed by common mental health**. His symptoms are felt to be **not under control**.

Vital Signs: Height: 170 Weight: 280 lbs BP: 120/80 Pulse: 80 Respirations: 20

Physical Examination:

General The patient **appears the stated age**.

Done My Computer

Metadata-based Semantic Annotations: Active Semantic EMR

Query View - DrugOnt_test_alt

◆ antiparkinsonian_agents
◆ neurological_agents
◆ interaction_with_cpnum_group_206_1
◆ cpnum_group_206
◆ comf_inhibitors
◆ cpnum_group_2118
◆ Tolcapone
◆ Tasmar
◆ formulary_1498

classification
classification
interaction
interaction
belongs to group
belongs to group
brand / generic

From: Tasmar
To: cpnum_group_2118
Type: has_cpnum_group

Semantic browsing and querying-- perform decision support (how many patients are using this class of drug, ...)

Nodes of Interest

Use Selected (none selected)
 Search By Name

Name:

case sensitive
 regular expression

Restrict Node Types

Defined Class
Enumeration
Individual
Logical Operation
Primitive Class
RDFS Class
Restriction

Arcs of Interest

Restrict Arc Types

equivalent to
has instance
has subclass
has_brandname_equivalent
has_brandname_equivalent (Domain>Range)
has_brandname_equivalent (Necessary and Sufficient)
has_brandname_prescription_drug
has_brandname_prescription_drug (Domain>Range)
has_brandname_prescription_drug (Necessary and Sufficient)
has_brandname_prescription_drug (Necessary)
has_cpnum_group
has_cpnum_group (Domain>Range)

Neighbours

Show Neighbours

Arc Direction

Outgoing and Incoming Arcs
 Outgoing Arcs
 Incoming Arcs

Restrict Levels

Levels:

Results

Query

Functional Requirements

- Use Cases
- Metadata-based Semantic Annotation
- *Recombinant Data*: Information Integration and Aggregation
 - Discovery Decision Support
- Ontology-driven Decision Support
- Knowledge Update and Maintenance

Information Integration: Ontology

OWL ontologies that blend knowledge from the Clinical and Genomic Domains

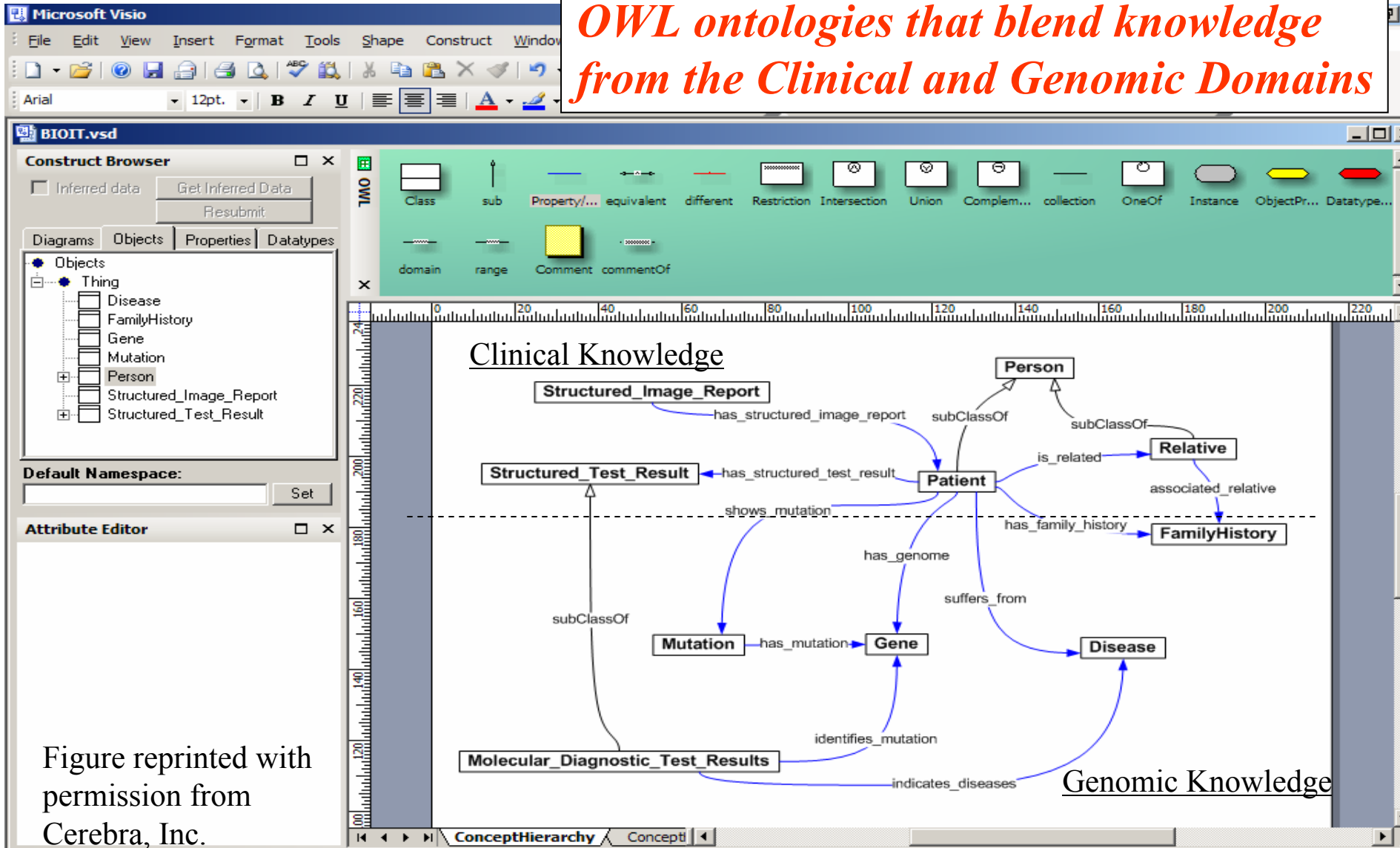
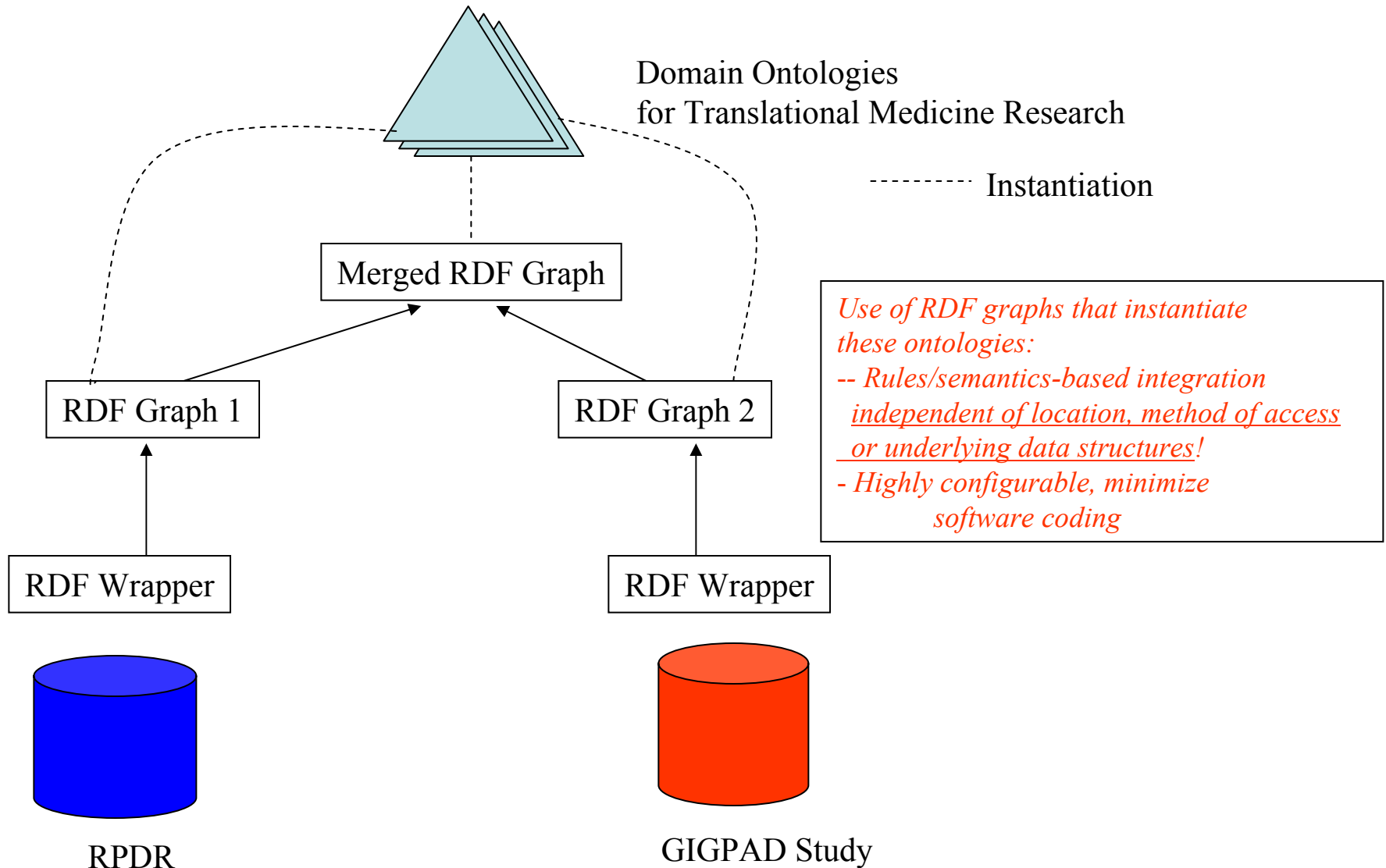
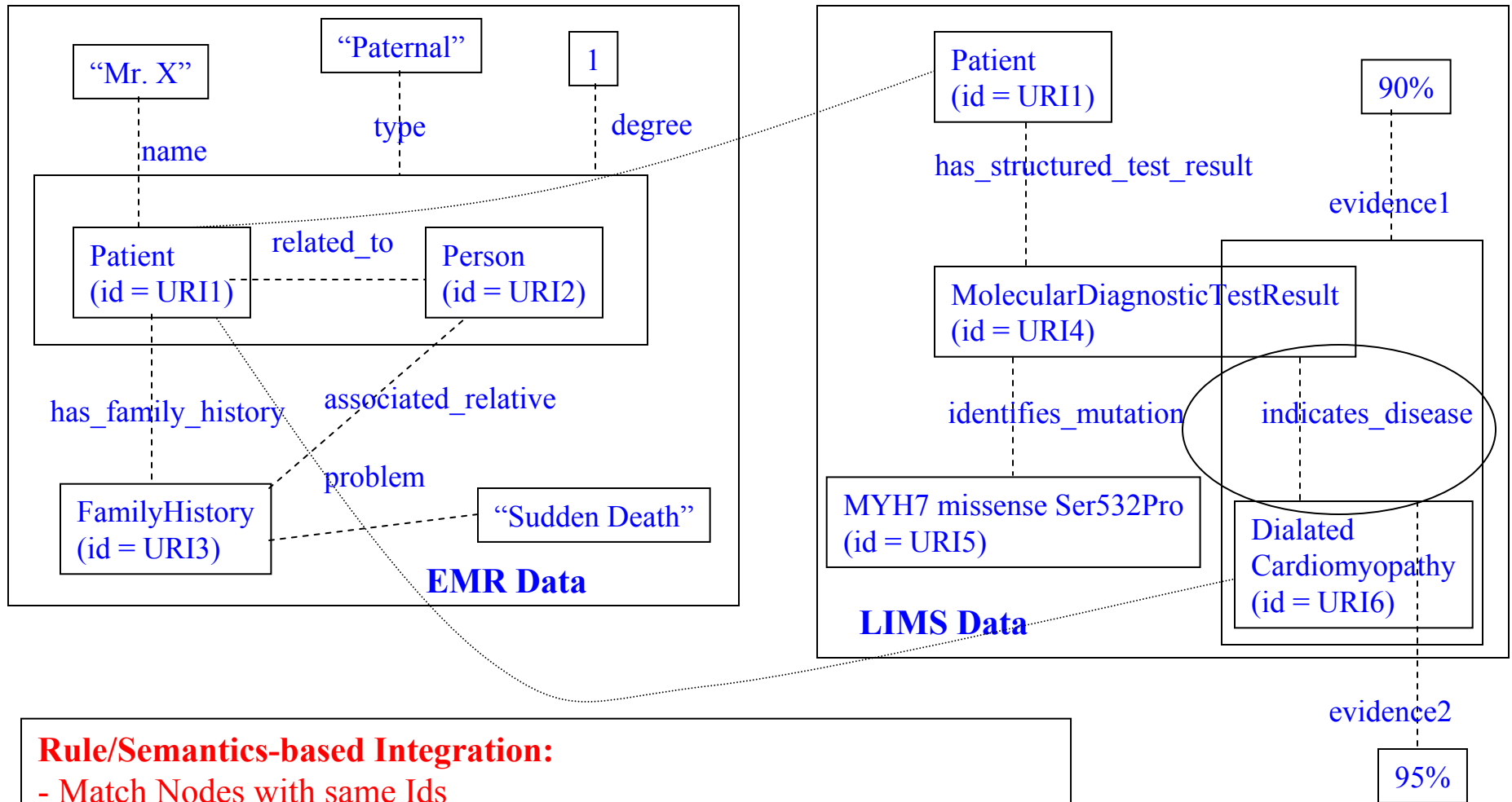


Figure reprinted with permission from Cerebra, Inc.

Information Integration: Architecture



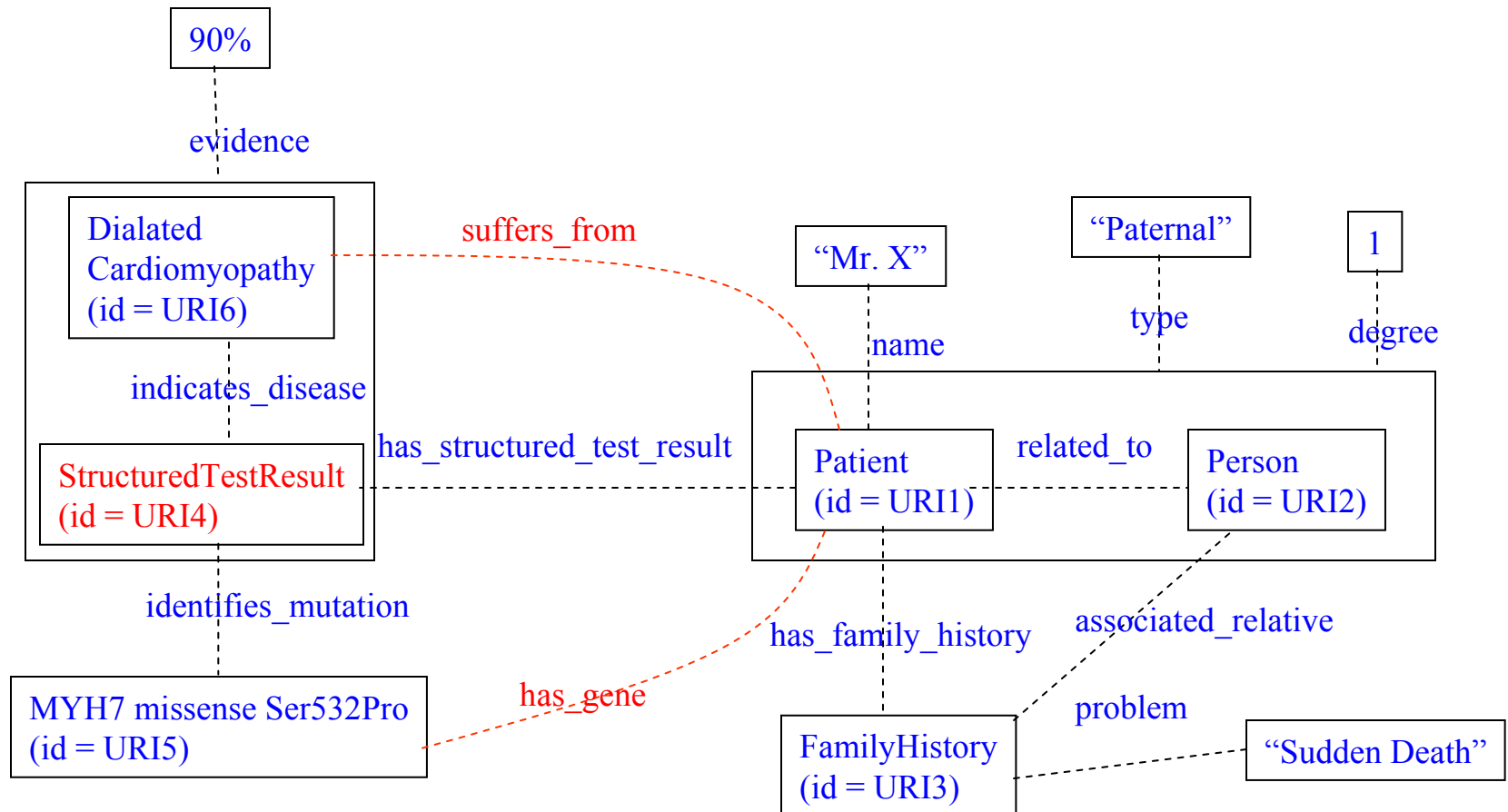
Bridging Clinical and Genomic Information



Rule/Semantics-based Integration:

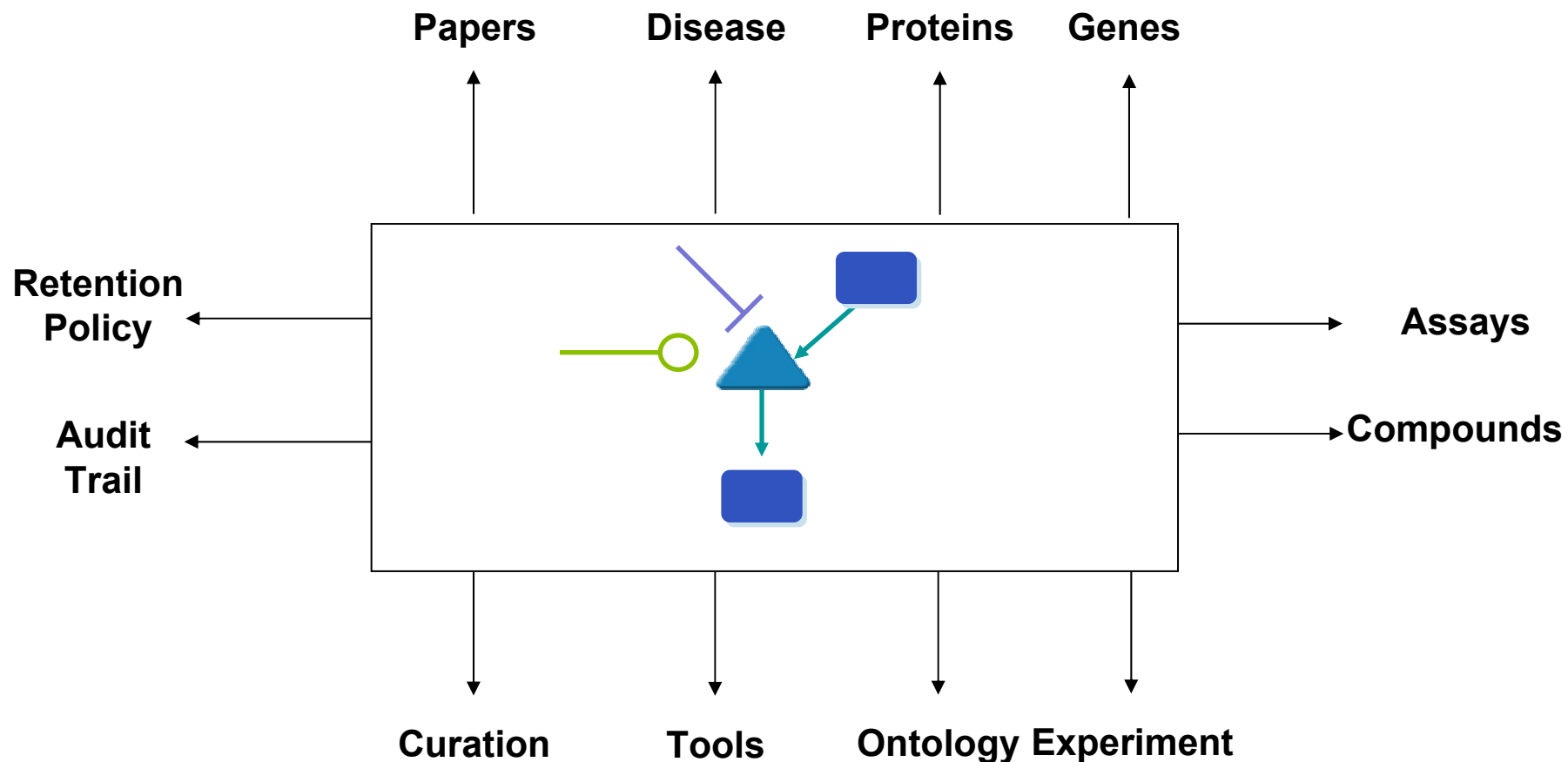
- Match Nodes with same Ids
- Create new links: IF a patient's structured test result indicates a disease THEN add a "suffers from link" to that disease

Bridging Clinical and Genomic Information

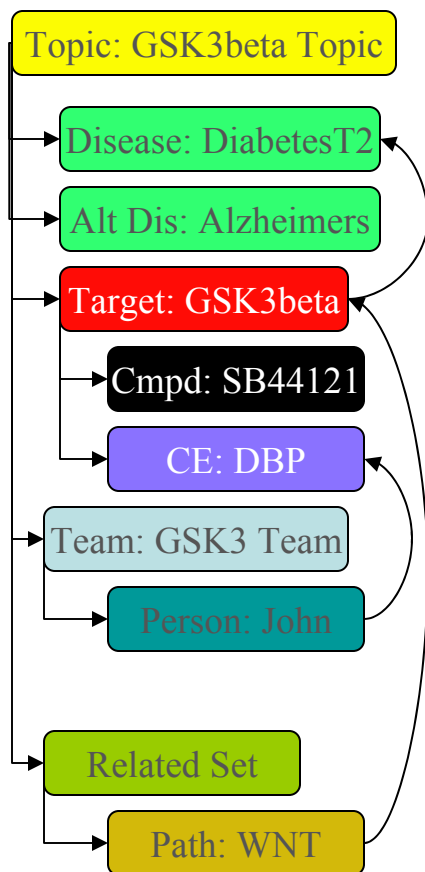


RDF Graphs provide a semantics-rich substrate for decision support. Can be exploited by SWRL Rules

Information Integration: Biology Requirements



Information Integration: Drug Discovery



Project Haystack Run Window Help

Haystack

GSK3beta Topic

Target overview

DBP Lead, SB..., A..., CHI..., AKAPaulic NCE, CHI... targets targets targets targets targets targets

GSK-3beta

Group members

| Title | role | Department | E-mail |
|-------------|-------------------|-----------------|-----------------|
| John Tegler | Medicinal Chemist | Chemistry | john.tegler@... |
| Steve Smith | Synthetic Chemist | Chemistry | steve.smit@... |
| Tim Gross | Molecular Modeler | Cheminformatics | tim.gross@... |

Primary disease

Type 2 Diabetes

#125853 [Links](#)

DIABETES MELLITUS, NONINSULIN-DEPENDENT; NIDDM

Alternative titles; symbols

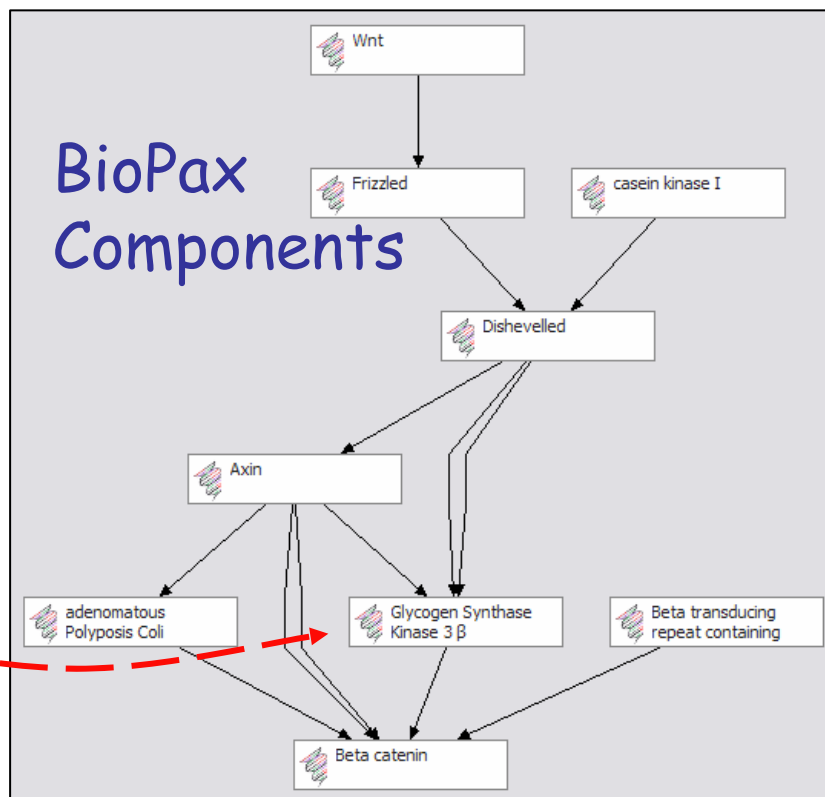
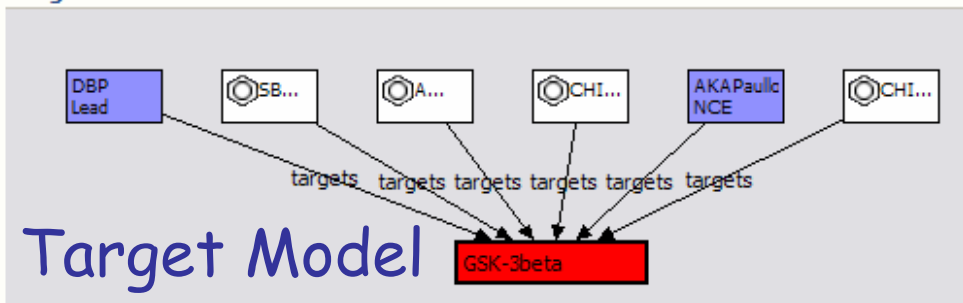
DIABETES MELLITUS, TYPE II
NONINSULIN-DEPENDENT DIABETES MELLITUS
MATURITY-ONSET DIABETES
INSULIN RESISTANCE, SUSCEPTIBILITY TO, INCLUDED

Gene map locus [20q12-q13.1](#), [20q12-q13.1](#)

Bridging Chemistry and Molecular Biology

Semantic Lenses: Different Views of the same data

Target overview



[urn:lsid:uniprot.org:uniprot:P49841](https://www.ebi.ac.uk/ols/ontologies/chem/targets/individuals/urn:lsid:uniprot.org:uniprot:P49841)

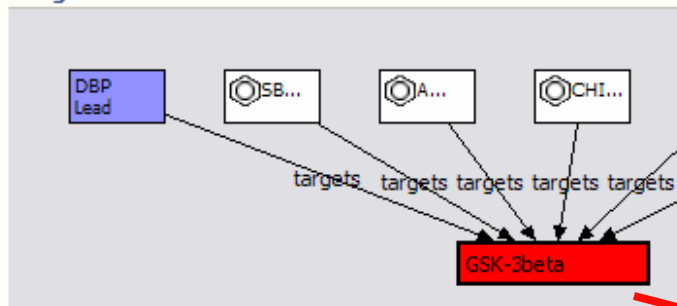
Apply Correspondence Rule:

if ?target.xref.lsid == ?bpx:prot.xref.lsid
then ?target.correspondsTo.?bpx:prot

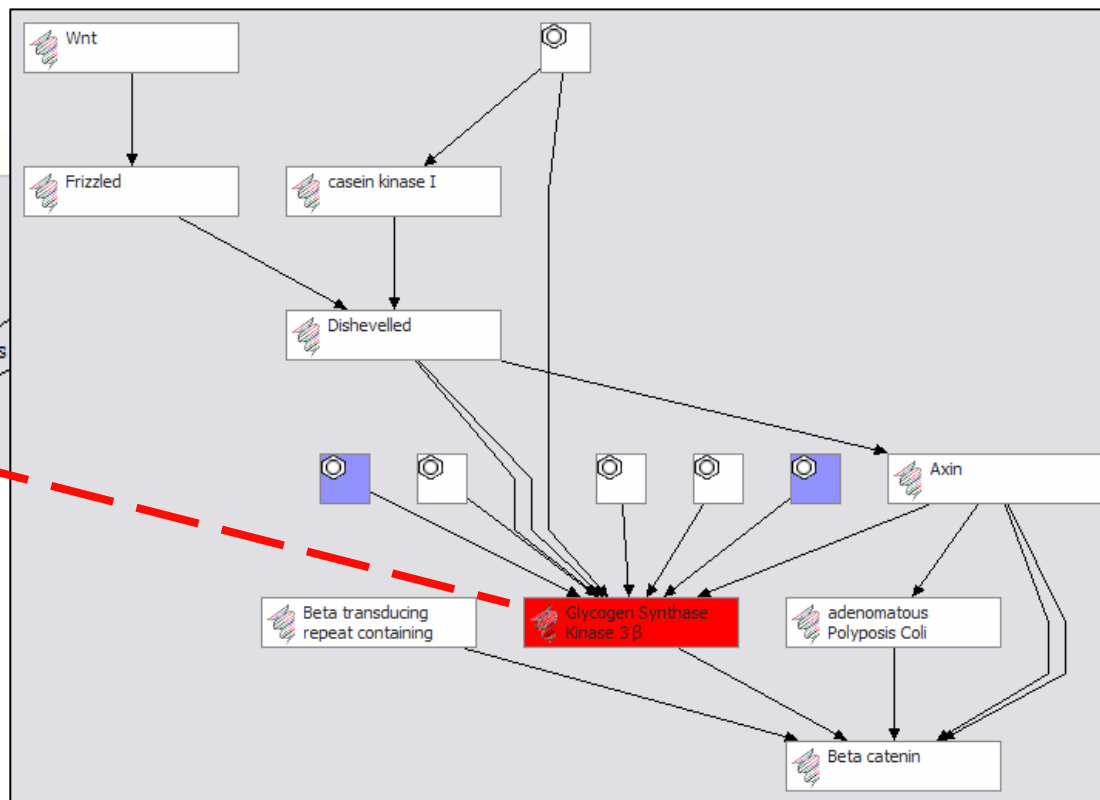
Bridging Chemistry and Molecular Biology

- Lenses can aggregate, accentuate, or even analyze new result sets
- Behind the lens, the data can be persistently stored as RDF-OWL

Target overview



- Correspondence does not need to mean "same descriptive object", but may mean objects with identical references



Case Study: Drug Safety


'Safety Lenses'

- Lenses can 'focus data in specific ways'
 - Hepatotoxicity, genotoxicity, hERG, metabolites
- Can be "wrapped" around statistical tools
- Aggregate other papers and findings (*knowledge*) in context with a particular project
- Align animal studies with clinical results
- Support special "Alert-channels" by regulators for each different toxicity issue
- Integrate JIT information on newly published mechanisms of actions

GeneLogic *GeneExpress* Data

- Additional relations and aspects can be defined additionally

Diseased Tissue



OVARY MALIGNANT

All Properties

label: OVARY MALIGNANT

normal form: OVARY NORMAL

omim: OVARIAN CANCER, EPITHELIAL

RDF Type: Sample

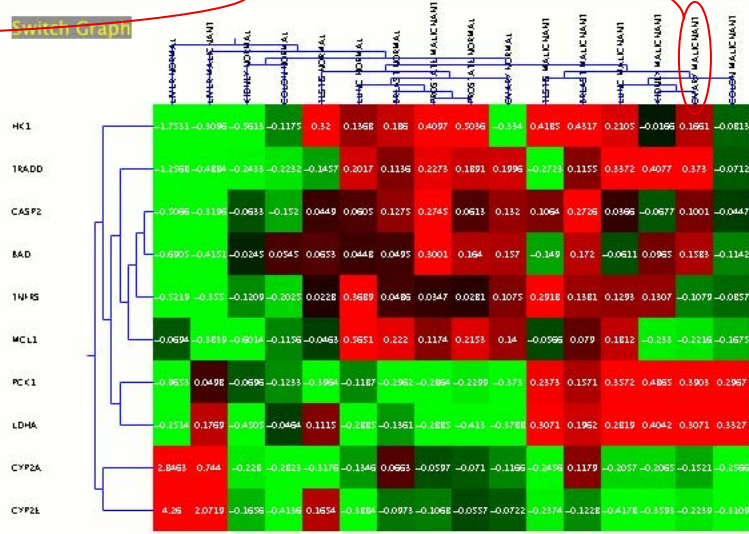
Links to OMIM (RDF)



Human GeneExpress Study

Tissue Study

Switch Graph

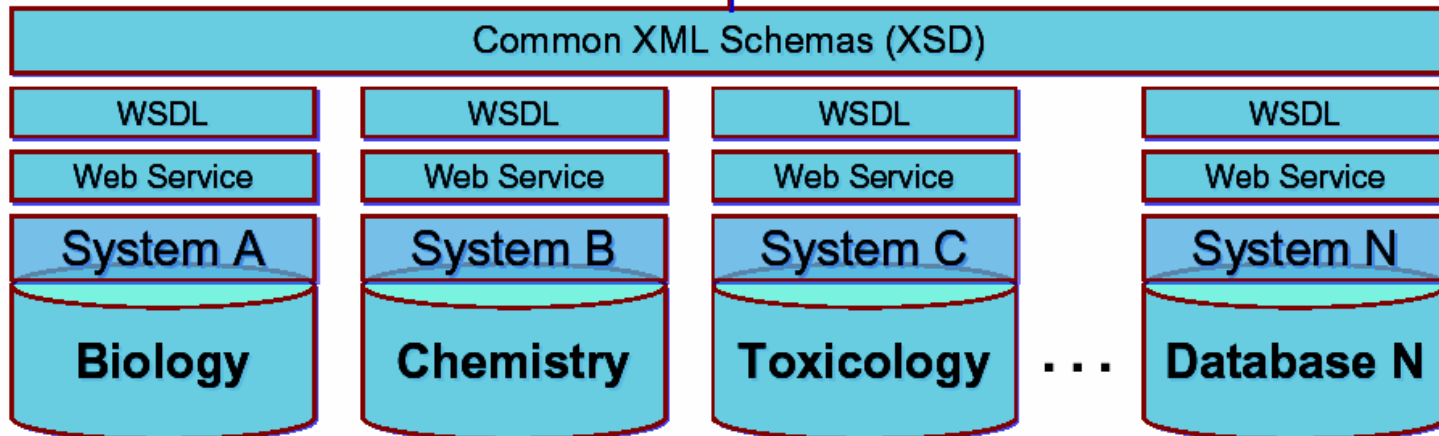
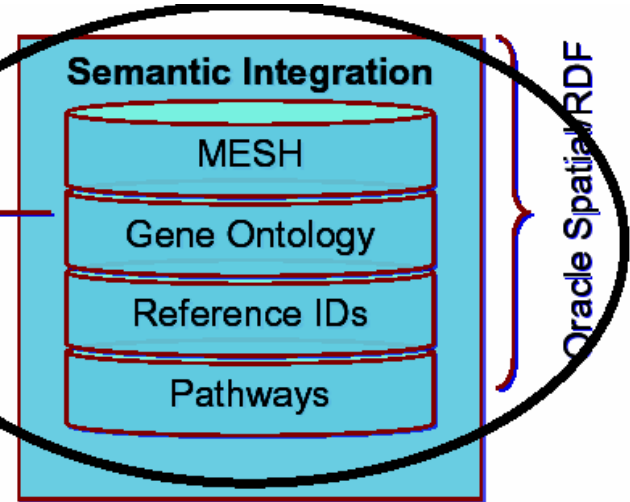
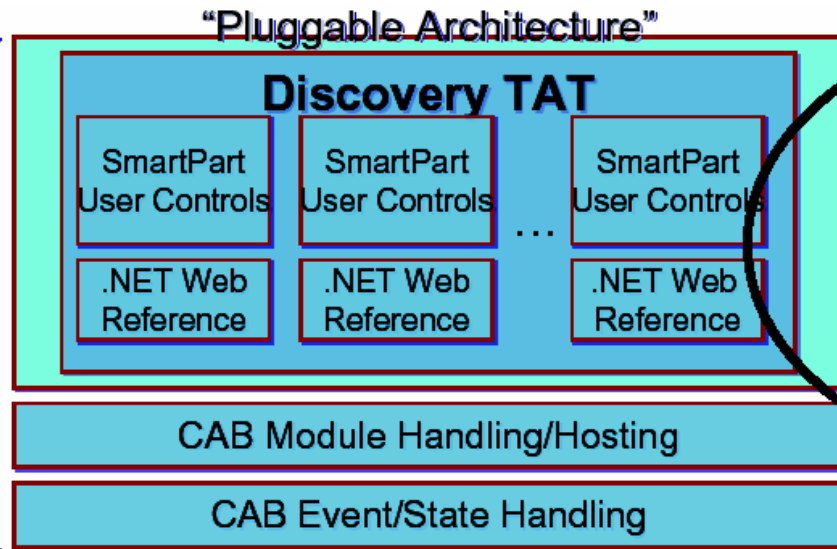


Tissue Study summary

| | | |
|---------------|---|-------------------------------------|
| Experimenter: | Li Liu*, Douglas M. Hawkins, Sujoy Ghosh and S. Stanley Y | <input type="button" value="Edit"/> |
| System: | Human Malignant and Normal Tissues | <input type="button" value="Edit"/> |
| Period: | 11/2/01-5/4/03 | <input type="button" value="Edit"/> |
| Design: | Human Tissue Gene Expression Protocol 7.3 | <input type="button" value="Edit"/> |

Drug Discovery Architecture

Visual Studio 2005/.NET 2.0
Microsoft Patterns & Practices
Composite Application Block



Perl CGI & Java
SOAP::Lite & Axis
Apache & Tomcat Web
.NET/IIS

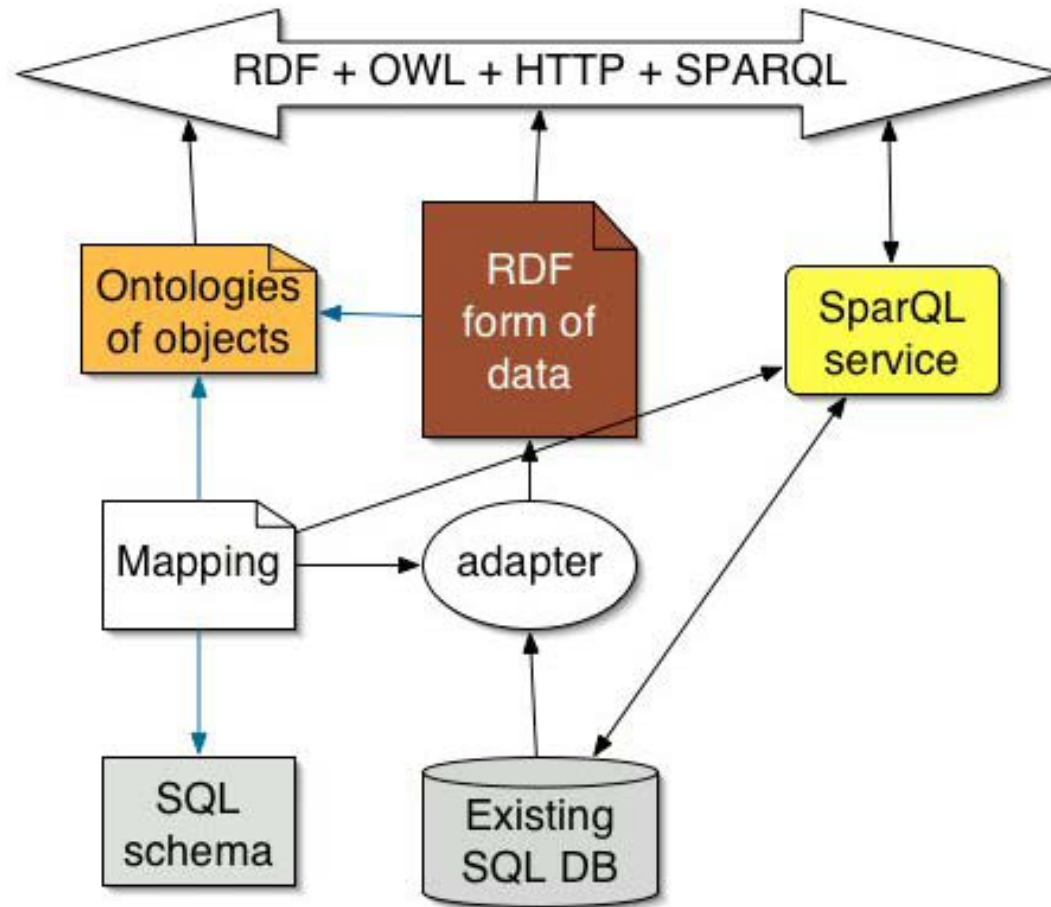
Advantages of Semantics

- RDF: Graph based data model
 - More expressive than the tree based XML Schema Model
- RDF: Reification
 - Same piece of information can be given different values of belief by different clinical genomic researchers
- Potential for “Schema-less” Data Integration
 - Hypothesis driven approach to defining mapping rules
 - Can define mapping rules on the fly
- Incremental approach for Data Integration
 - Ability to introduce new data sources into the mix incrementally at low cost
- Use of Ontology to disallow meaningless mapping rules?
 - For e.g., mapping a gene to a protein...

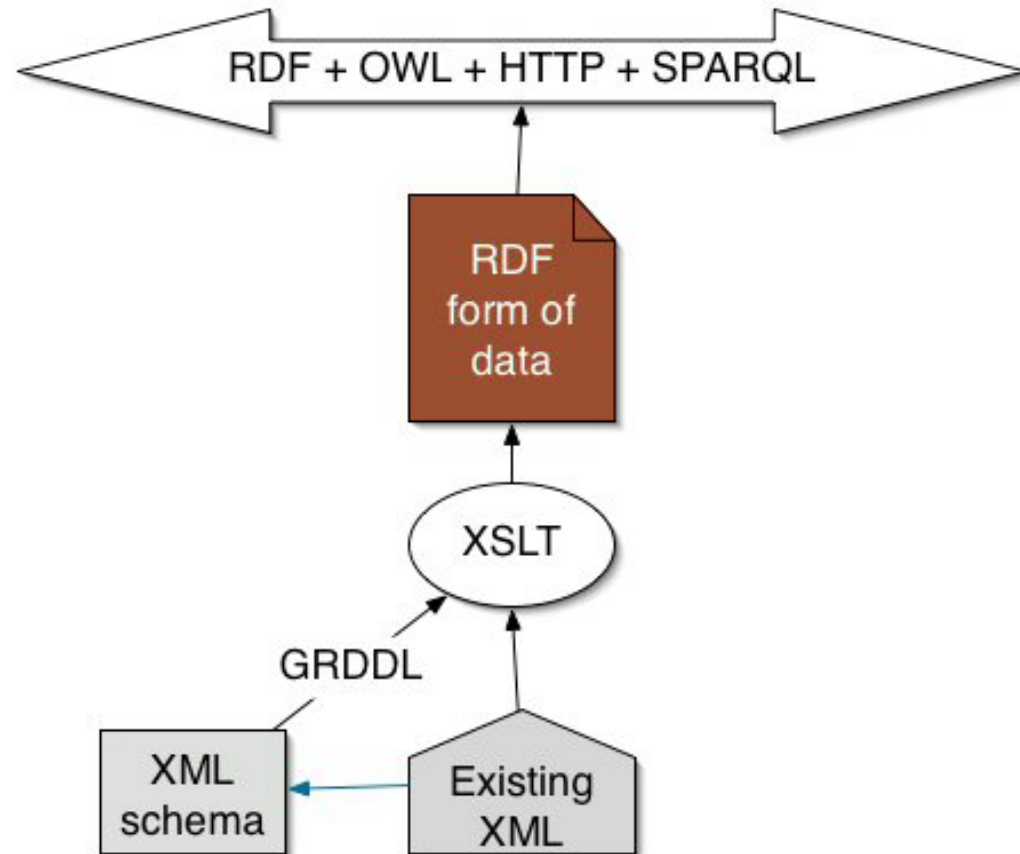
Information Integration: Two Approaches

- RDF Data Warehousing
 - Data is stored in a centralized data repository
 - Example: Oracle RDF Data Store
- RDF Wrappers/Interfaces
 - The data remains in the source repository
 - Create RDF Wrappers and Interfaces to expose and RDF view over the repository
 - Example: Jena RDF Engine and D2R Mapping Tool

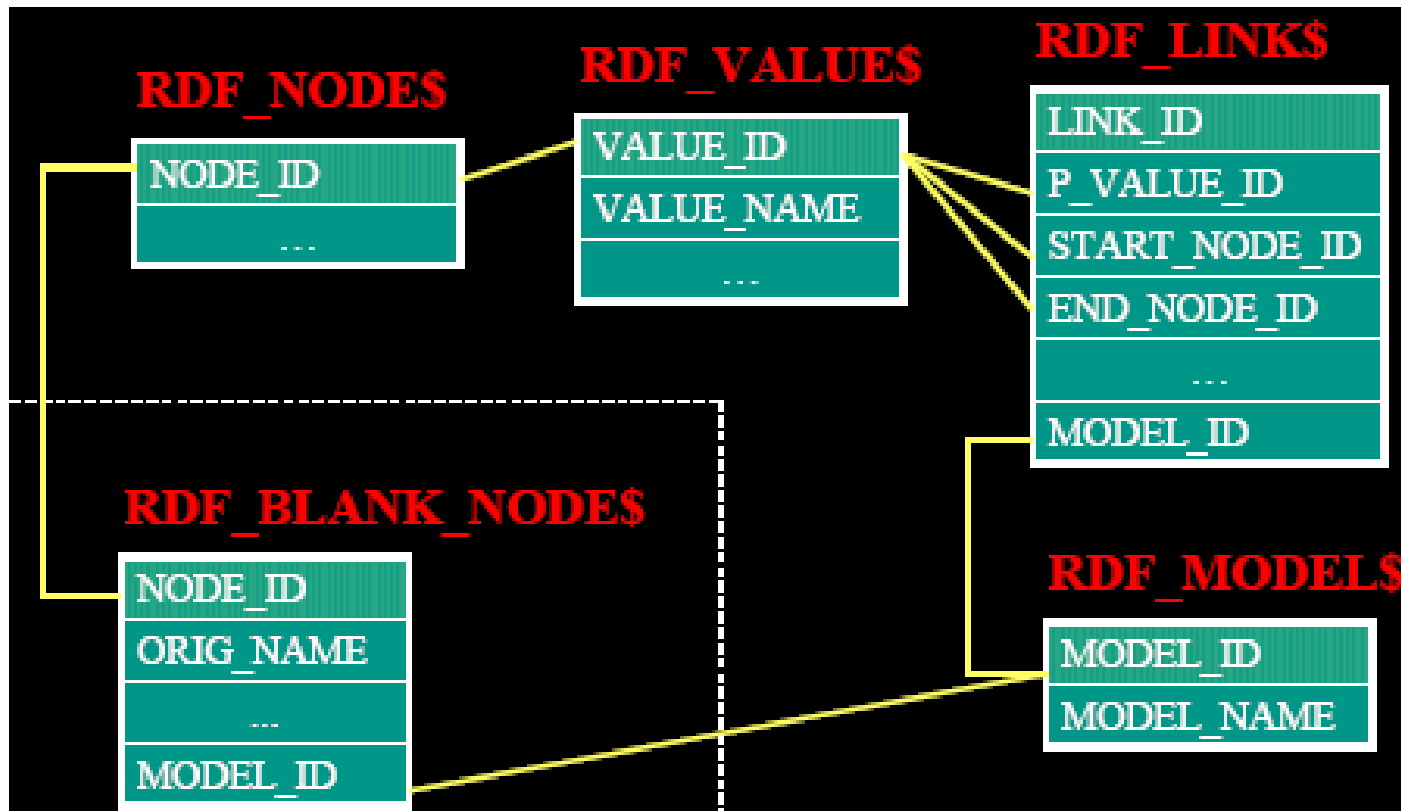
Adapting RDBMS for RDF



Adapting XML Schema to RDF



Oracle RDF Data Store: Schema



Oracle RDF Data Store: Hypothetical RDB to RDF Mappings

Employee Table

| EmpID | Name | SiteID | Role |
|--------------|-------------|---------------|-------------|
| 110 | S Hagan | 1 | Development |
| 120 | B Shimp | 2 | Marketing |
| 130 | K Bock | 3 | Sales |

SQL Component

Table EMPLOYEE

Column EMPLOYEE. EMPID

Row with EMPID=110

NAME cell in that row

URI relative to database

EMPLOYEE/schema#EMPLOYEE

EMPLOYEE/schema#empid

EMPLOYEE/rowBy/empid/110

EMPLOYEE/rowBy/empid/110#name

Oracle RDF Data Store: SPARQL-like Query Capability

- A table function allows a graph query to be embedded in a SQL query
- Searches for an arbitrary pattern against the RDF data
- Includes inferencing, based on RDF, RDFS, and user-defined rules
- Automatically resolve multiple representations of the same point in value space, e.g. 1 vs. 1.00

RDF Querying Problem

- Given
 - RDF graphs: the data set to be searched
 - Graph Pattern: containing a set of variables
- Find
 - Matching Subgraphs
- Return
 - Sets of variable bindings: where each set corresponds to a Matching Subgraph

RDF Query Example

Family Data: (:John :brotherOf :Mary)
(:Mary :parentOf :Matt)
(:John :name "John")
(:Mary :name "Mary")
(:Matt :name "Matt")

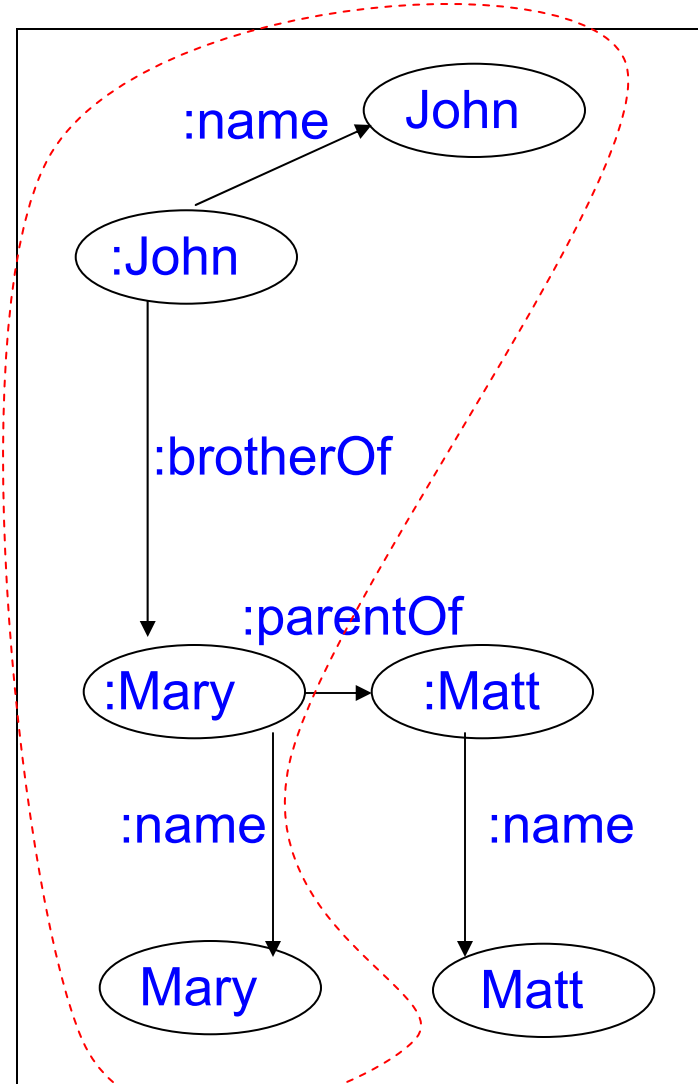
Graph Pattern: (names of Mary's brothers)
(?x :brotherOf ?y)
(?y :name "Mary")
(?x :name ?n)

Variable Bindings:

x = :John, y = :Mary, n = "John"

Matching Subgraph:

(:John :brotherOf :Mary)
(:Mary :name "Mary")
(:John :name "John")



Matching Subgraph

RDF Querying Issues

- Support specification of graph pattern-based SQL query
- Occurrence of same variables in multiple triples of graph pattern: Processing requires self-join
 - e.g.

| | | |
|-----|------------|---------|
| (?x | :brotherOf | ?y) |
| (?y | :name | “Mary”) |
| (?x | :name | ?n) |
- Query processing (e.g for filter conditions, ORDER BY) requires datatype-specific comparison semantics

Schema Triple: (:age rdfs:range xsd:int)

Graph Pattern: (?x :age ?a)

Filter Condition: a > 60

ORDER BY: a DESCENDING

RDF Querying Issues: Inference

- Query processing may involve Inferencing
- Example:

Data: (:Jim :brotherOf :John) (:John :fatherOf :Mary)
Graph Pattern:(?x :uncleOf ?y)
Result: Empty

Rule:
(?x :brotherOf ?y) (?y :fatherOf ?z)
→ (?x :uncleOf ?z)
Inferred data: (:Jim :uncleOf :Mary)
Result: x = :Jim, y = :Mary

RDF Querying Approach

- General Approach
 - Create a new (declarative, SQL-like) query language
 - e.g.: RQL, SeRQL, TRIPLE, N3, Versa, SPARQL, RDQL, RDFQL, SquishQL, RSQL, etc.
- SQL-based Approach
 - Introduces a *SQL Table Function* `RDF_MATCH` that uses SPARQL-like graph pattern to express RDF queries
- Benefits of SQL-based Approach
 - Leverages all the powerful constructs in SQL (e.g., `SELECT / FROM / WHERE`, `ORDER BY`, `GROUP BY`, aggregates, Join) to process graph query results
 - RDF queries can easily be combined with conventional queries on database tables thereby avoiding staging

Embedding RDF Query in SQL

- **SELECT ...**
FROM ...,
TABLE (

RDF Query
(expressed as RDF_MATCH
Table Function invocation)

) t, ...
WHERE ...;

- Use of RDF_MATCH Table Function allows embedding a graph query in a SQL query

RDF_MATCH Table Function

- Input parameters

```
RDF_MATCH (  
  Pattern,           ← graph pattern  
  Models,           ← Data (set of RDF graphs)  
  RuleBases,       ← Rules (0 or more rulebases)  
  Aliases           ← list of prefixes for namespaces  
)
```

- Returns a set of columns containing variable bindings
 - Variable matching URI returned as single VARCHAR2 column with the same name (e.g. x for ?x)
 - Variable matching literal returned as a pair of VARCHAR2 columns with a name (e.g. x for ?x) and the type (x\$type for ?x)

RDF_MATCH Example

- Example: student reviewers less than 25 years old

```
SELECT t.r reviewer, t.c conf, t.a age
```

```
FROM TABLE (
```

```
  RDF_MATCH (
```

```
    '(?r rdf:type :Student)
```

```
    (?r :reviewerOf ?c)
```

```
    (?r :age ?a)',
```

```
    RDFModels('reviewers'),
```

```
    NULL,
```

```
    RDFAliases(...))
```

```
) t
```

```
WHERE t.a < 25;
```

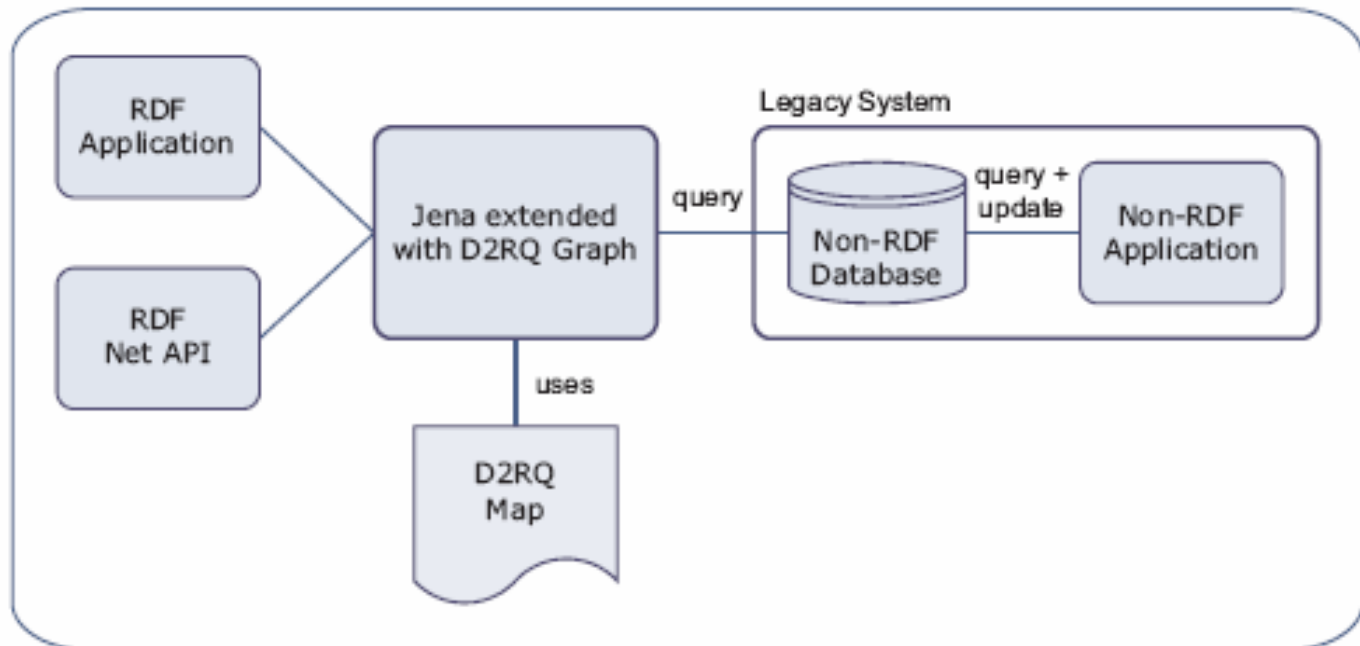
UniProt Sample Queries using Oracle RDF Data Store

UniProt Sample Queries

| Description | Pattern | Projection | Result limit |
|--|--------------------------------------|------------|--------------|
| Q1: Display the ranges of transmembrane regions | 6 triples 5 vars | 3 vars | 15000 rows |
| Q2: List proteins with publications by authors with matching names | 5 triples 5 vars 1 LIKE pred. | 3 vars | 10 rows |
| Q3: Count the number of times a publication by a specific author is cited | 3 triples 2 vars | 0 vars | 32 rows |
| Q4: List resources that are related to proteins annotated with a specific keyword | 3 triples 2 vars | 1 var | 3000 rows |
| Q5: List genes associated with human diseases | 7 triples 5 vars | 3 vars | 750 rows |
| Q6: List recently modified entries | 2 triples 2 vars 1 range pred. | 2 vars | 8000 rows |

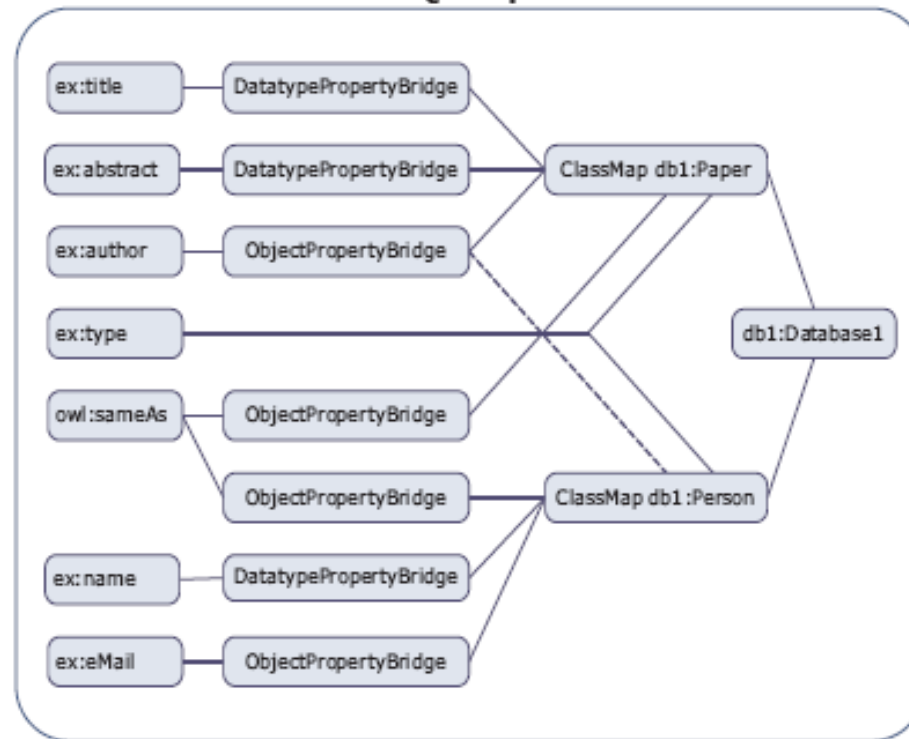
(Chong et al, VLDB Conference, Sep. 2005 – from Erik Jain's work?)

RDF Applications: Jena



Mapping RDBMS to RDF: D2RQ

Structure of the D2RQ Map



RDF Applications: Jena

RDQL Query

```
SELECT ?x
WHERE (<http://www.papers.org/3465>,
      ex:author, ?x)
```

SQL Query



```
SELECT Persons.Email
FROM PersonPaper, Persons
WHERE PersonPaper.PersonID =
      Persons.PersonID
AND PersonPaper.PaperID = 3465;
```

Databases Tables

Persons

| PersonID | Name | Email |
|----------|-------------------|-----------------------|
| 1 | Andy Seaborne | andy.seaborne@tpc.com |
| 2 | Christian B. Lenz | christ@dbi.acr.de |

Papers

| PaperID | Title |
|---------|---|
| 3465 | RDQL-Treating Non RDF Databases as Virtual RDF Graphs |
| 3728 | The Impact of Semantic Web Technologies on Job Search |

PersonPaper

| PaperID | PersonID |
|---------|----------|
| 3465 | 2 |
| 3465 | 1 |
| 3728 | 2 |

Advantage: “Schema-free” data integration

- Low cost approach for data integration
- No need for maintenance of costly schema mappings
- Ability to “merge” RDF graphs based on simple declarative rules that specify:
 - Equality of URIs
 - Connecting nodes of same type
 - Connecting two nodes associated by a “path”
- Disadvantage: Potential for specifying spurious non-sensical rules

“Schema-free” Integration Example

- Match nodes with the same URIs
 - Can be represented using OWL same-as
- If a patient’s structured test result indicates a disease, then the patient suffers from the disease
 - `has_structured_test_result.indicates_disease = suffers_from`
 - Needs Rules for representation: Role composition not supported by OWL.

Rule-based approaches

- Rule ML
- Semantic Web Rules Language (SWRL)
- Production Rules Systems
 - ILOG
 - Blaze
- Open Source Rule Engines
 - JESS (Forward chaining, production rules)
 - XSB (Backward chaining, prolog implementation)

Rule ML

```
<imp>
  <_head>
    <atom>
      <rel>suffers_from</rel>
      <var>patient</var>
      <ind>"Hypertrophic Cardiomyopathy"</var>
    </atom>
  </_head>
  <_body>
    <and>
      <atom>
        <rel>has_structured_result</rel>
        <var>patient</var>
        <var>test_result</var>
      </atom>
      <atom>
        <rel>indicates</rel>
        <var>test_result</var>
        <var>"Hypertrophic Cardiomyopathy"</var>
      </atom>
    </and>
  </_body>
</imp>
```

SWRL

```
Implies(Antecedent(Patient(I-variable(x)),  
                  TestResult(I-variable(y)),  
                  has_structured_result(I-variable(x), I-variable(y)),  
                  indicates(I-variable(y), "Hypertrophic Cardiomyopathy")),  
        Consequent(suffers_from(I-variable(x), "Hypertrophic Cardiomyopathy")))
```

Production Systems: ILOG

Class Patient: Person

method get_name(): string;
method has_genetic_test_result(): StructuredTestResult;
method has_mutation(): string;
method suffers_from(): Disease;
method set_suffers_from(Disease): void;

Class StructuredTestResult

method get_patient(): Patient;
method indicates_disease(): Disease;
method identifies_mutation(): set of string;
method evidence_of_mutation(string): real

the_result = the_patient.has_genetic_test_result();

IF the_patient.has_genetic_test_result().indicates_disease()
= "Hypertrophic Cardiomyopathy"

THEN

the_patient.set_suffers_from("Hypertrophic Cardiomyopathy")

Some thoughts on Rules

- Ease of use: SWRL is comparatively easier to use
- Expressivity: SWRL expressivity is less than RuleML
- Ontological Underpinnings: SWRL enables incorporation of OWL classes and constraints.
- Re-uses Rule ML and OWL standards, e.g., tags, XML, etc.
- Production Systems such as ILOG and Blaze have their own internal XML-based representation format.
 - SWRL could be a potential interchange language
 - Issues currently being investigated in the W3C RIF working group

Reification: Annotation of existing data

- Level of accuracy of test result.
 - Sensitivity and Specificity of lab result
 - Level of confidence in genotyping or gene sequencing
- Probabilistic relationships
 - Likelihood that a particular test result or condition is indicative of a disease or other medical condition
- Level of trust in a resource
 - Results from a lab may be trusted more than result from another
 - Results from well known health sites (NLM) may be trusted more than others
- Belief attribution
 - Scientific hypotheses may be attributed to appropriate researchers

Functional Requirements

- Use Cases
- Metadata-based Semantic Annotation
- *Recombinant Data*: Information Integration and Aggregation
 - Discovery Decision Support
- Ontology-driven Decision Support
- Knowledge Update and Maintenance

Clinical and Genomic Decision Support

IF the patient's LDL test result is greater than 120
AND the patient has a contraindication to Fibric Acid
THEN

Prescribe Zetia Lipid Management Protocol

Contraindication to Fibric Acid: Clinical Definition (Old)

The patient is contraindicated for Fibric Acid if he has an allergy to Fibric Acid or has elevated Liver Panel

Contraindication to Fibric Acid: Clinical+Genomic Definition (New)

The patient is contraindicated for Fibric Acid if he has an allergy to Fibric Acid or has elevated Liver Panel or has a genetic mutation Missense: XYZ3:Ser@\$#Pro

Please note: Hypothetical – assume a genetic variant is a biomarker for patients contraindicated to Fibric Acid.

Clinical and Genomic Decision Support: A Rules-based Implementation

Business Object Model Design

Class Patient: Person

```
method get_name(): string;
method has_genetic_test_result(): StructuredTestResult;
method has_liver_panel_result(): LiverPanelResult;
method has_ldl_result(): real;
method has_contraindication(): set of string;
method has_mutation(): string;
method has_therapy(): set of string;
method set_therapy(string): void;
method has_allergy(): set of string;
Method get_category(): set of string;
```

Class StructuredTestResult

```
method get_patient(): Patient;
method indicates_disease(): Disease;
method identifies_mutation(): set of string;
method evidence_of_mutation(string): real;
```

Class LiverPanelResult

```
method get_patient(): Patient;
method get_ALP(): real;
method get_ALT(): real;
method get_AST(): real;
method get_Total_Bilirubin(): real;
method get_Creatinine(): real;
```

Clinical and Genomic Decision Support: A Rules-based Implementation

Rule base Design

```
IF the_patient.has_ldl_result() > 120
AND ((the_patient.has_liver_panel_result().get_ALP() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_ALT() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_AST() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_Total_Bilirubin() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_Creatinine() ≥ <NormalRange>)
OR "Fibric Acid Allergy" ∈ the_patient.has_allergy()
OR "Missense: XYZ3:Ser@ $#Pro" ∈ the_patient.has_mutation())
THEN
  the_patient.set_therapy("Zetia Lipid Management Protocol")
```

Definition of “Fibric Acid Contraindication”

Clinical and Genomic Decision Support: Definitions vs Decisions

Commonly occurring design pattern:

- The definition of a “Fibric Acid Contraindication” is represented using rules.
- The decision related to therapeutic intervention is also represented using rules.

Currently, both these inferences are performed by the rules engine.

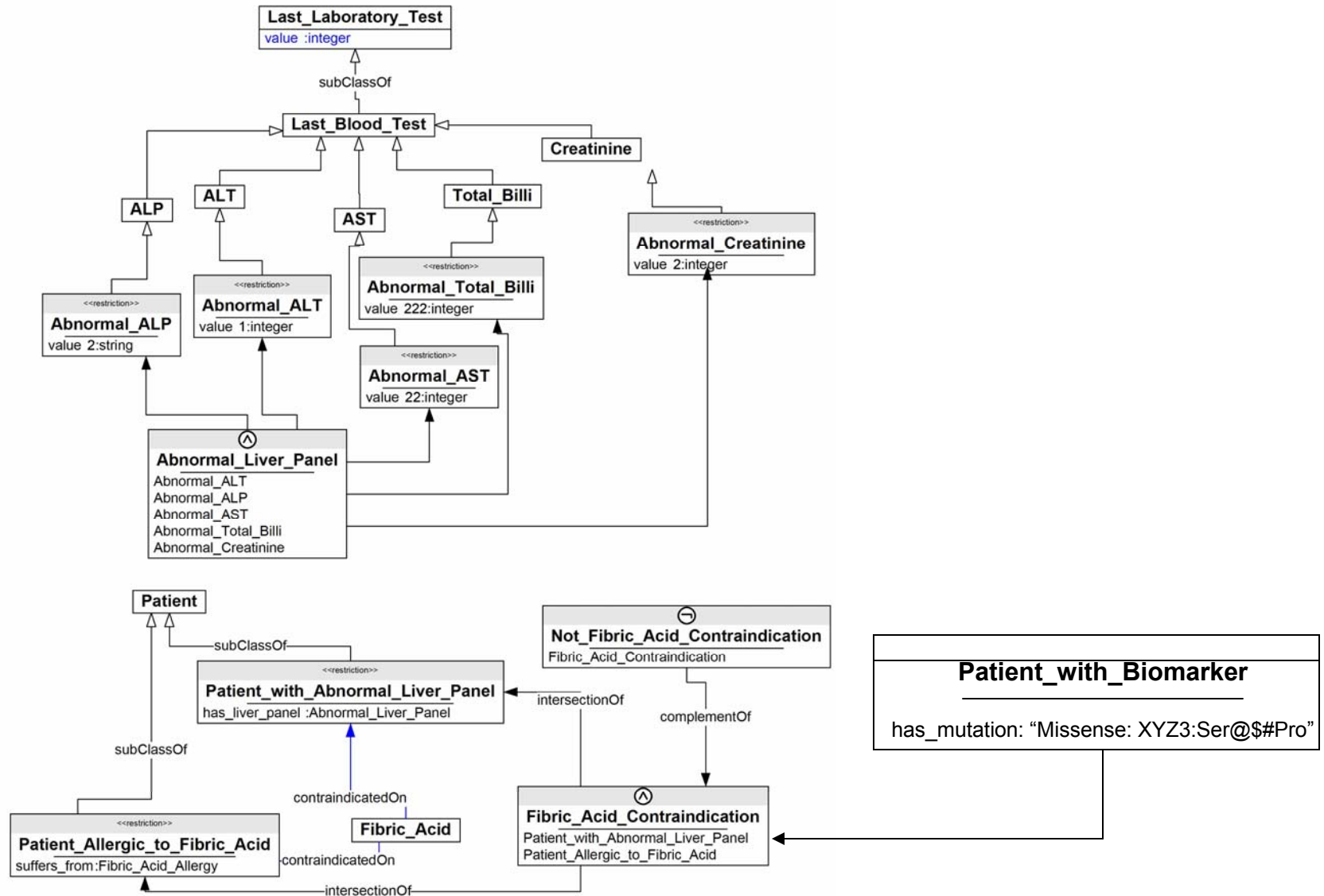
Clinical and Genomic Decision Support

Role of Ontology Engine

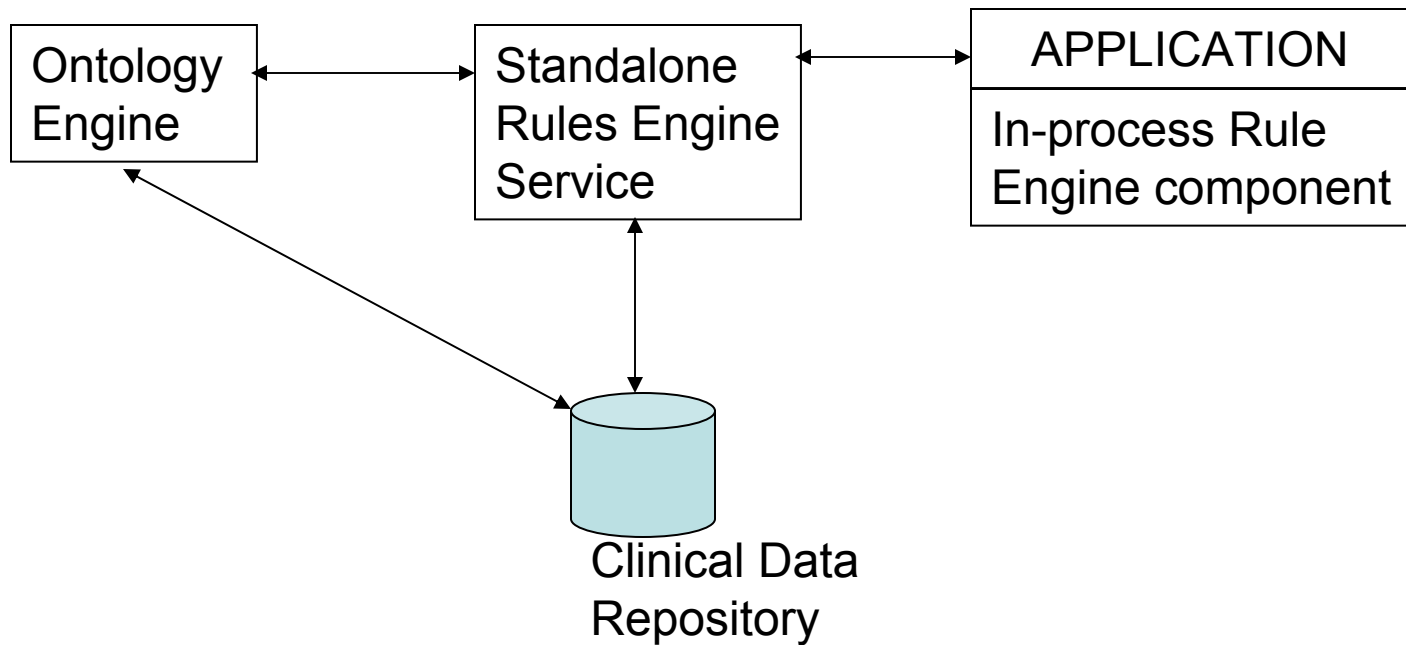
```
IF the_patient.has_ldl_result() > 120
AND the_patient.get_category() = PatientWithFibricAcidContraindication
THEN
set the_patient.has_therapy("Zetia Lipid Management Protocol")
```

implemented in an OWL-based
ontology engine

OWL representation of Fibric Acid Contraindication



Ontology Driven Clinical Decision Support: Architecture



Clinical Decision Support:

Decoupling definitions vs decisions

- Classification inferences (does patient have a fibrin acid contraindication?) can be evaluated by an ontology engine.
- Reduces overhead on Rule Engine
- Opens up the possibility of plugging-in other specialized inference engines (e.g., spatio-temporal conditions)
- Makes knowledge maintenance easier
 - Each definition may be referred to in 100s of rules..

Decision Support: Statistical vs Symbolic Approaches

- Symbolic:
 - Knowledge Driven: Needs input of Subject Matter Experts
 - Not scaleable: Knowledge Bases can get huge in case of interacting conditions
 - Example:
 - Set of Rules for “CAD”
 - Set of Rules for “Diabetes”
 - What about rules for “Diabetes” and “CAD”
 - In general for N conditions, the Knowledge base size can be of the order of 2^N .
- Statistical:
 - Data Driven: Models can be “learned” from the data
 - More scaleable
 - Probabilistic conclusions, Thresholding required
 - **Blackbox: No explanations possible!**
- Hybrid: Need some combination of the two...

OWL Reasoners

- CEL
 - Polynomial time classifier for the description logic EL+
 - EL+ is specially geared towards biomedical ontologies
- Cerebra
 - Commercial C++ reasoner, Support for OWL-API
 - Tableaux based reasoning for TBoxes and ABoxes
- Fact++
 - Free open source reasoner for DL reasoning
 - Support for Lisp API and OWL API
- KAON2
 - Free Java based DL reasoner with support for SWRL fragment
 - Support for DIG API
- MSPASS
 - A generalized theorem prover for numerous logics, also works for DLs
- Pellet
 - Free open source Java based reasoner for DLs
 - Support for OWL, DIG APIs and Jena Interface
- RacerPro
 - Commercial lisp based reasoner for DLs
 - Support for OWL APIs and DIG APIs

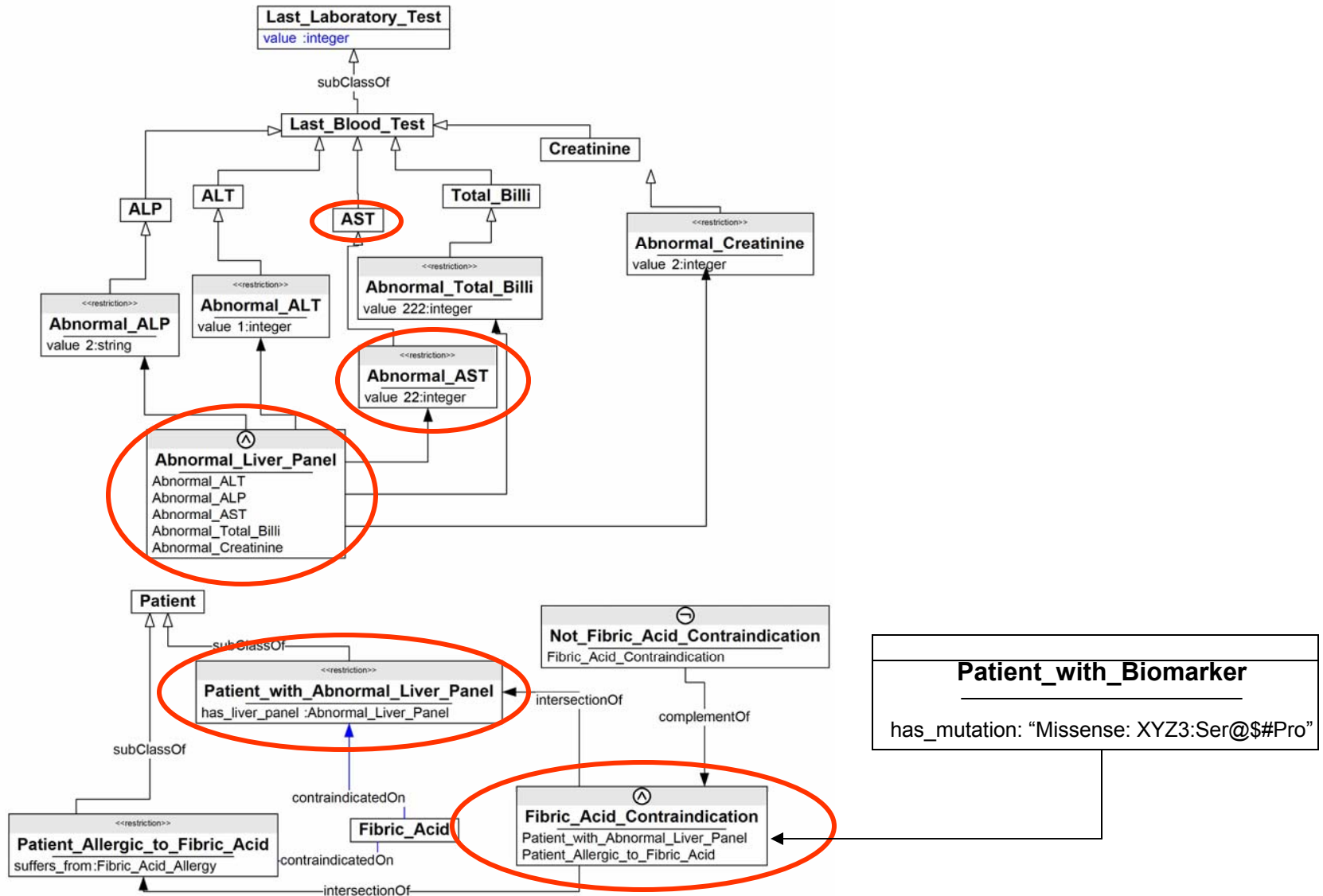
Functional Requirements

- Use Cases
- Metadata-based Semantic Annotation
- *Recombinant Data*: Information Integration and Aggregation
 - Discovery Decision Support
- Ontology-driven Decision Support
- Knowledge Update and Maintenance

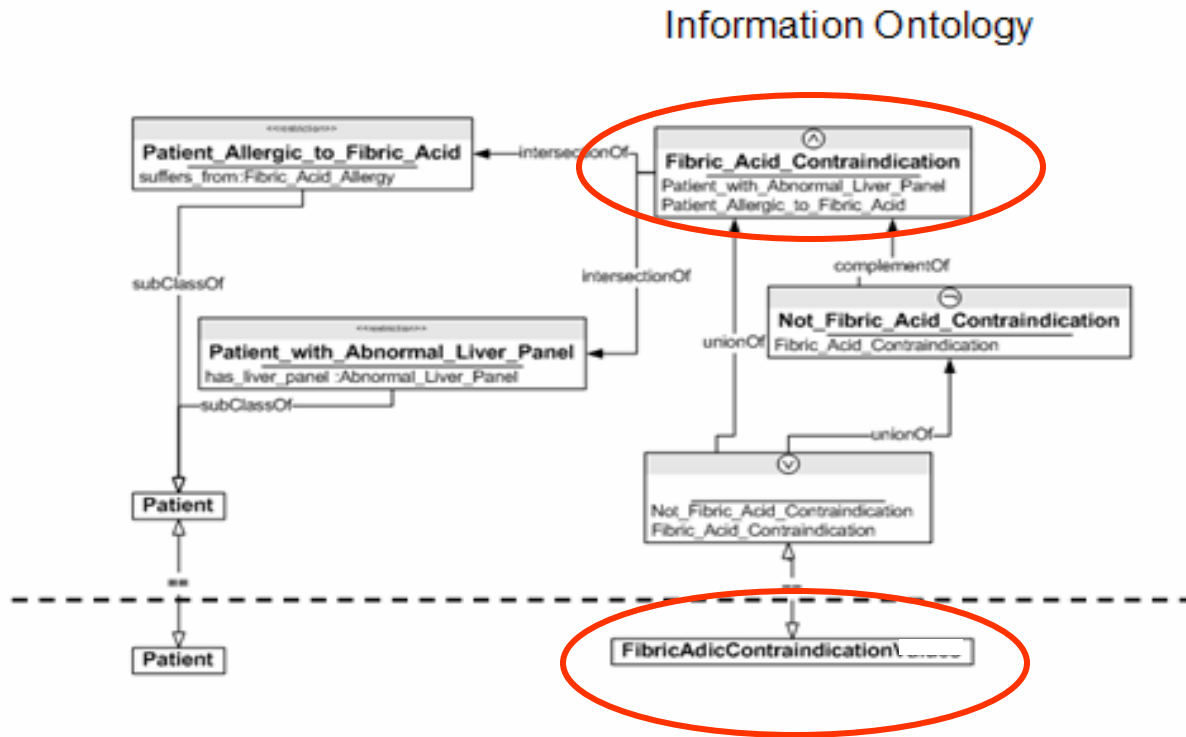
Knowledge Update and Maintenance

- There is rapid knowledge discovery and evolution in the Healthcare and Life Sciences
- Provenance is an important aspect of maintaining knowledge consistence
- There is a close interrelationship between knowledge change and provenance
 - What has changed? – Change
 - Why did it change? – Provenance
 - Did someone change it? – Provenance
 - Did its components change? – Change
 - Who changed it? – Provenance

Domain Ontology



Bridge – Composition Ontology



Rule base

Knowledge Change and Provenance

- At each stage, Knowledge Engineer gets notified of:
 - What has changed?
 - The definition of Fibrin Acid Contraindication
 - Why did it change?
 - Fibrin Acid Contraindication ← Patient with Abnormal Liver Panel ← Abnormal Liver Panel ← Abnormal AST ← Change in AST Values
 - Fibrin Acid Contraindication ← Patient with Biomarker
 - Who was responsible for the change?
 - Knowledge Engineer who entered the changed AST values?
 - Change in a Clinical Guideline?
 - New Molecular Diagnostic Test appears in the market?

Knowledge Update and Maintenance

- Knowledge Dependency Propagation
 - If the definition of a concept changes,
 - What other concepts does it impact?
 - What other clinical decision support rules does it impact?
- Assertion Dependency Propagation
 - If a clinical decision support rule is changed, how does this impact potential decisions made for a patient?
 - How and when should such decisions be updated?
 - What if the decision involved a drug which has already been administered?

Implementation Options

- Relational Databases
 - Set up RDBMS triggers or stored procedures
 - The semantics of the various dependencies are hard-coded in the application code associated with RDBMS triggers or stored procedures
- Rule Engines
 - Encode the semantics of the various dependencies of into rules.
 - Better than previous option as rules can be changed easier than application code
 - However, potential dependencies may be very high and depend upon the underlying KR language
- Ontology Engines and Reasoners
 - The reasoner can identify dependencies
 - Requires modeling knowledge appropriately in the OWL-DL model

Current State: Healthcare KM

- Market is very siloized
 - Different Vendors for Different Knowledge Types
- Order Sets and Templates
 - Next Gen, Zynx
- Vocabularies
 - HLi, IMO, Apelon
- Documentation Templates
 - IMDSOft
- Not Supported
 - Clinical Guidelines, Clinical Decision Support Rules, etc.
- Lack of a comprehensive semantics-based platform from KM

Semantics-based Knowledge Maintenance

- Managing change and provenance is a very difficult problem
- Semantics can play a crucial role in it:
 - A reasoner can navigate a semantic model of knowledge and propagate change
 - One can declaratively change the model at any time
 - The reasoner will compute the new changes!
- Configuration v/s coding. Could read to a huge ROI!
- Could be the potential “killer app” for the Semantic Web

Conclusions

- Healthcare and Life Sciences is a knowledge intensive field. The ability to capture semantics of this knowledge is crucial for implementation.
- Incremental and cost-effective approaches to support “as needed” data integration need to be supported.
- Scalable and modular approaches for decision support need to be designed and implemented.
- The rate of Knowledge Updates will change drastically as Genomic knowledge explodes. Automated Semantics-based Knowledge Update and Propagation will be key in keeping the knowledge updated and current
- Personalized/Translational Medicine cannot be implemented in a scalable, efficient and extensible manner without Semantic Web technologies