Slide 1
Biomedical Informatics at the University of Chicago: A Multidisciplinary Approach

Umberto Tachinardi, MD, MS
Assistant-Dean for Academic and Research Informatics
Research Associate - Associate Professor

The University of Chicago
Cancer Research Institute

Slide 2
The Biomedical Research "Cloudscape"

Funding
Regulatory
Clinical Research
Healthcare
Technology
Basic Sciences Research
Organizational

Slide 3
The science machine

Hypothesis ???
Ideas
Collaboration
Funding
Publications
Patents
Science "knowledge"
Personalized Medicine
"Real-time-matrix" Science
Analysis
Data Collection
Research

Clinical
Labs
Molecular
Pubs
Epi

Recycled Data/
Knowledge

Slide 4

UC Research IS structures

UCCRC
DoM
Peds
Health Studies

CBIS
iBi
CI

CaBIG
CTSA

Velos

Slide 5
The CBIS Academic and Research Applications Group

CIO (Eric Yablonka)
Clinical Applications
Governance PM
Infastructure (IT)
Academic & Research

Slide 6
The CBM Cosmos

Velos
Tridon KidsGenes
TraM

Clinical Research Support IS
Grants IS
IRB IS
Contract Mgt
Special Support
(Servers/Storage/Computation)

Education

Researchers Advocacy
Policies
Demo Projects
Storage
Special Projects
BI Research
CTSA

CI
BSDS
iBi
UCMCIT
NSIT
Networking

Back Office
Pritzker IS
Identified Clinical Data
Security
Support
Servers

Slide 7
Current Situation

eSphere
Velos

KidsGenes
TraM
Tridom

Oacis
Other legacy

Epic

IRB
HIPAA

Consented Patients
De-identified Patients
Identified Patients

Research

Clinical

Radiology

Slide 8
Future???

De-Identification (Research)
Ontology Services
NLP Services

Local
PubMed

eSphere
Velos
IRB

Clinical Research Datawarehouse

KidsGenes
TraM
Tridom

SOA

Research
Clinical

Oacis
Other legacy

Epic
Clarity

Radiology
Honest Broker (Clinical)

Slide 9
Clinical Research Data Integration

Epic

Labs
Path
Rad

I
HL7

Velos

II
SOA

SOA/ESB
CBIS

SOA/ESB
iBi

CaGrid Node
CaTissue
CaExchange

III
SOA/Grid

NCI

Slide 10
The BSD/BRI Initiative
Data exchange architecture

UCH
SOA pipeline

BSD
Patient Demographics

Lab data
Alerts (ie SAE)
Clinical Annotation
Images
Administrative messages (CT match)
Maintenance messages
Etc

Slide 11
Infra-Structure demands
Storage
Networking

Slide 12
Graphic: Bar Chart of Growth of Storage by Fiscal Year and Terabytes Storage

Slide 13
Integrated Solution Workflow
Data Acquisition: User entered, LIMS, Collected by Instrumentation, etc.
Researcher

Application Server
PACs
Proteomic
Clinical
Ontology
eVelos
LIMS
Microarray
Sequencing
NLP Data

Databases
Nth

Code Mapping
Microarray DB
Images

Disk Storage
High Speed Disk (SAS)
Moderate Speed Disk (SATA)
Moderate Speed Cache for Archival Disk
(BluRay & Worm Tape)

Slide 14
Integrated Solution Workflow

Proteomic
Clinical
Ontology
Microarray
Sequencing
NLP Data

Databases
Nth

Disk Storage
High Speed Disk
Moderate Speed Disk
Moderate Speed Cache for for Archival Disk

High Speed Networking

Data Center
Network Speeds
1GB, 2GB,4GB
10GB ,20GB

Computational Power
Specialized Servers
Shared Large Memory Model Servers
MPI Cluster Servers

Shared Programs

Slide 15
Integrated Solution for Modern Biomedical Research
1+ Years

Data Acquisition: User entered, LIMS, Collected by Instrumentation, etc.

Application Servers (100s)
PACs
Proteomic
Clinical
Ontology
eVelos
LIMS
Microarray

Sequencing
NLP Data

Databases
Nth

Application Server < 100

SOA ESB (SOAP/Web Services)
Ontology Services
Trusted Broker
NLP Services
Security Broker

GRID Service

External Collaboration

Slide 16
Velos Usage
The University of Chicago is currently using Velos in many of their operational processes.

Daily Use
Clinical Trials Regulatory Management
Clinical Trials Data Management
Listing of Open Protocols on the Intranet
Consent Assess For Clinics

Database
Clinical Trials Review Committee
Scientific and Accrual Monitoring Committee
Regulatory Manager Dashboard
Supervisory Dashboard

Reports
Clinical Data Upload System (CDUS)
Grant Reporting
Study IRB Expiration Report
NCI Summary 4 Report

Slide 17
Velos Current Stats

Patients ~ 26,000

Active Studies (2,849)
Surgery 213; Neurology 31; Medicine 930; Pediatrics 15; Hem/Onc 1670 (1280 - open to accrual, 390 -closed to accrual)
Users ~ 270
Suregery 21, Neurology 7, Medicine 65, Hem/Onc 143, Biostats 4, Investigational Pharmacy 3, OCR 5, CBIS 2

Slide 18
Current use of Velos at UofC

Departments
Health Studies
CBIS
Cancer Research Center
General Clinical Research
Gynecology/Oncology
Hematology/Oncology
Pulmonary/Critical Care
Medicine/Oncology
Office of Shared Research Facilities
Clinical Research Support Office
Neurology
Gynecology/Obstetric
Office of Clinical Research
Psychology
Pediatrics
Pediatrics Oncology
Radiology
Radiology/Oncology
Surgery
Surgery/Oncology
University of Chicago Pharmacy
Medicine
Neurology/Oncology
Cancer Clinical Trials Office

Slide 19
Graphic: Screen Grab of the University of Chicago Velos eResearch web page

Slide 20
Graphic: Screen Grab of the University of Chicago Velos Communication web page

Slide 21
Graphic: Form Samples

Slide 22
Graphic: Form Samples – Adverse Events

Slide 23
Graphic: Form Samples – BP Form

Slide 24
Diagram: Research Nurses/Data Managers and Regulatory managers different roles

Slide 25
Velos
Report Manager
Prots/Cons
CTRC
SAM
Ruby on Rails

Slide 26
Graphic: Web Application – Protocol Listing

Slide 27
Graphic: Web Application – Protocol Listing; continued

Slide 28
Graphic: Web Application – Protocol/Consents

Slide 29
Graphic: Web Application – Regulatory Dashboard Application

Slide 30
Graphic: Web Application – Regulatory Dashboard Application; continued

Slide 31
The Velos "Core Team" at UofC

UCCRC
Marcy List
Amber Burnett
Jia Cheng
Eugene Limb
Swapna Bapat

Arthur Christoph
Consuelo Skosey

OCR
Dionisia Saner
Bethany Martell

Department of Medicine
Don Saner
Nick Shank
Tim Holper

Department of Pediatrics
Eneida Mendonca
Michael Rose

CBIS
Seigmund Johnson
Ryan Crist
Clyde Danganan

& many more….

Slide 32
Translational Data Mart (TraM):
What Is TraM and What Can TraM Do?

Xiaoming Wang
April 22, 2009

Slide 33
What Is TraM and What Does It Do?
TraM is a healthcare data integration warehouse with a web-based application interface: https://tram.uchicago.edu

TraM is a "generic" system that can be used with a broad range of healthcare data and has been customized for all types of cancer data, not just breast cancer

TraM can provide an integration environment for various biomedical domain databases

TraM is scalable and configurable but still a work in progress

Slide 34
Data Unification and Integration Workflows at UCMC

Cancer Registry (IMPAC)

Trial Registry (eVelos)
Family Genetics (Progeny)
Basic Research (various sources)
Chemo Records (BEACON)
Clinics (EPIC)
Medical Survey (various DBs)
Specimen Bank (eSphere)
Image (Stentor)
Image Metadata (radiology labs)

Slide 35
Chart: Dynamic Statistics (not yet synced to source data)

Slide 36
Screen Grab: Support Data Mining (Longitudinal)

Slide 37
Screen Grab: Support Data Mining (cross-sectional)

Slide 38
Screen Grab: Support Data Mining (zone in study)
currently underdevelopment

Slide 39
Screen Grab: Locate Sample for Basic Research

Slide 40
Screen Grab: Support Full Scale Curation
(editing survey questionnaire)

Slide 41
Screen Grab: Multiple Level of Data Privacy Control (per IRB, project, and user role)

Slide 42
Acknowledgements
Computation Institute
Xiaoming Wang
Lili Liu
Greg Cross
Ian Foster

Center for Cancer Genetics
Funmi Olopade
Kisha Hope

Shelly Porcellino
Jim Fackenthal
Others

Dept. of Radiology
Paul Chang
Gilliam Newstead
Sunny Arkani Janson

Other
Breast Cancer SPORE
Cancer Registry (Cassie Simon)
Specimen Bank (Leslie Martin)

Slide 43

Department of Medicine

T.R.I.D.O.M.

Total patients asked: 6943
Patients consented: 5370
Patients refused : 1499
Percent consented : 77.34%
Samples received : 3446

Department	Consented	Samples Received
CARDIOLOGY	155	118
DERMATOLOGY	2	1
EMERGENCY MEDICINE	2	2
ENDOCRINOLOGY	214	166
GASTROENTEROLOGY	1312	932
GENERAL MEDICINE	800	527
HEMATOLOGY/ONCOLOGY	310	225
HOSPITALIST MEDICINE	5	5
INPATIENT	168	64
NEPHROLOGY	52	48
NEUROPHYSIOLOGY	10	4
PULMONARY/CRITICAL CARE	1335	521
RHEUMATOLOGY	975	842
Unknown	30	17

Slide 44

Graphic: Chart based on ethnicity

Slide 45

Tridom

Slide 46

The Chicago BioMedicine Information Services (CBIS)
Service Oriented Architecture (SOA)

Translational Research Support Services:

Patient Data Services
(Laboratory, Pathology, Radiology)
Electronic Honest Broker
DICOM iBroker

Paul J. Chang, M.D., FSIIM
Professor & Vice-Chairman, Radiology Informatics
Medical Director, Pathology Informatics
University of Chicago School of Medicine

Medical Director, Enterprise Imaging
Architect, CBIS SOA Initiative
University of Chicago Medical Center

Slide 47

Typical Legacy-based IT Environment:
"Using humans to integrate workflow"
Patient EHR
RIS
Path
PACS

Slide 48

Integration using a"Single Vendor Solution"
Single Vendor Application "Suite"
EHR Module
RIS Module
Path Module
PACS Module

Slide 49

Using "Edge" Protocols (DICOM & HL7) to Orchestrate Integrated Workflow:
IHE PACS – RIS Integration Workflow Model

Slide 50
Integration Engine Approach Using Edge Protocols (HL-7, DICOM) to Transfer State
Example: PACS – RIS Integration Workflow Model

Slide 51
Screen Grab: Amazon.com homepage

Slide 52
Screen Grab: Netflix.com homepage

Slide 53
Screen Grab: igoogle.com homepage

Slide 54
Enterprise Integration Model: Towards a Service Oriented Architecture

Pathology
HIS
RIS
PACS

Middle ("Business Logic") Layer (Agents, ORBs, Web Services, etc.)

Slide 55
Patient Data Services
(Laboratory, Pathology, Radiology)

Slide 56
Electronic Honest Broker

Slide 57
Process Overview

Slide 58
High-Level Logical Architecture

Slide 59
Technologies Overview

Java – SE6, EE5, JAX-WS 2.1
SOAP Service and Client – JAX-WS generated from WSDL
Database agnostic; Oracle implementation
LDAP agnostic; AD implementation
Servlet Container agnostic; Tomcat implementation

Slide 60
Screen Grab: Login Page

Slide 61
Screen Grab: Electronic Honest Broker

Slide 62
Screen Grab: Electronic Honest Broker: Add Study

Slide 63
Screen Grab: Electronic Honest Broker: New Study Added

Slide 64
Screen Grab: Electronic Honest Broker: Adding a New Patient

Slide 65
Screen Grab: Electronic Honest Broker: Verify Patient

Slide 66
Screen Grab: Viewing PHI

Slide 67
Screen Grab: Honest Broker Study View

Slide 68
DICOM iBroker

Slide 69
DICOM iBroker
A portal and related services that provide end-users simple and powerful HIPAA compliant access to clinical image datasets
Leverages Electronic Honest Broker

Slide 70
Screen Grab: DICOM iBroker

Slide 71
Screen Grab: DICOM iBroker

Slide 72
Screen Grab: DICOM iBroker

Slide 73
Screen Grab: DICOM iBroker

Slide 74
Screen Grab: DICOM iBroker

Slide 75
Screen Grab: DICOM iBroker

Slide 76
Screen Grab: DICOM iBroker

Slide 77
Diagram: First integrated use of SOA

Slide 78
caBIG
caBIG uses caGrid as the foundation which uses various components of Globus Toolkit
WS-RF Framework
Security Framework
Index Service
We lead the Workflow Project of caGrid
We are part of the Architecture workspace

Acknowledgement: Ravi Madduri

Slide 79
UCCRC-caBIG
We have a caGrid node up and running on caBIG
We evaluated caArray and we deployed it in UCCRC
We are working with Velos in an effort to do CDUS submissions using a caGrid Service

Slide 80
caGrid Environment

Slide 81
IRIS
Integrated Research Information Services

Slide 82
Current IT Research Environment
ARG – Academic and Research Group (PSMTrak,
CI – Computational Institute (TRAM)
CRC – Cancer Research Center (Velos)
DID – Development, Integration & Databases (Clinical SOA)
HSD – Health Studies Department (Sandstone)
iBi – Institute for BioInformatics
MAD – Medicine Application Developers (TRIDOM)
NSIT – Networking Services and Information Technologies

Slide 83
Functional Academic & Research Domains

Slide 84
Current ARG Applications & Systems

Slide 85
Current Databases and Feeds

Slide 86
ARG Ecosystem

Slide 87
SOA Layers

Slide 88
Diagram: Consumers

Slide 89
UC|ißi

Slide 90
Genome Science and Biomedical Informatics increase the "experimental space" and "clinical space", respectively one can "search" for biomedical solutions

Genetics, Genome Science, Systems Biology
Phenomics

In "genetic" studies we seek to correlate heritable changes in DNA to Disease Phenotype. In "genomic" studies we seek to correlate DNA changes with other changes in "molecular state" to predict Disease Phenotype. In Biomedical Informatics we seek to characterize "disease" by its component parts with the goal of predicting disease susceptibility by correlation to individual differences in "molecular states".

Slide 91
As of Jan 2008, *GenBank* houses just under 1 Terabyte (1012) of DNA sequence information. Though huge by pre-genome sequencing standards, the data storage challenge is really a consequence of the new "experimental space" that genome science and genome technology enable.

Since 2007 Genbank doubles every 6months
Shorter than Moore's law (computer power doubling every 20 months!)

Slide 32

Automated literature mining to generate a Molecular Interaction Network

We are currently screening 250,000 journal articles…
2.5M reasoning chains
4M statements

Slide 93

Discovering Pathogenic Pathways: why do we need to integrate information?

All currently stored in iBi repository.

New "experimental space"

Slide 94

PhenoGO : A Resource for the Multiscale Integration of Clinical and Biological Data

Lee T. Sam, Eneida A. Mendonça, Carol Friedman,
Yves A. Lussier, M.D. (Corresponding Author)
Director, Dept. of Medicine Center for Biomedical Informatics
Associate Director for Informatics, Cancer Research Center
Co-Director, Clinical Translational Science Award (CTSA) Informatics Core
Associate Professor of Medicine, Biological Science Division
Judith Blake, Jackson Lab

Slide 95

The PhenoGO Encoding Pipeline

Slide 96

Diseases and Disorders

Slide 97

Web Query: Basic

Runs a query analogous to an SQL OR for the search terms

Slide 98

Web Query: Advanced

Runs a query analogous to an SQL AND for the search terms

Slide 99

Conclusions

Substantial additions to a high-quality, wide-raging gene-GO-phenotype resource drawn from the biomedical literature and existing knowledge bases
It's a fantastic resource and I urge everyone to make use of it
http://www.phenogo.org

Slide 100

Acknowledgements
University of Chicago
Tara Borlawsky (now at Ohio State)
Yang Liu
Jianrong Li

Grants
NLM, 1K22 LM008308-01 (YL)
NLM, R01 LM007659-01 (CF)
NLM 1U54 CA121852-01A1 National Center: Multi-Scale Study of Cellular Networks

Slide 101

Discovery of Protein Interaction Networks Shared by Diseases
Lee Sam, Yang Liu, Jianrong Li, Carol Friedman, Yves A. Lussier

Slide 102

PhenoGO.org
Lussier Y, Borlawsky T, Rappaport D, Liu Y, Friedman C. Pac Symp Biocomput. 2006;:64-75.
Database of Phenotypic context for
GO annotations
Gene-phenotype-GO
NLP (BioMedLEE) and computational terminology-derived
3,102 distinct diseases and phenotypes (# human)
32,911 distinct proteins (7,016 human)
532,407 total records (# human)
Precision: 85% [n=120]
Recall: 76% [n=120]

Slide 103

Overall process

Slide 104

What do we have data on?
PhenoGO
7016 proteins (101,711 entires)

Reactome
1140 proteins (126,147 entires)

Shared
176 Proteins

Slide 105

Acknowledgements

University of Chicago
Yang Liu
Jianrong Li
Tara Borlawsky
Yves A. Lussier
Rosie H Xing

Columbia University
Carol Friedman

Synergistic Collaborations:
GO/PATO consortium
Michael Ashburner, Flybase
Judith Blake, MGI
Suzanna Lewis, UB
Joanna Amburger, OMIM

Grants
NLM, 1K22 LM008308-01 (YL)
NLM, R01 LM007659-01 (CF)
NLM 1U54 CA121852-01A1 National Center: Multi-Scale Study of Cellular Networks (YL/Phenotypes & CF/NLP Team Leaders)

Slide 106
Probing genetic overlap among complex human phenotypes

Slide 107
$\Sigma=1.5 \times 10^6$ patient records

Slide 108
Data: ICD9 codes in Columbia University clinical database

Slide 109
Results: Autism

Slide 110
Results: Schizo-phrenia

Slide 111
Diagram: Number of patients

Slide 112
Diagrams: Diseases A, B, C

Slide 113
CIQR ("Seeker"):
Context-Initiated Question and Response
Eneida Mendonca, MD, PhD

Slide 114
History
Linking medical records to information resources
Making retrieval specific to a particular patients
Using statistical and semantic models to identify relevant information in patient's records
Re-ranking retrieval of scientific literature based on the patient's map
Problems: not a good time for retrieval, questions not always related to the part of the record the physician is looking at.

Slide 115
CIQR: Objectives
Capture information needs when they occur, without disrupting workflow
Deploy mobile devices using data and voice input
Translate high-level information needs into search strategies adapted to user needs capabilities of resources
Develop models of search strategies using human search expertise (reference librarians)
Study how librarians classify, clarify and refine questions
Automate using speech processing and natural language parsing
Deliver materials relevant to information needs in an accessible format and in a timely manner
Organize materials retrieved by librarians and transmit them to clinicians for display on mobile devices

Slide 116
Where we are going?

Slide 117
From Bench to Bedside:
Lessons Learned in the University of Chicago CTSA

Re-Engineering Translational Research at The University of Chicago
Clinical and Translational Science Award U54 RR023560

Julian Solway, MD
University of Chicago

Slide 118
Image: Satellite photo of the South Side of Chicago

Slide 119
Univ of Chicago Medical Center Primary Service Area (PSA) – 1.1 M residents

Slide 120

The health status of the 1.1 M residents of the UCMC PSA is extremely poor, as reflected in extraordinarily high rates of common complex diseases and of infant mortality. Furthermore, 10-15% of UHI adult residents are disabled.

It is in this context that the University of Chicago CTSA program seeks to translate research discoveries into real, effective therapies.

Slide 121

A comprehensive approach to the challenges:

Institute for Translational Medicine (ITM) – a new University structure to collect, integrate, and disseminate the intellectual, organizational, and resource infrastructure needed to reduce barriers to translational research interactions among UC investigators. The ITM:
provides new modes of "research navigation" assistance to help UC and Community faculty and trainees identify and contact collaborators
will incentivize multidisciplinary collaborations with pilot and collaborative seed grant funding
support the actual research with a wide spectrum of core and intellectual capabilities

Urban Health Initiative (UHI) – principal goals are:
fully engage our Community in developing clinical and community research agendas
involve Community organizations and practices in the conduct, interpretation and dissemination of that research
accelerate the translation of health knowledge into and out of our Community
most importantly, reduce the marked health disparities in our Community through a combination of service, education, and research initiatives. CTSA Community Translational Science Cluster is the principal research infrastructure component of the UHI.

Slide 122

Three 3 bold steps that will overcome these obstacles.

Training in Clinical and Translational Science
Greatly expanded our training for careers in clinical and translational research, now providing research training and career development opportunities for physician and non-physician scientists, allied health professionals, undergraduates, and high school students, both at UC and in our Community.
Previous barrier to such research training goal as one of "activation energy", in that we have had the essential capability for such training and career development, but failed to realize our full potential due to inadequate mentor incentives, inadequate course offerings, and inadequate recognition of the need for comprehensive training for careers throughout the translational research spectrum.

ITM Training Cluster – broad research training/career development mission, enhanced faculty/mentor incentives, more clearly defined training pathways, and wide buy-in throughout the CTSA community, operating across divisional, departmental, professional school, and university lines to bring together CTSA faculty and students from UC, our allied institutions, and the Community.
Committee on Clinical and Translational Science – new academic entity to advance multidisciplinary PhD/MS training in clinical and translational research.

Slide 123
Org Chart: The University of Chicago CTSA

Slide 124
Org Chart: The University of Chicago ITM

Slide 125
Screen Grab: The University of Chicago ITM intranet

Slide 126
Graphic: Pangea
Permian 225 million years ago

Slide 127
Informaticians and other related contributors at UofC

Slide 128
Contact address:

Umberto Tachinardi
utachina@bsd.uchicago.edu