Applications and Architectures: Exploring Rate-Limiting Challenges to Interoperability

Workshop on Long Term Preservation and Management of Electronic Health Record

NIH/National Library of Medicine
Bethesda, Maryland, April 5-6, 2011

Lynn H. Vogel, Ph.D.
Vice President and Chief Information Officer
Associate Professor
Bioinformatics and Computational Biology
The University of Texas
MD Anderson Cancer Center
Houston, Texas

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

The fundamental healthcare question for Patients, Clinicians and Scientists

What happened **to the last 100 pati**

who had the ***same diagnosis,***

who have ***characteristics similar to mine,*** **and**

who were treated with the ***regimen*** **that the**

clinician is recommending?

Outcomes from Evidence-based Medicine

Personalized Medicine

Translational Research

In order to respond to this question, our historical views of business drivers will have to change. . .

- **Clinical Care and Research are becoming increasingly intertwined through more "personalized" medicine;**
- **New data sources, forms and structures are emerging rapidly, many without established standards**
- **New risks have entered our infrastructure**
  - **Internal and External**
  - **Consequences of not managing these risks can be significant**
- **Continuing cost pressures question how much data redundancy we can support over the longer run.**

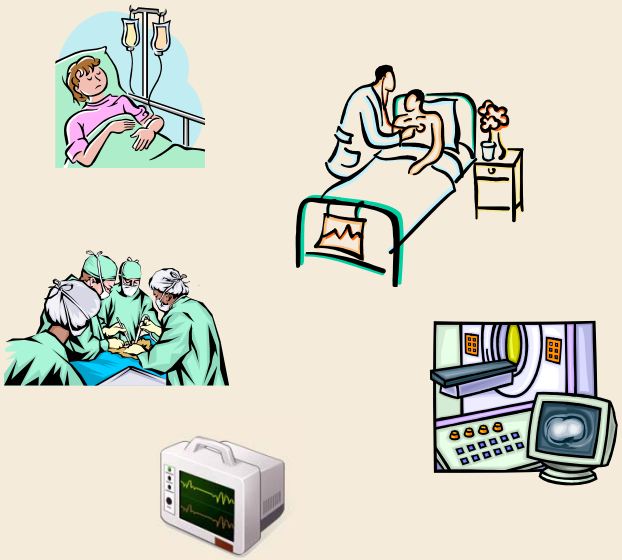
What interoperability means in Healthcare . . .

The ability for providers to have access to
the right data,
in the right place,
in the right form
when they need it !

. . . regardless of their location . . .
or the location of the source of the data!

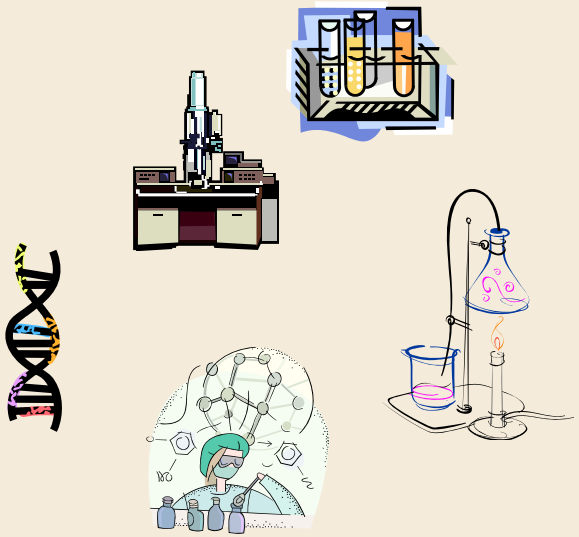
Historical Organizational Separation in Healthcare

Hospital/Office-based Clinical Care



Older architectures, acute care-focused, commercial vendors, software packages predominate, push to buy from single vendor, proprietary code

University-based Scientific Discovery

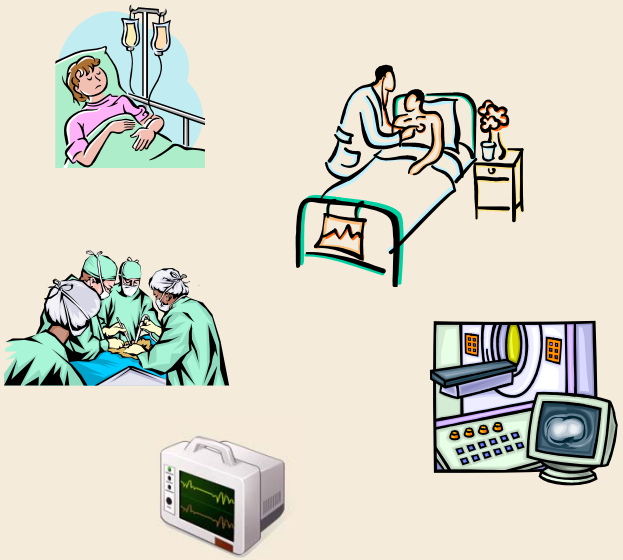


Isolated, Investigator-focused, few accepted data standards, software focused on small niches, open source coding preferred

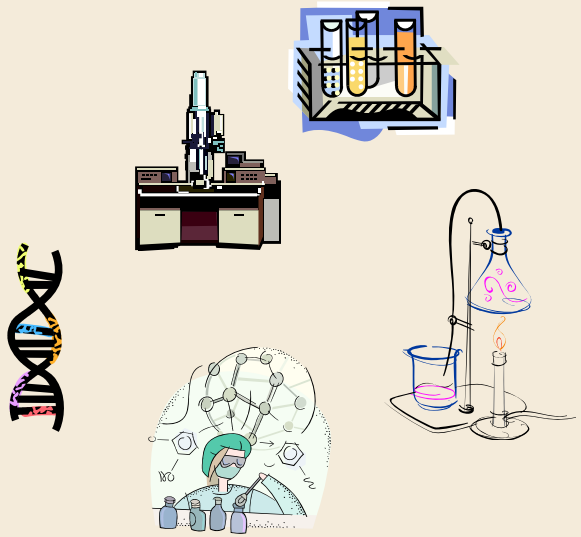
Information Technology

Personalized Medicine Challenges these Historical Separations

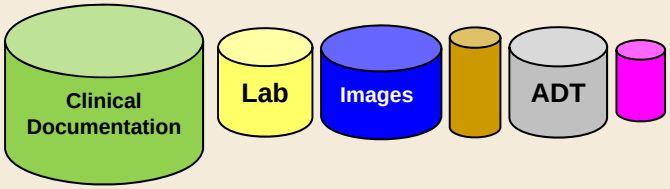
Hospital/Office-based Clinical Care



University-based Scientific Discovery

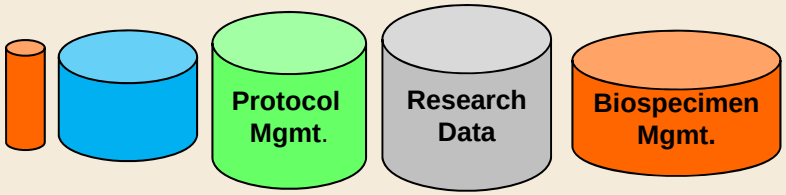


Clinical Data Sources

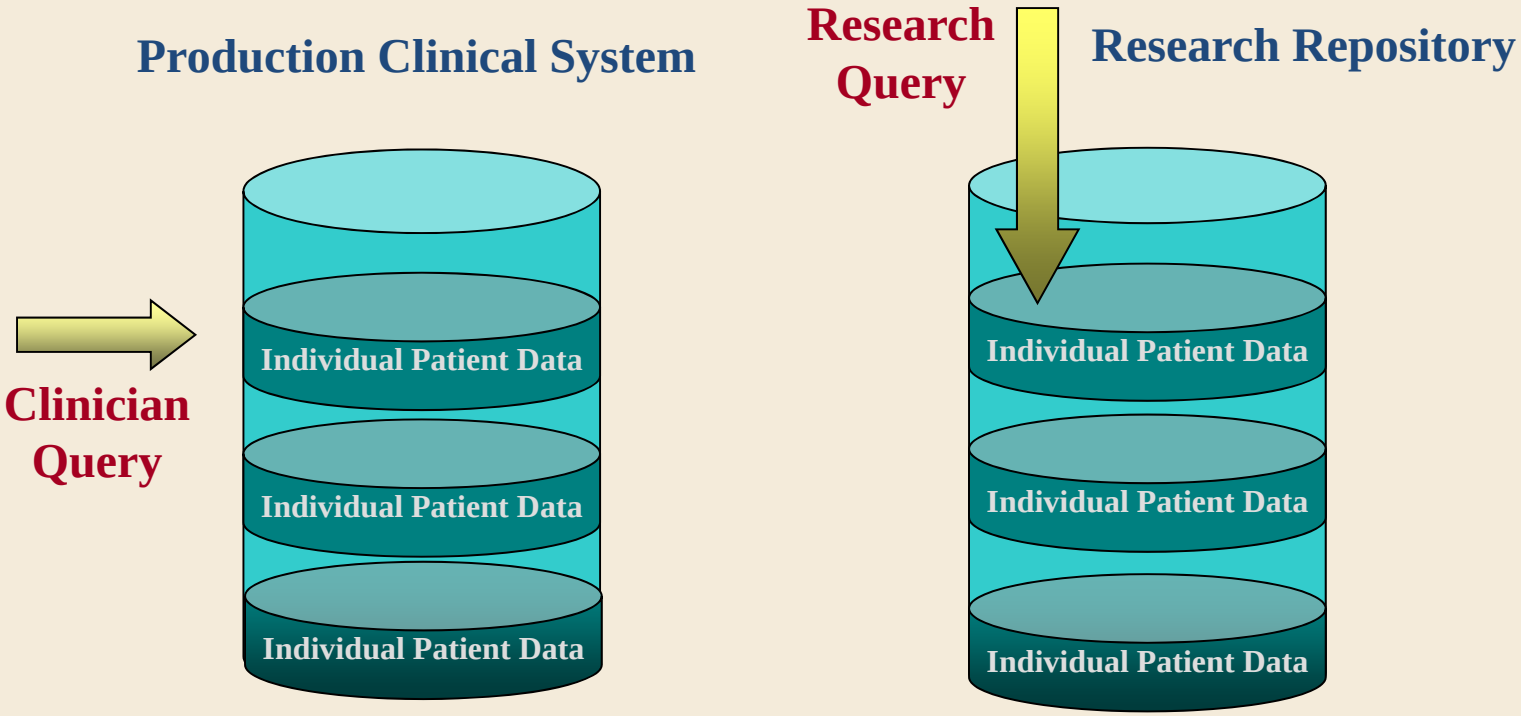


Information Technology

Research Data Sources



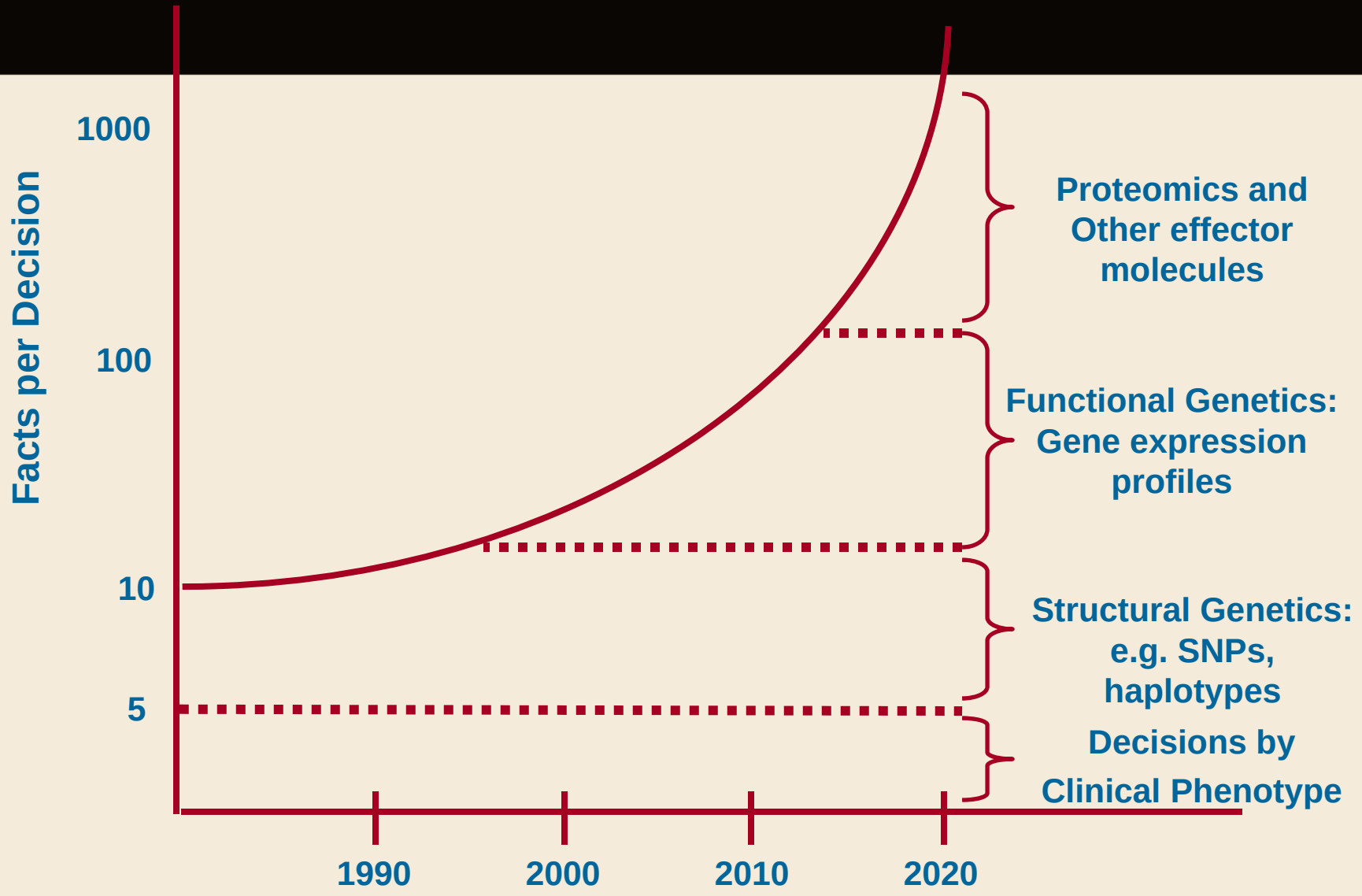
Challenges of integrating Data from Clinical Practice and Scientific Discovery



Focus on Single Patient, Many Attributes

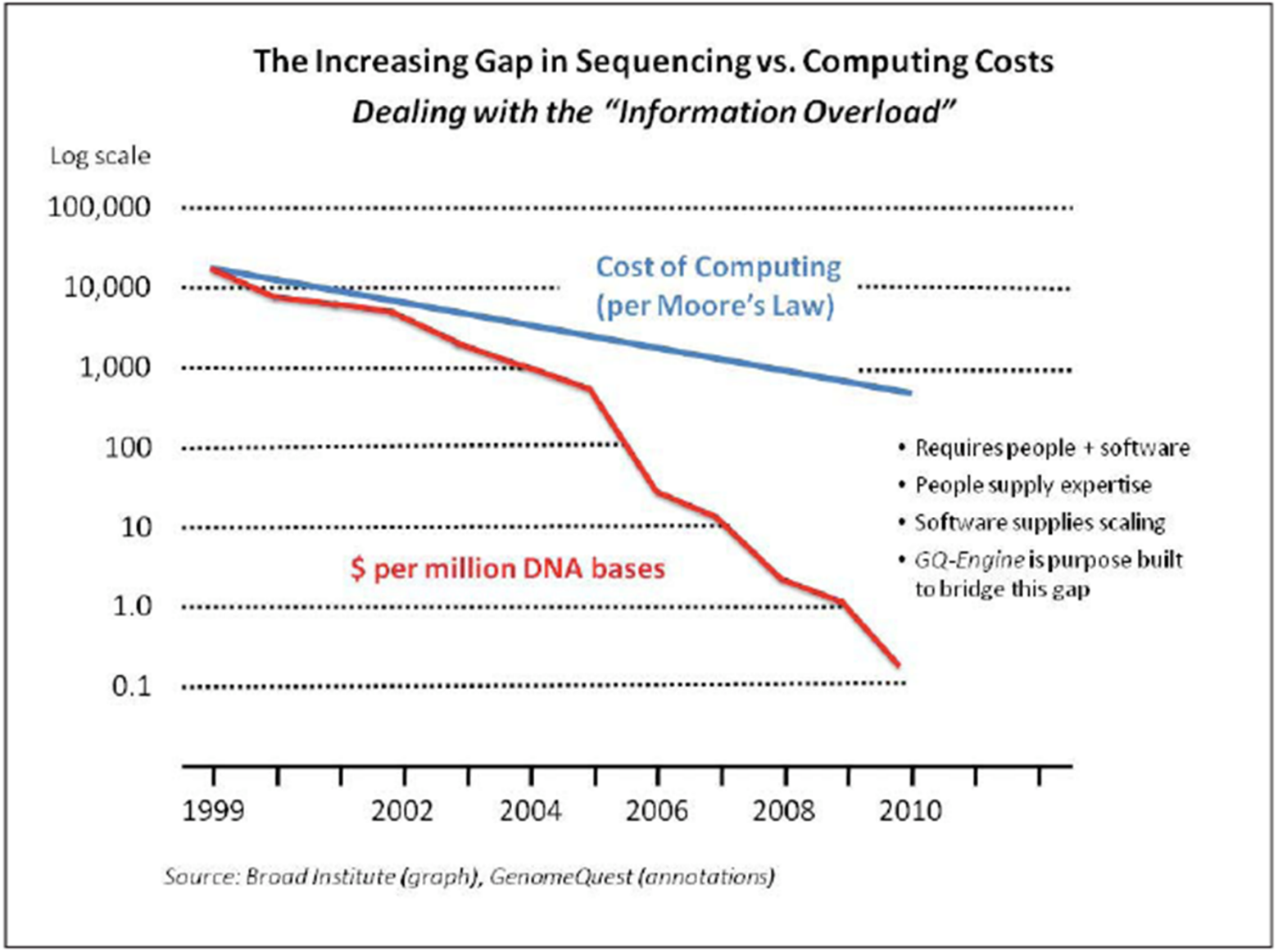
Focus on Many Patients, Few Attributes, Algorithm-based

Clinical decision-making is becoming more complicated



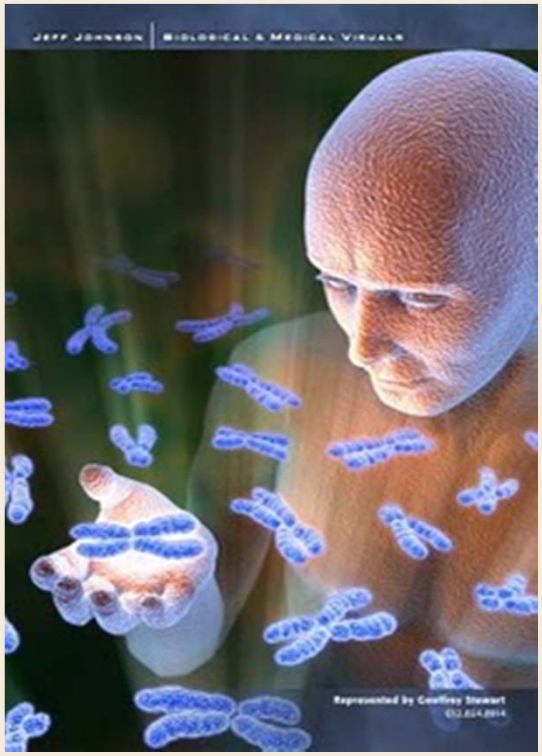
With appreciation to William W. Stead, M.D., 2007 AMIA Panel Presentation, "Why We Need Internal Development", November 11, 2007

A PERFECT STORM

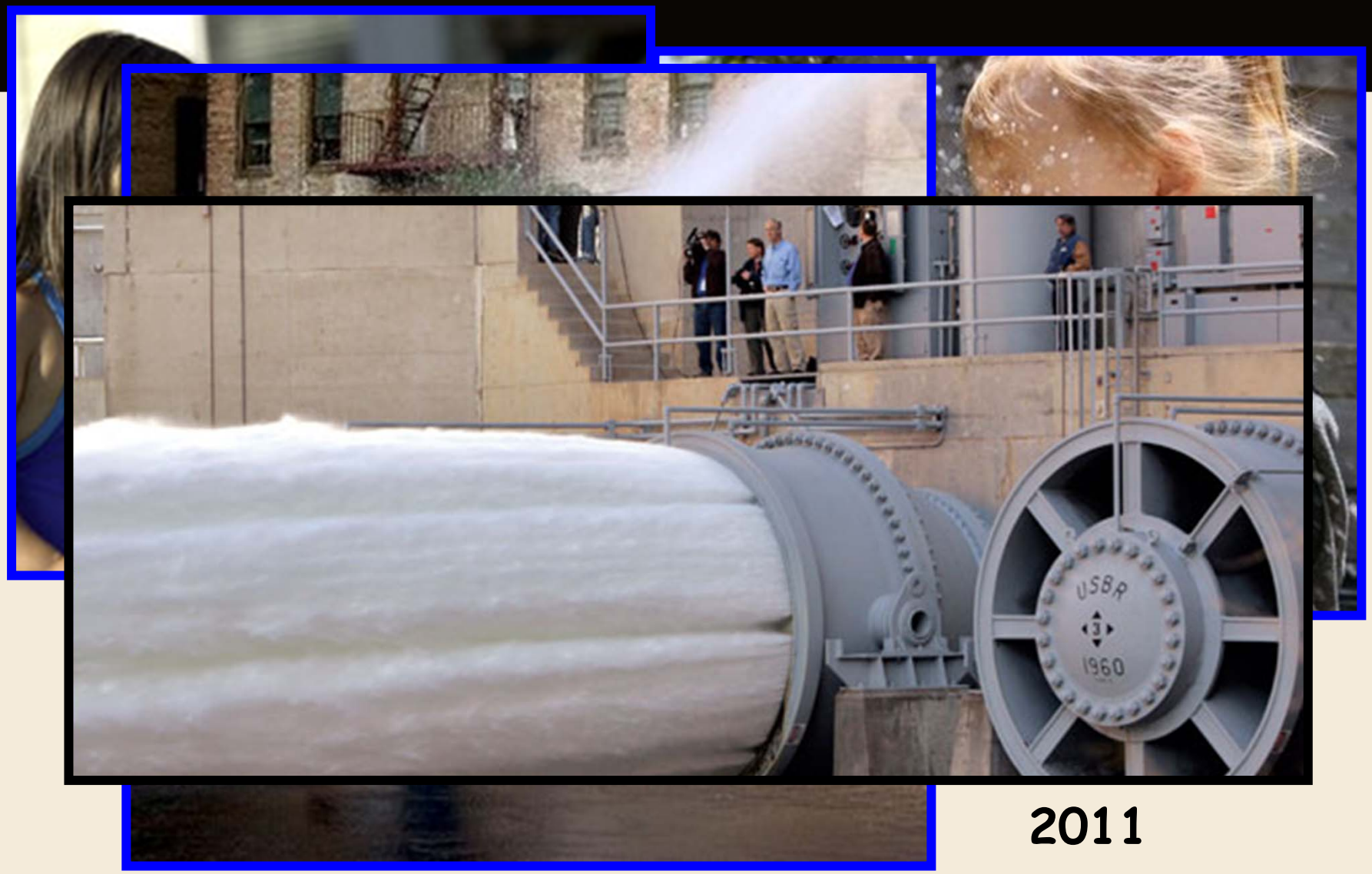


So how much data are we talking about?

- **It depends . . .**
- **Raw data for later analysis = >2<30 Terabytes**
- **Store a file containing each and every DNA letter in your genome, = ~1.5 gigabytes**
- **With very accurate sequence data and access to a high-quality reference genome = ~20 megabytes**



Another way to look at this challenge . . .



2008

2011

With appreciation to Dr. John Weinstein

Medicine is also challenged by the vocabularies we use. . .

Area	Vocabulary
Anatomy	SNOMED CT; NCI Thesaurus
Demographics	HL7 2.4+
Diagnosis	SNOMET CT
Procedures (non lab)	SNOMED CT
Procedures (laboratory)	LOINC
Laboratory Results	SNOMED CT
Oncology-specific terms	caDSR
Immunizations	HL7 2.3.1 + (CVX, MVX)
Clinical Drugs	RxNorm
Drug Classifications	National Drug File Reference Terminology
Manufactured Dosage Form	FDA/CDER
Drug Product	FDA NDC
Nursing	SNOMED CT
Text Based Reports	NL7 CDA 1.0
Clinical Trials	CDISC
Multimedia Reports	None Yet

And what we call our drugs . . .

Klonopin for seizures or Colchicine for gout
Celebrex for pain or Celexa for depression
Reminyl for Alzheimers or Amaryl for diabetes

Zyprexa for schizophrenia or Zyrtec for congestion
Lamictal for Epilepsy or Lamisil for antifungal
Omacor for Cholesterol reduction or Amicar for blood clots

And how we use the vocabularies we have . . .

1. 140/90;
2. Accelerated hypertension;
3. Arteriolar nephrosclerosis;
4. Benign hypertension;
5. Benign Intracranial Hypertension;
6. HBP;
7. Borderline High Blood Pressure;
8. Borderline Hypertension;
9. Chronic hypertension;
10. Essential hypertension;
11. Elevated BP;
12. Familial hypertension;
13. Familial Primary Pulmonary Hypertension;
14. Genetic hypertension;
15. High blood pressure;
16. Hypertension - essential;
17. Hypertension - malignant;
18. Hypertension - renovascular;
19. Hypertensive crisis;
20. Htn;
21. Hyperpiesia, Idiopathic hypertension;
22. Idiopathic pulmonary hypertension;
23. Malignant hypertension;
24. Nephrosclerosis - arteriolar;
25. PPH; Pregnancy-induced hypertension;
26. Primary Obliterative Pulmonary Vascular Disease;
27. Primary Pulmonary Hypertension;
28. Primary pulmonary hypertension (PPH);
29. Primary pulmonary vascular disease;
30. Pulmonary Arterial Hypertension, Secondary;
31. Pulmonary hypertension;
32. Renal hypertension;
33. Secondary pulmonary hypertension;
34. Severe hypertension;
35. Toxemia;
36. Toxemia of pregnancy

And this doesn't identify context: whether measurements were standing, sitting, in the morning, in the afternoon, or in the presence of a "white coat"!

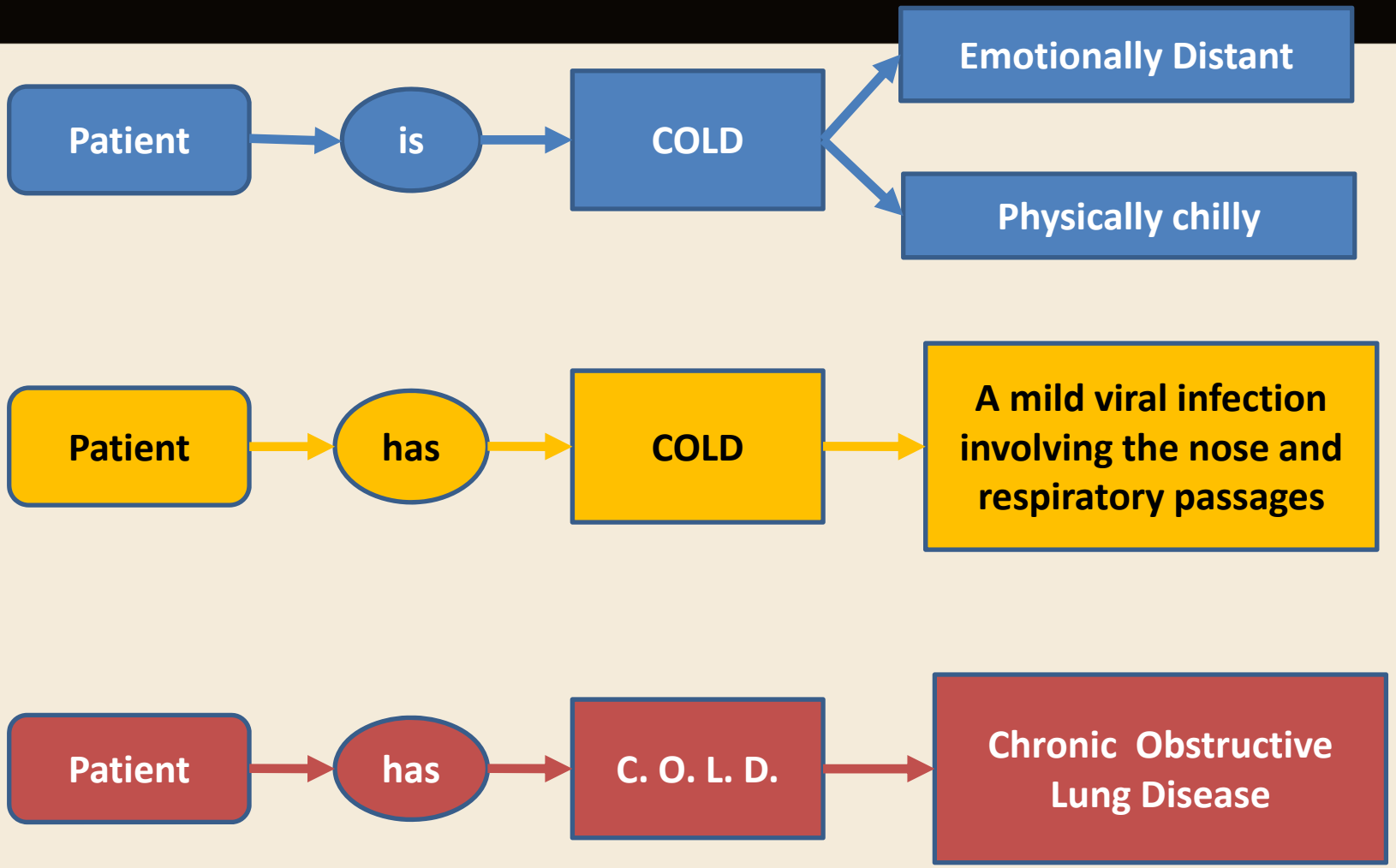
Yet another challenge: the context in which terms are used

1. You are a physician;
2. You are seeing a patient and recording your observations;
3. You are pressed for time;
4. You are writing your observations in the patient's medical record;
5. You write:

C O L D

What does this term mean?
Can a computer understand what you have written?

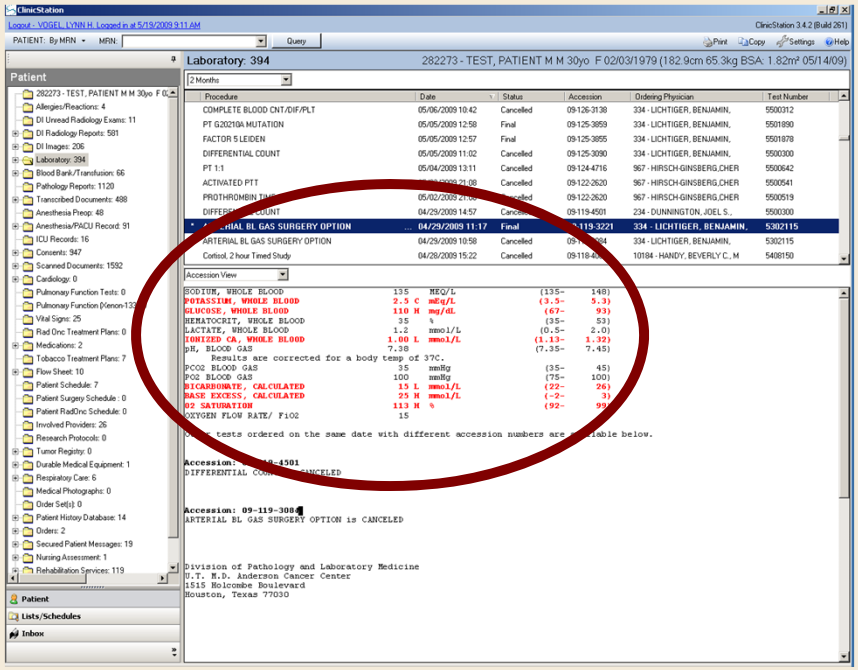
Meaning of COLD?

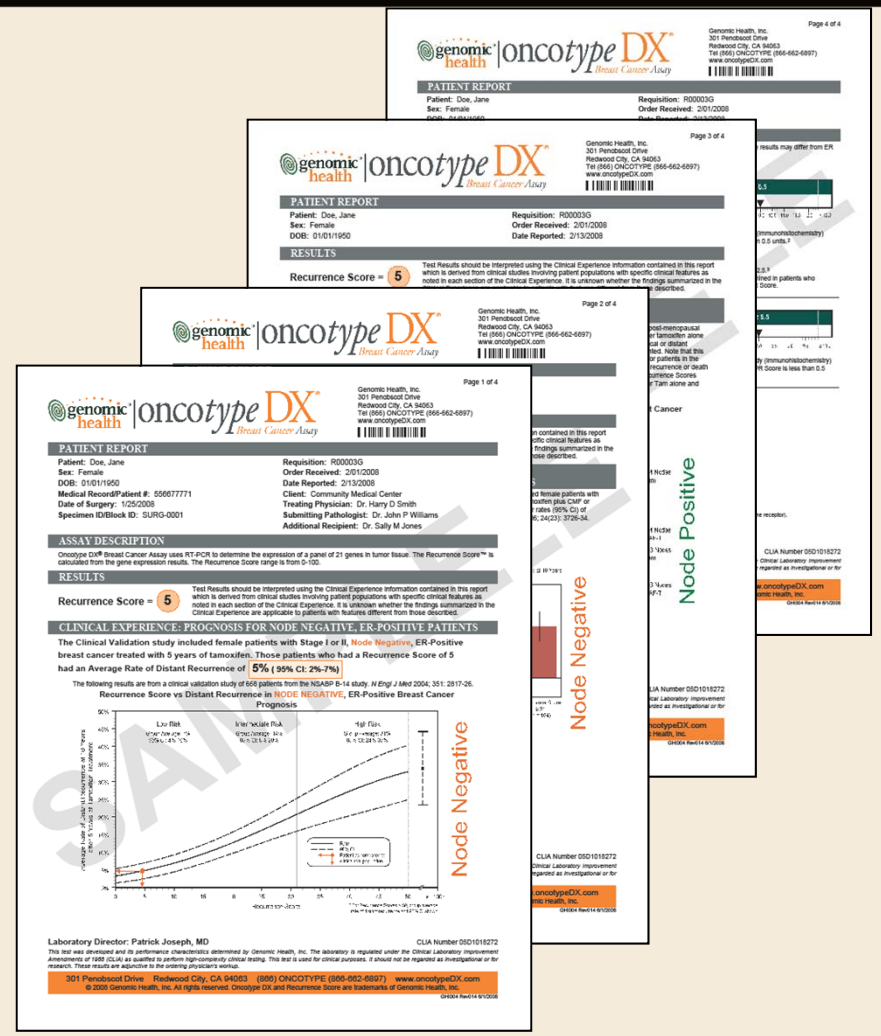


Thus begins the challenge of semantics!!!

What does a “clinical lab result” look like?

Typical Lab Result Presentation





Oncotype DX® Breast Cancer Assay

Discovering "semantics" in the OncotypeDX® report

genomic health | oncotype DX® Breast Cancer Assay

Genomic Health, Inc.
301 Penobscot Drive
Redwood City, CA 94063
Tel (866) ONCOTYPE (866-662-6897)
www.oncotypeDX.com

PATIENT REPORT

Patient: Doe, Jane
Sex: Female
DOB: 01/01/1950
Medical Record/Patient #: 556677771
Date of Surgery: 1/25/2008
Specimen ID/Block ID: SURG-0001

Requisition: R00003G
Order Received: 2/01/2008
Date Reported: 2/13/2008
Client: Community Medical Center
Treating Physician: Dr. Harry D Smith
Submitting Pathologist: Dr. John P Williams
Additional Recipient: Dr. Sally M Jones

ASSAY DESCRIPTION

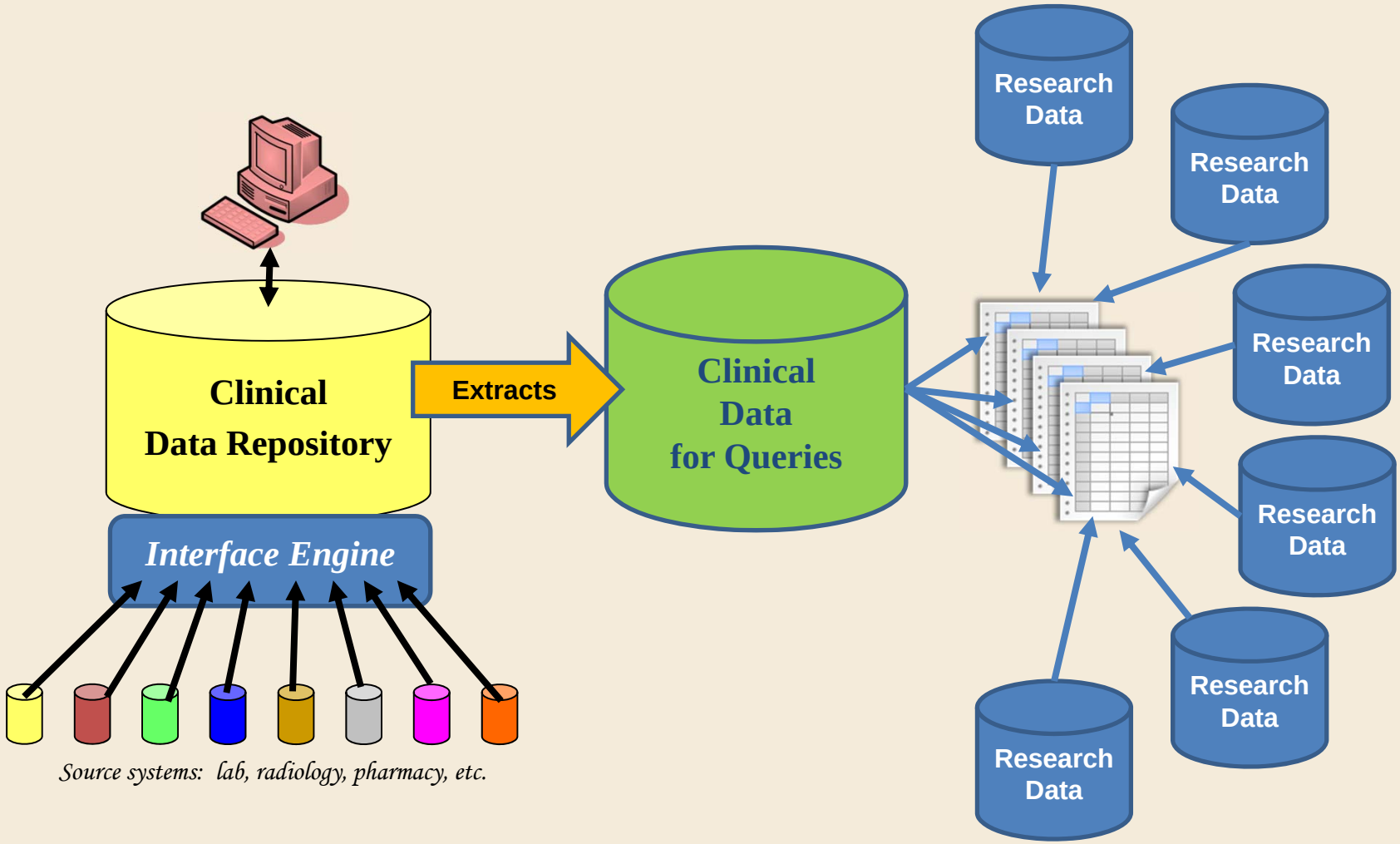
Onco*type* DX® Breast Cancer Assay uses RT-PCR to determine the expression of a panel of 21 genes in tumor tissue. The Recurrence Score™ is calculated from the gene expression results. The Recurrence Score range is from 0-100.

RESULTS

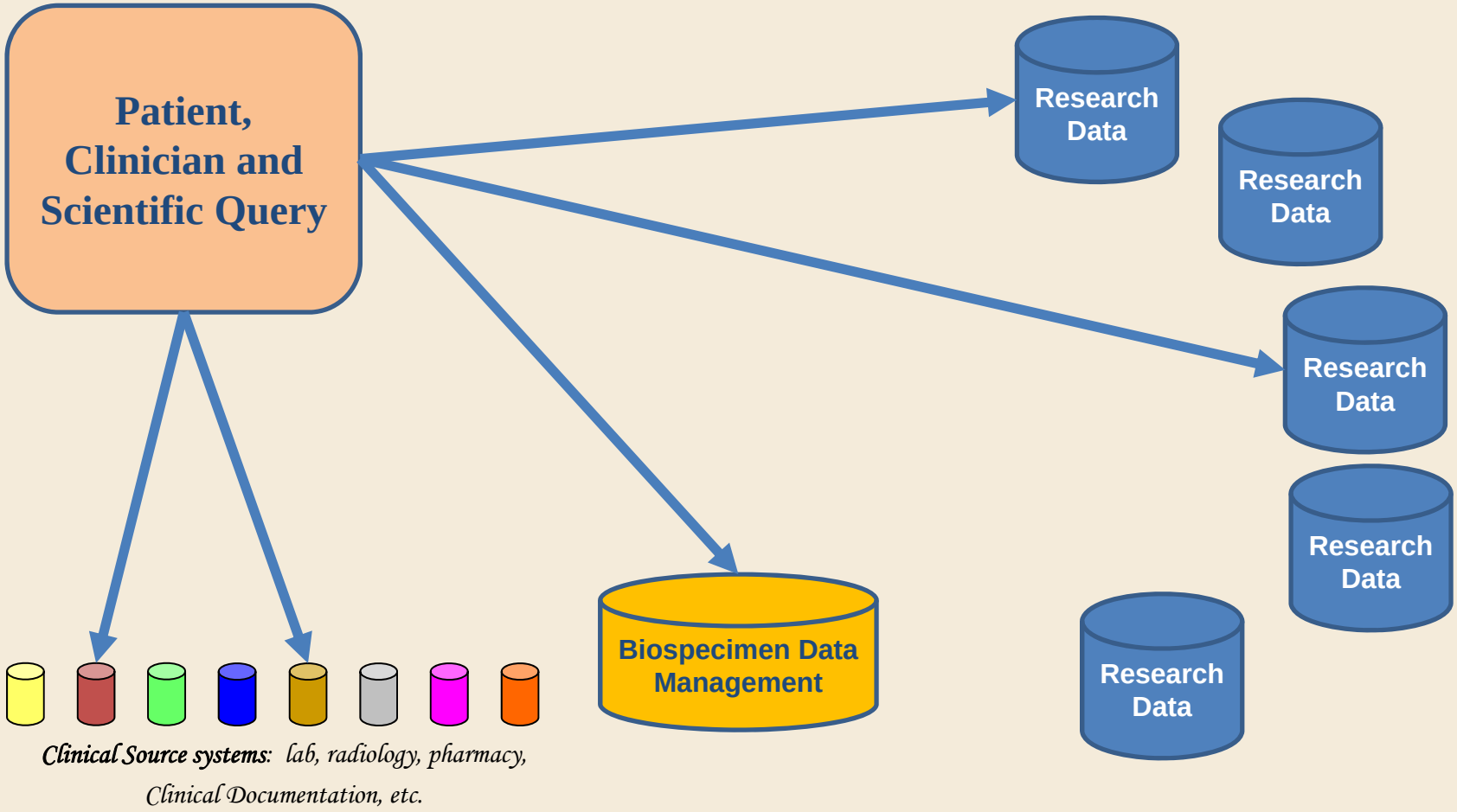
Recurrence Score = 5

Test Results should be interpreted using the Clinical Experience information contained in this report which is derived from clinical studies involving patient populations with specific clinical features as noted in each section of the Clinical Experience. It is unknown whether the findings summarized in the Clinical Experience are applicable to patients with features different from those described.

Historical Model (or Paradigm) for "Interoperability" between Clinical and Research Data



What if you could do this?

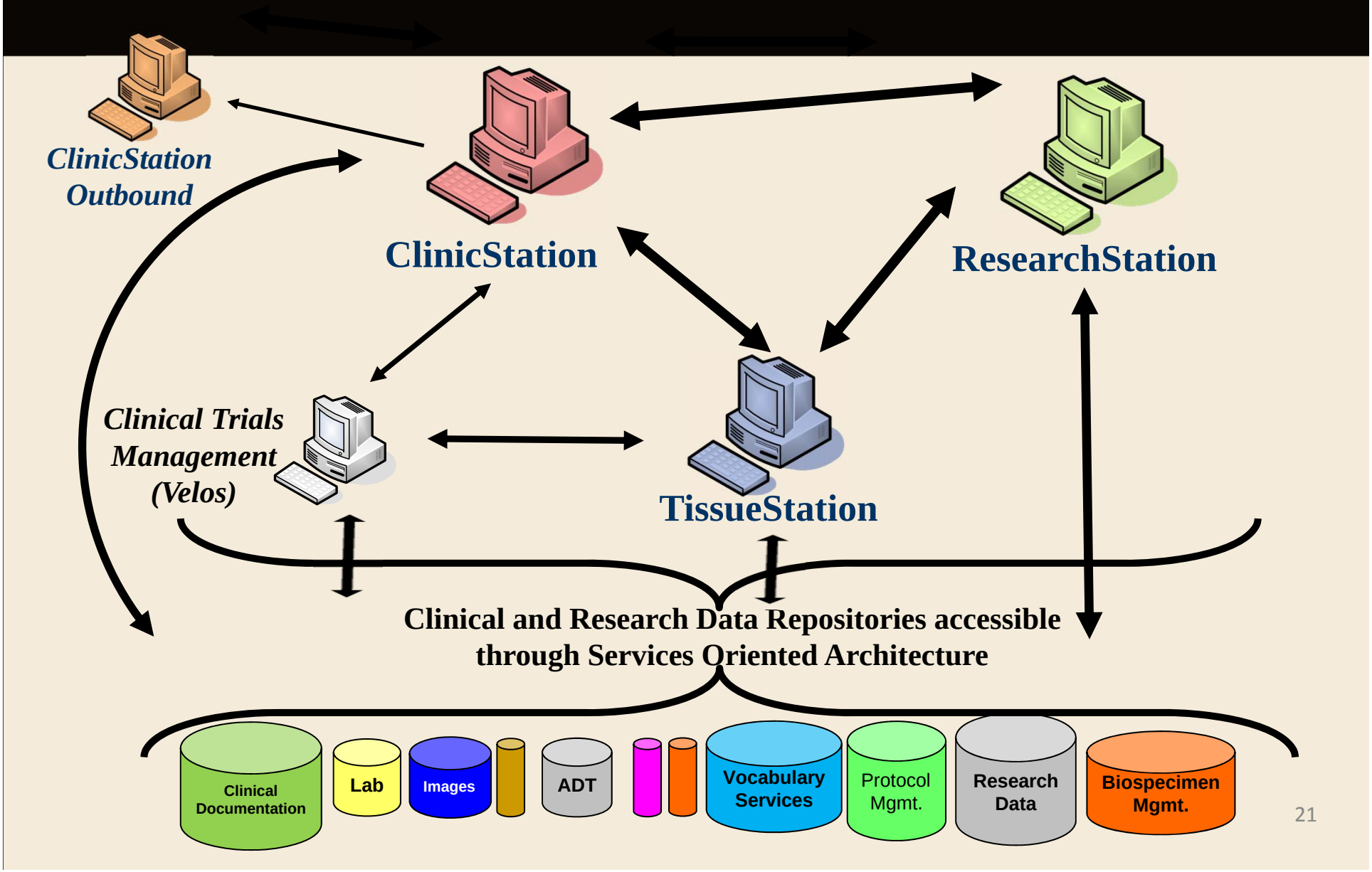


To accomplish this, you need an appropriate architecture.

What is MD Anderson doing to respond to these challenges?

- In-house software development
  - ClinicStation
  - RadStation
  - PathStation
  - TissueStation
  - ResearchStation
- Fully implemented Services Oriented Architecture
- Dramatically increased computational resources
  - Processing
  - Storage

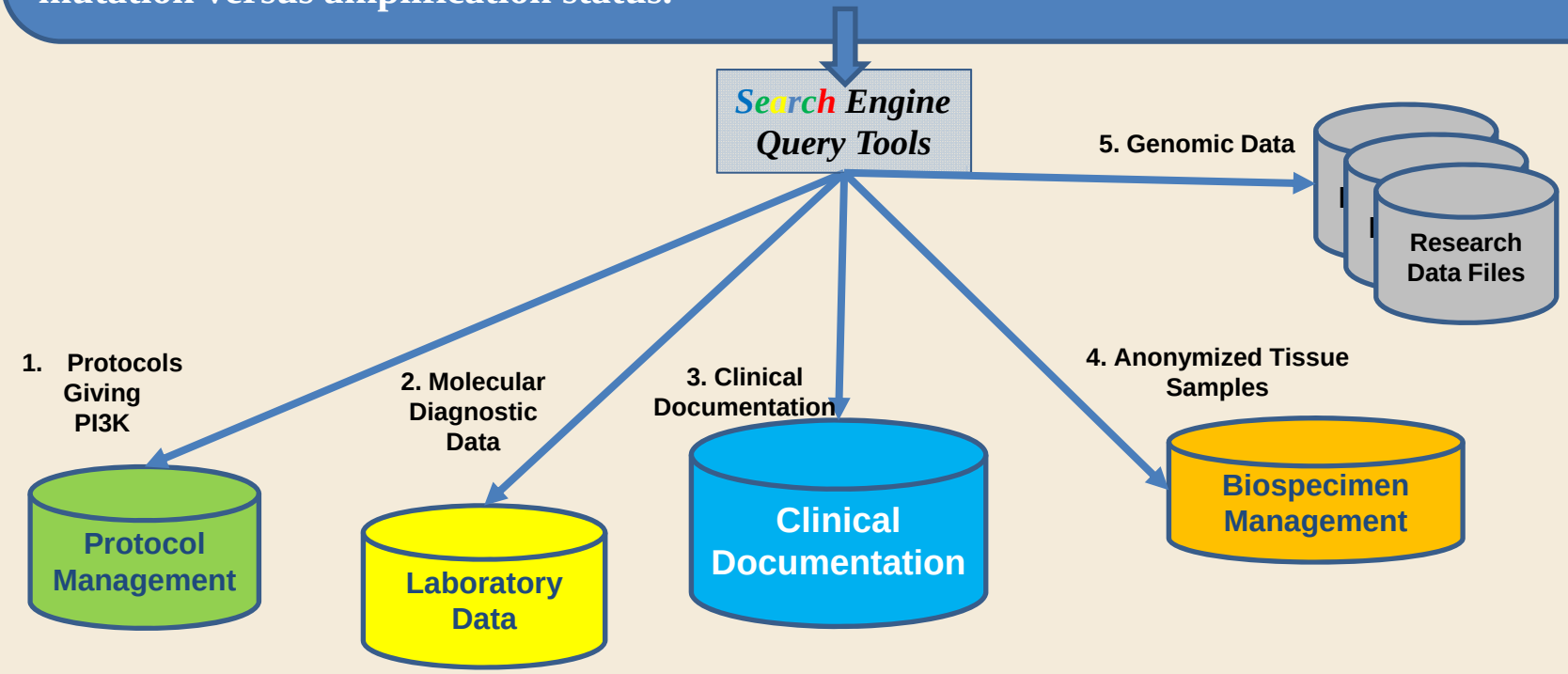
Core Components for MDACC's EHR



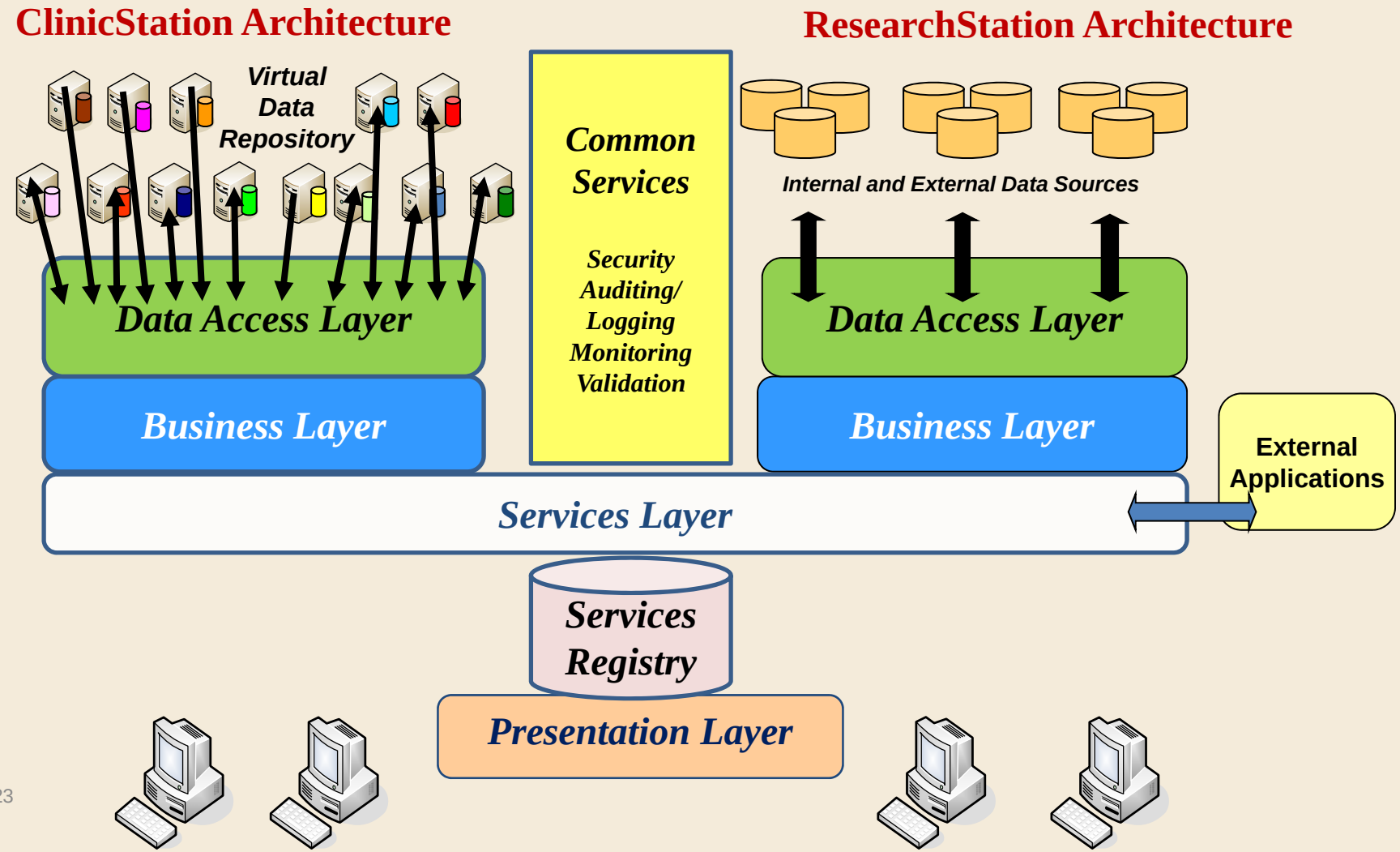
Crossing the Clinical/Research Divide

QUERY: Match the molecular profiles that are common across multiple disease types to determine clinical outcomes depending upon the nature of a specific gene alteration.

Use gene expression and proteomic data to explore classification patterns to predict response to therapy, e.g., with PI3K inhibitors, and determine whether this has any relationship to mutation versus amplification status.



How we bring it all together . . .



What SOA permits us to do

- **Leverage existing assets;**
- **Support multiple forms of integration;**
- **Support specialized implementations;**
- **Focus on building standard components which can be reused;**
- **Permit the incorporation of new computing models as they are needed, e.g., portal-based models, grid computing, etc.**
- **Focus on building systems that will change, not systems that will last!**

Why Healthcare IT is often considered "out of date"

- **Enterprises of all sizes using SOA – 70% . . . *but not in healthcare***
- **77% of this group are "satisfied" with their SOA efforts;**
- **Enterprises of all sizes cutting back on SOA – 1%**
- **Industries leading the SOA investment include utilities, telecom, finance and insurance; SMB public sector and healthcare are lagging.**

Statistics from Forrester Research, Inc., based on 2165 respondents in Q42010, cited in www.zdnet.com on March 24, 2011, accessed on April 4, 2011.

Some closing thoughts . . .

- **With "warehouse-centric, intra-organizational" architectures, are we approaching the classical engineering tradeoff between**
  - **A design that is all encompassing**
  - **A design that can actually work?**
- **With so much investment in decades old architectures and standards, how long will it really take to provide a "meaningfully used", truly interoperable EHR?**
- **Personalized medicine complicates issues of anonymization and de-identification**
- **How do we introduce alternative architectural models into a highly politicized process?**
- ***The true rate limiting factor to interoperability in healthcare is not the applications we use, but the complexity and semantic challenges of our vocabularies, and the architectures within which they exist.***

Thank You!

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center

Making Cancer History®