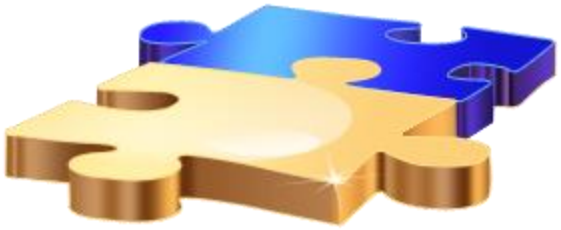
Changing the Model in Pharma and Healthcare – Can We Afford to Wait Any Longer?

Erich A. Gombocz
egombocz@io-informatics.com
IO Informatics, Berkeley, CA, USA

INFORMATICS

Outline



1 Introduction

2 State of the Industry

3 Methodology for Change

4 Use Cases of Adoption

- Pharmaceutical Industry
- Government
- Clinical Decision Support

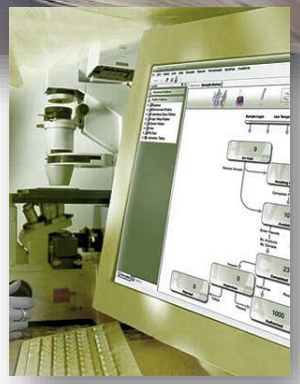
5 Future Outlook

- Historic Models in Life Sciences
- Rise of New Technologies and Machines
- Economic Importance of Data
- Staying Competitive

1 Introduction

INFORMATICS

Historic Models



- Paper records, Spreadsheets
- LIMS (acquisition, workflow, sample tracking, data exchange interfaces, audit, compliance)
- Traditional ETL, partially connected
- Relational Data Warehouses, proprietary schemas
- ELN - specific solutions
- Strict separation of departments, target areas

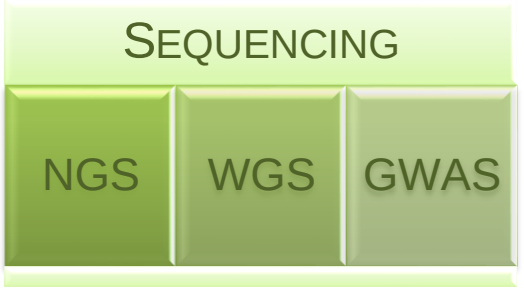
Closed data views make more & more data less & less usable.
Consequence: adverse effects are missed, clinical trial efficiency is low

IO INFORMATICS

Rise of New Technologies

INFORMATICS

Advances in Sequencing



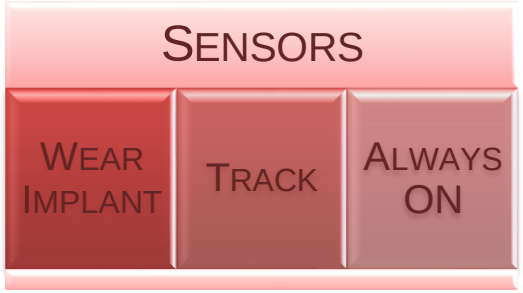
- massive next generation sequencing (NGS), de novo assemblies
- clinical genome sequencing and interpretation by leading organizations (Mayo Clinic, Cancer Genome Atlas)
- exome sequencing in CLIA- and CAP-certified labs for diagnostic use
- RNA sequencing, genome-wide association studies (GWAS)
- gene-based biomarker discovery: personal genomics for precision medicine, and genetic selection of population cohorts for clinical trials
- NGS use to determine causality of variants in genetic diseases
- Numbers of academic and commercial clinical genomics providers rising, more and better genome interpretation software, maturation of testing

Adoption of clinical NGS is not trivial: setting standards, analytical and clinical validity of tests, reimbursement

INFORMATICS

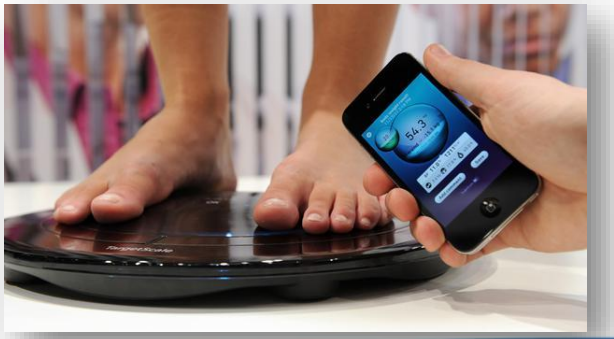
K. Davis: Insight Pharma Reports (2013):
Advances in Clinical Genome Sequencing and Diagnostics

Advances in Sensors



- New sensor technologies plus advances in mobile computing: sensors everywhere, on everything, real-time internet connected
- "The internet of you" - wearable medical computing is a hot commodity
- Devices have great potential to help patients and clinicians monitoring vital signs and symptoms
- Personal health: constantly tracking sensors are changing data collection to a continuous monitoring stream – giving individuals and physicians more detailed data about influence parameters on health or disease state
- Lifestyle, such as exercises, habits and environments have been recorded similarly

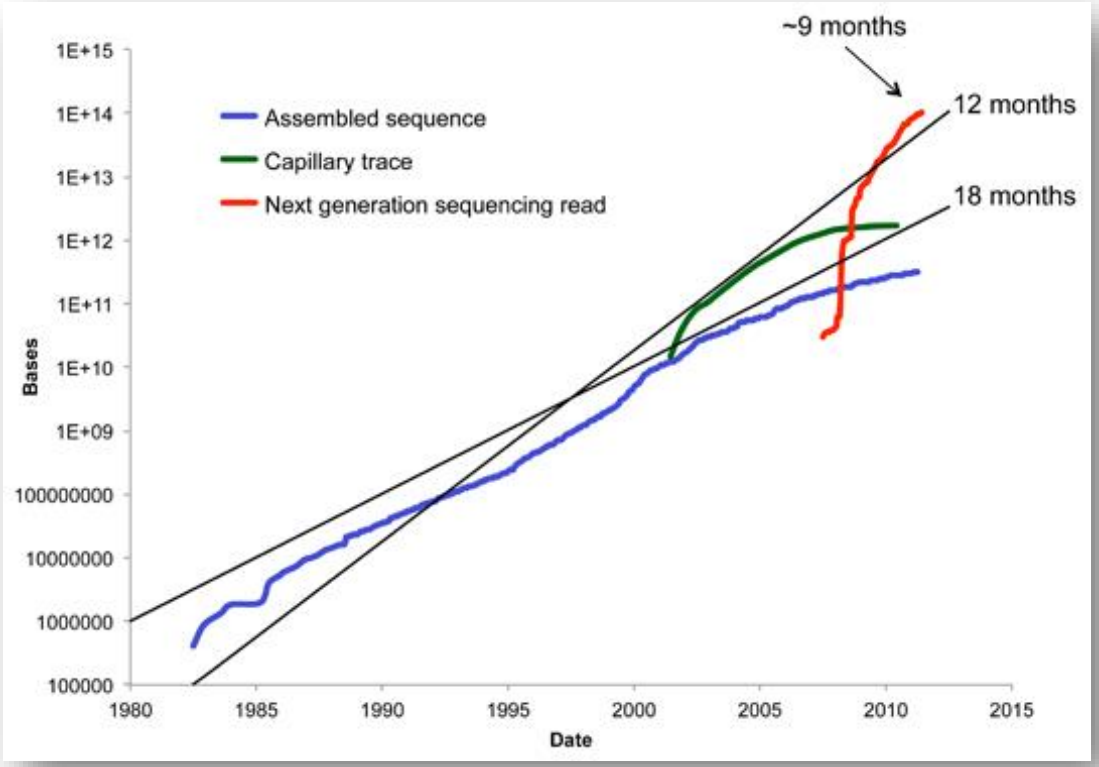
Continuous (more accurate?) data about influence parameters on health state



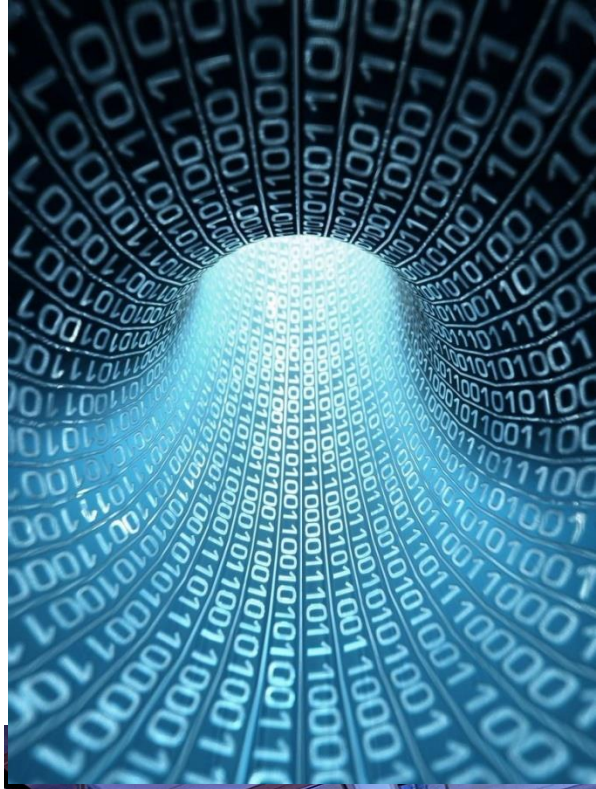
INFORMATICS

Really Big Data

- Just sequencing alone …



Source: G. Cochrane, EMBL (2011)

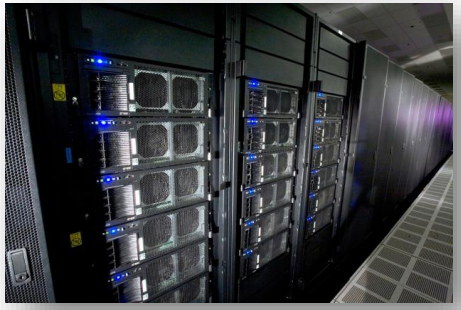


INFORMATICS

Consequences for Computing

- Data size, computational efforts for analysis and high data dynamics require investment in High Performance Computing (HPC)
- Tradeoffs between inexpensive storage and NAS, SAN or Cloud services (CEPH, OpenStack, Amazon)
- Budget-driven compromises: raw data lost for much smaller analyzed data
- Algorithmic transformations are changing, making review for verification impossible.
- massive parallel computing and distributed clusters needed for analysis
- management of 'big data' has become a complex, expensive and demanding task at scales beyond forecast expectations.

Computing is a bottleneck in practical data use, and interoperability, provenance and versioning are equally important



INFORMATICS

Economic Importance of Data

- Infonomics big data analytics as way to turn data into money
- access to and use of vast interconnected data is power, information as asset
- Expanding social networks drive individuals to take care of their needs for better prognosis and treatment
- scarcity of public availability of medical data motivates patients to make their data publicly accessible
- data access = power: implications in shift from revenue and margins driven industrial models towards customer-centric, health outcomes for patients motivated strategies
- consumer influence is driver for new healthcare systems

Moving to evidence-based, outcomes-focused, behavior-driven life sciences affects insurers, payers, providers, drug discovery and drug development

INFORMATICS

Staying Competitive

- Low ROI on research and development has provided little incentives to innovation or change
- The unwillingness to share data even within tightly controlled consortia has reduced the competitiveness of the industry
- Pharma and Healthcare are realizing benefit from collaborations with complimenting data
- Meaningful collaboration and crowd-sourced analysis require interoperability between commercial and academic entities
- Interoperability makes big data BIG OPEN DATA
- Openness assures rapid progression in scientific discovery and provides a solid foundation for Pharma and healthcare to stay competitive

Drivers for collaborative data sharing are cost and efficiency which will dominate the future of data-driven life sciences

INFORMATICS

- Data Generation vs. Knowledge Gain
- Traditional Data Mining
- Cost of Research vs. Outcomes
- Data Ownership:
  Closed Data vs. Patient-Shared Access

2 State of the Industry

INFORMATICS

Data Generation vs. Knowledge

- Massive data - 1 NGS run ~900 GB raw data, 1 machine = 10 TB/day
- Complex testing costs went down, but analysis lags behind:
  - "$1000.- genome at $1 Mio interpretation." – only a small fraction of information is used to build knowledge
- Required proficiency of range of experts and specialists is not easy to establish
- Despite genome annotations and interpretation of causative variants, clinical verification and ramifications for physician and patient require still large efforts
- Widespread sequencing for diagnosis and quality of life improvement is still distant (same in proteomics, transcriptomics and microbiome)
- Bioethics pros and cons of WGS of every newborn child need to be sorted out
- Socio-economic disparities in healthcare must be addressed
- Standardization of analytical methods, algorithms and QC requirements are mandatory for data use across laboratories

Rapid data generation from automated low cost high throughput sample processing does not provide the achievable knowledge gain

Traditional Data Mining

- Relational data warehouses and object data bases require upfront data models (schemas). excellent for final datasets and highly performant on optimized queries, solutions are demanding in support due to their rigid and static structure
- However, many mining tools are available as RDBS technology has been established long ago and is embraced by big players in information technology
- At the scale of growth, additional data warehouses with traditional data mining approaches cannot keep up with today's dynamic information requirements.
- In Life Sciences, dynamic, agile solutions are required to keep pace with changing data types, formats, instrumentation and analytical requirements. questions to ask and use cases are moving targets, so any inflexible solution limits its applicability
- traversing data, inferring from other data and searching complex pattern across all resources to find clues what questions you can answer is essential in life sciences

As proprietary static schemas prevent cross-resource queries and any change is labor intense, the limitations of traditional approaches are obvious

INFORMATICS

Cost of Research vs. Outcomes

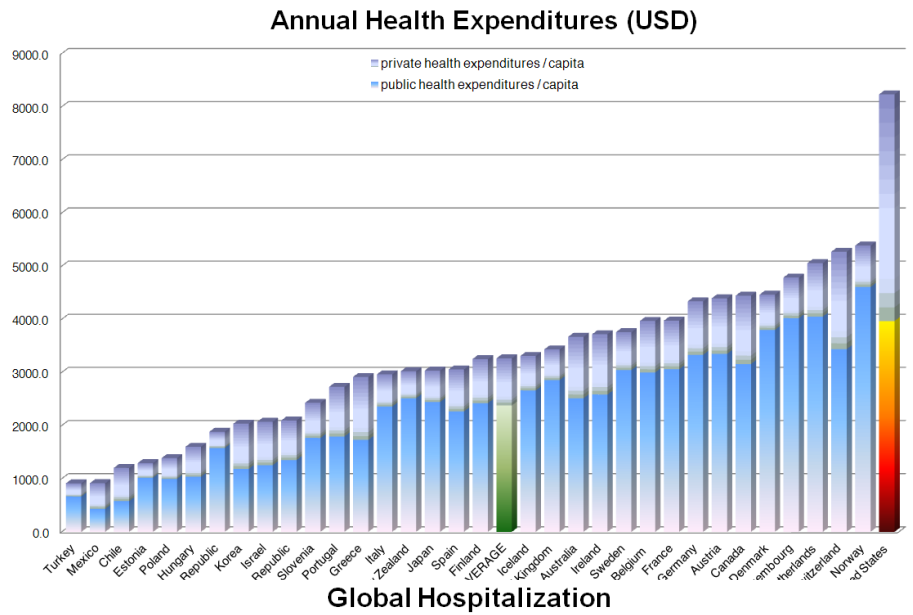
- Staggering costs of new drugs – stunning, but well known within the industry
  - Estimated average cost of bringing a new drug to market is $1.3 billion
  - A drug developed by a major pharmaceutical company costs at least $4 billion in R&D
- Comparing healthcare cost and quality
  - The US spends $8,233 per person/year (2.5 x more than most of developed nations) and uses 17.6% of GDP for healthcare
  - The US had 2.4 practicing physicians per 1,000 people (OECD average 3.1)
  - The US had 2.6 hospital beds per 1000 people (OECD average of 3.4)
  - The US Life expectancy increased by 9 years between 1960 and 2010 (Japan's by 15 years, OECD average by 11 years)
- Per patient clinical trial costs have risen on average by 70 percent across all development phases since 2008

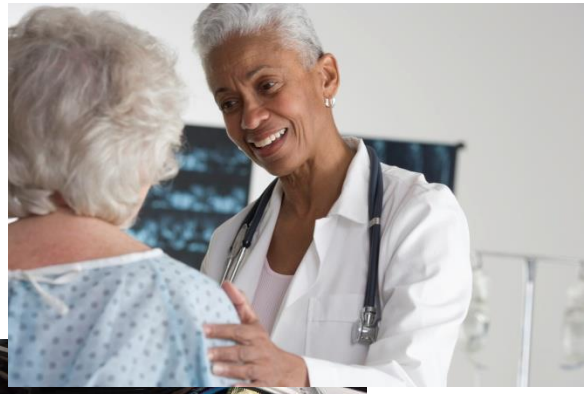
The numbers are clear indicators that the cost vs. outcome ratio needs improvement and the current models require adjustments

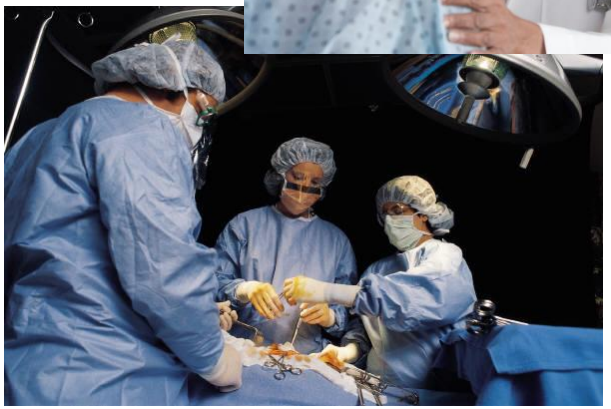
INFORMATICS

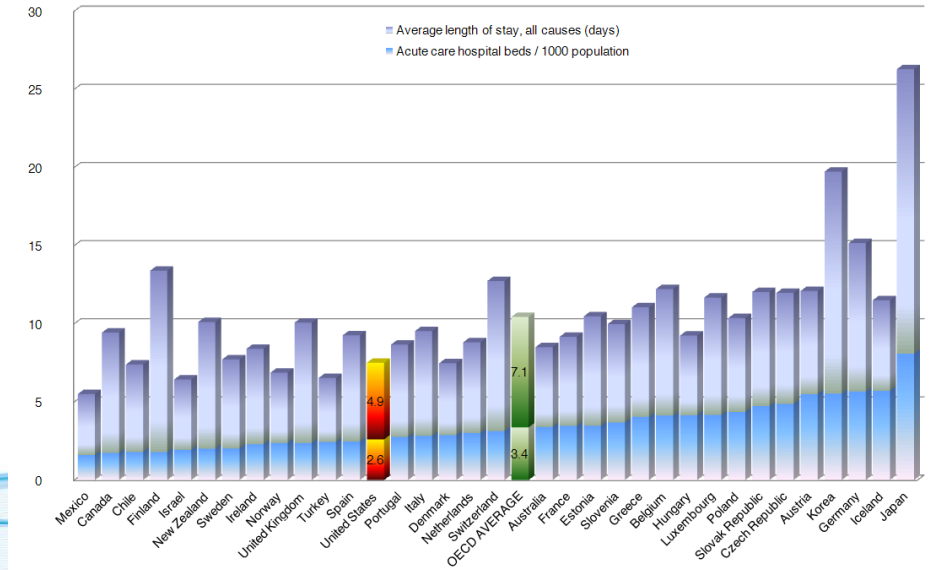
Research Cost vs. Outcomes

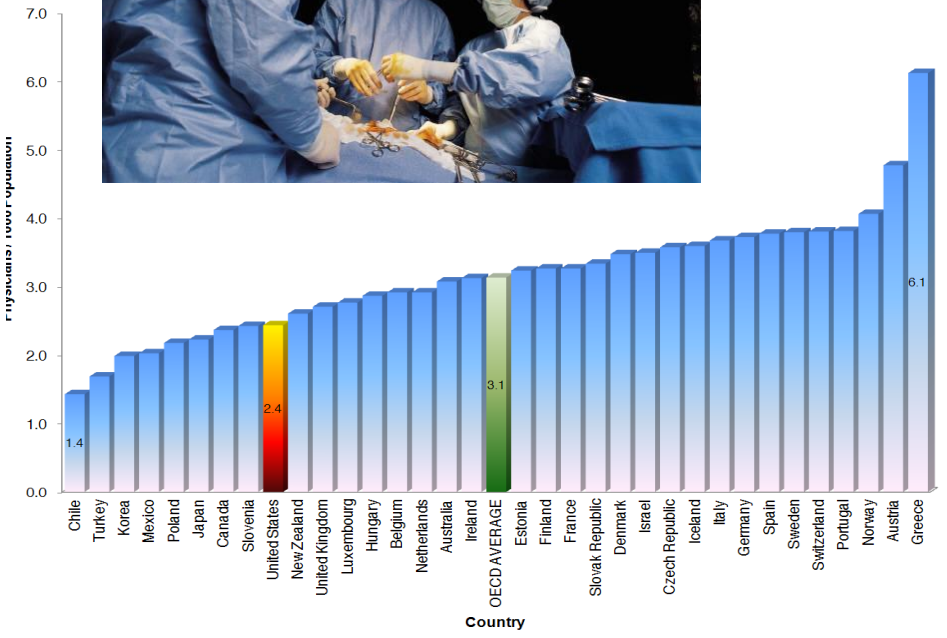
- Example: Healthcare



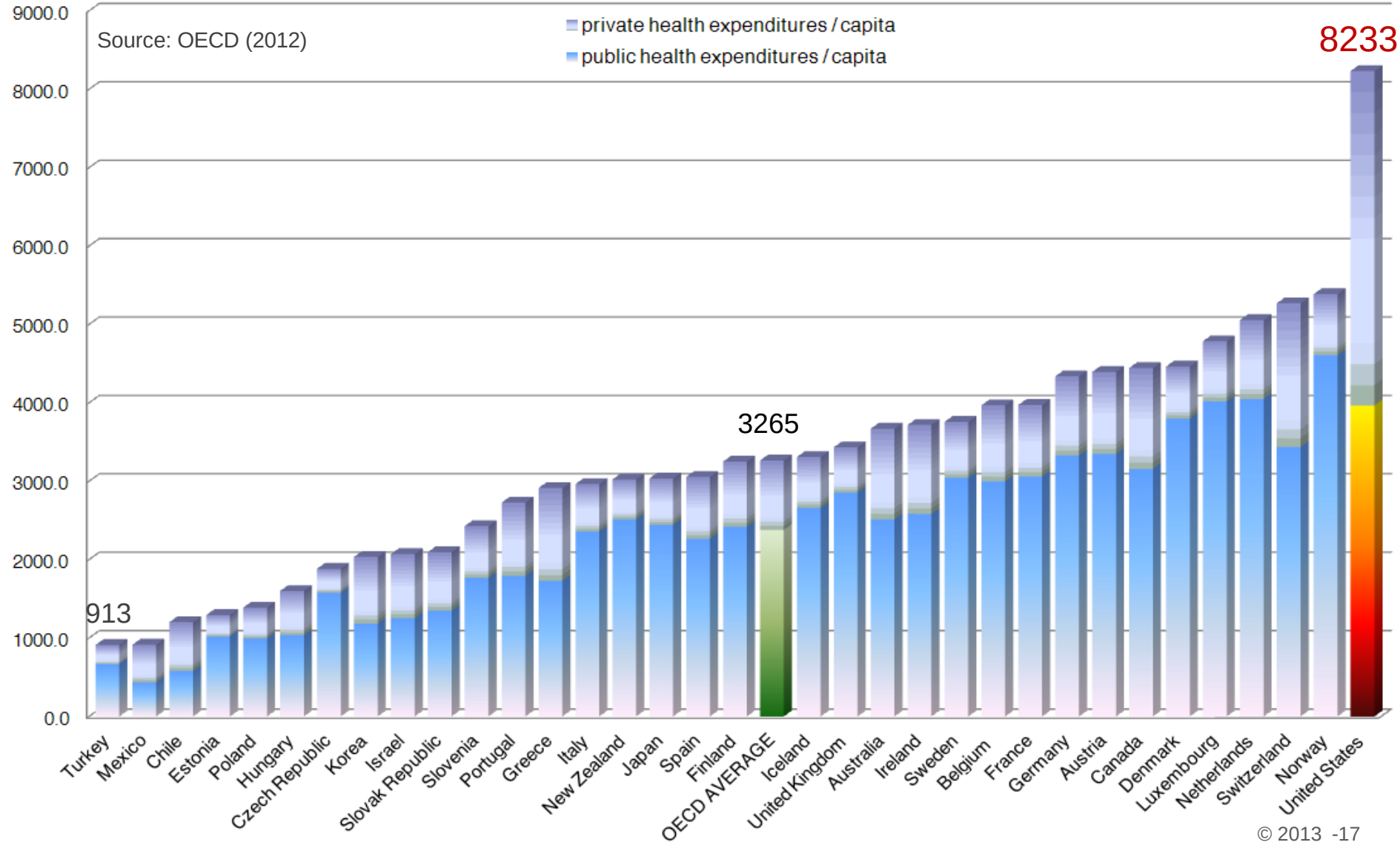




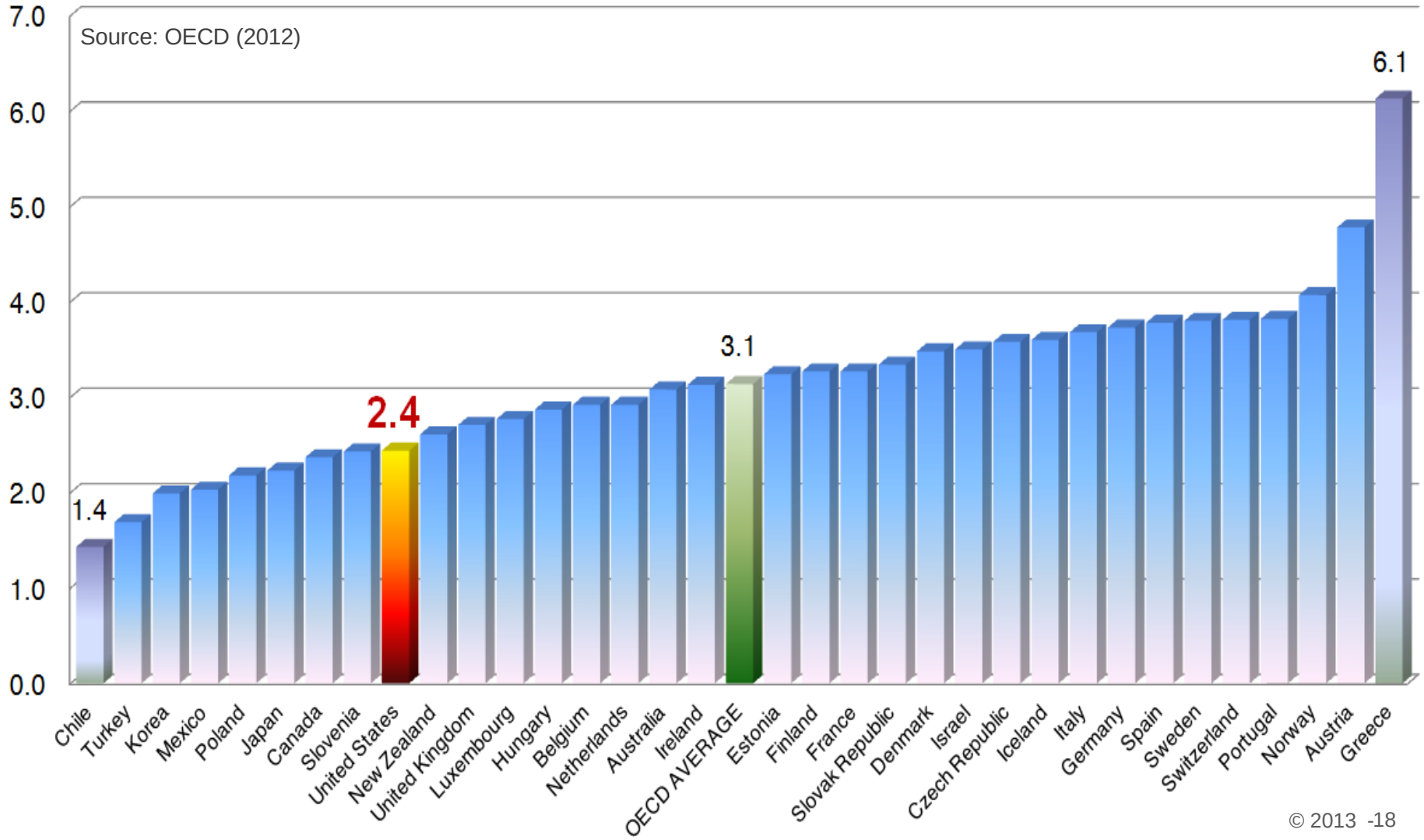




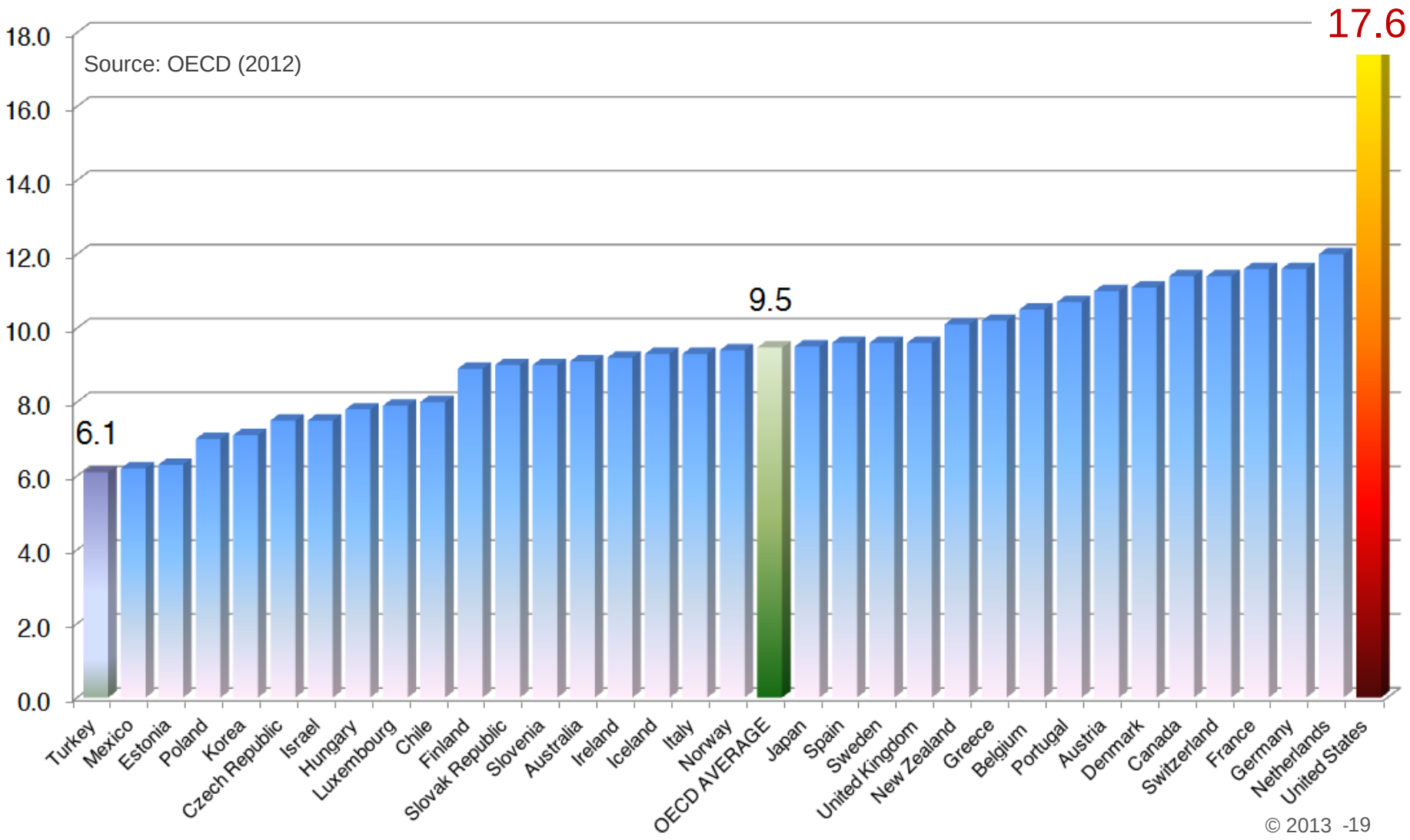
HEALTH EXPENDITURES (PER CAPITA, US $ / YEAR)



HEALTHCARE QUALITY (PHYSICIANS / 1000 POP.)



HEALTH EXPENDITURES (IN %GDP)

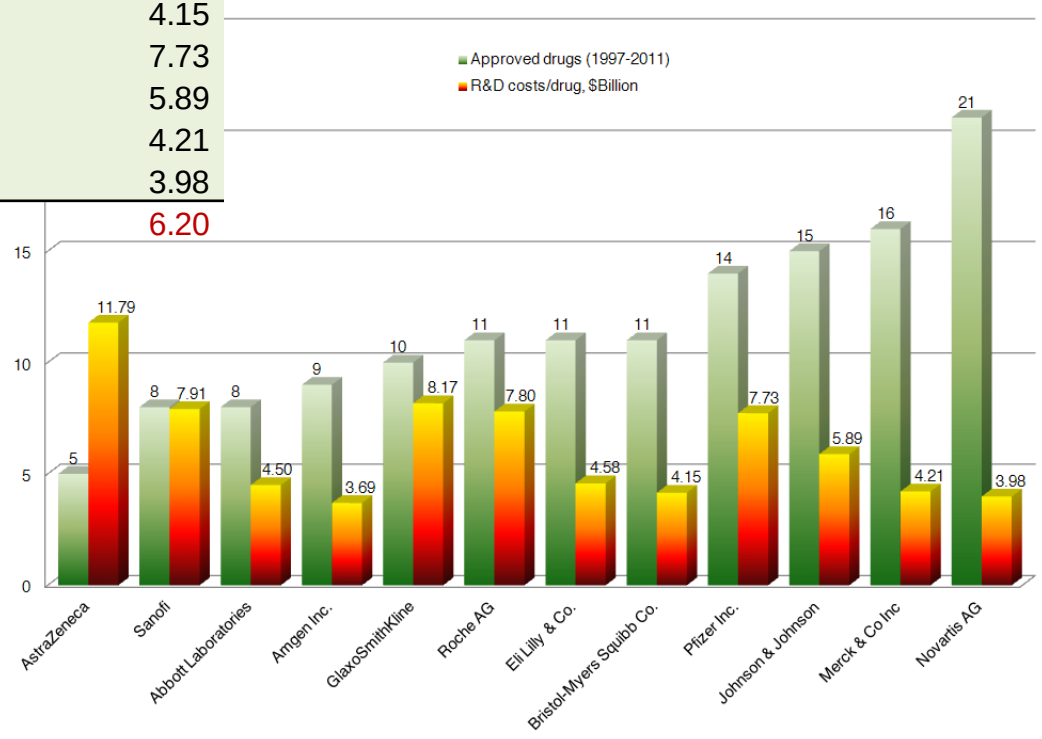


Research Cost vs. Outcomes

- Example: Pharmaceutical Industry

Company	Approved drugs (1997-2011)	R&D costs/drug ($Billion)
AstraZeneca	5	11.79
Sanofi	8	7.91
Abbott Laboratories	8	4.50
Amgen Inc.	9	3.69
GlaxoSmithKline	10	8.17
Roche AG	11	7.80
Eli Lilly & Co.	11	4.58
Bristol-Myers Squibb Co.	11	4.15
Pfizer Inc.	14	7.73
Johnson & Johnson	15	5.89
Merck & Co Inc	16	4.21
Novartis AG	21	3.98
Average	11.58	6.20

Source: Forbes (2012)



INFORMATICS

Change in Data Ownership

- Social media has arrived in healthcare ! Closed data vs. patient-shared is a shift in mindsets
- As patients are sharing publicly their own data, no restrictions on scientific use apply
- Government incentives to actively share information (blue button initiative)
- Consequences: shifts of data ownership from hospital and providers to patient: crowdsourcing for integrated research
- Extension to life style data - more and more personal health devices in use, strong movement towards prevention and wellness

Sharing real-time monitored vital functions among individuals and physicians indicate closed data lose ground against patient-shared open access

Blue Button

a simple concept
with transformative potential

To improve healthcare, we must empower patients. Join the effort to give all Americans the information they need to become active participants in their own care.

IO INFORMATICS

- Semantic Approach to Data Integration
- Linked Life Data, Linked Open Data
- Complexity and Change Require Dynamic, Adaptable Models
- Understanding Biology: Shift To Interoperable, Integral Systems
- Towards New Models: Pharma 3.0, Healthcare 3.0

3 Methodology for Change

INFORMATICS

Semantic Approach to Integration

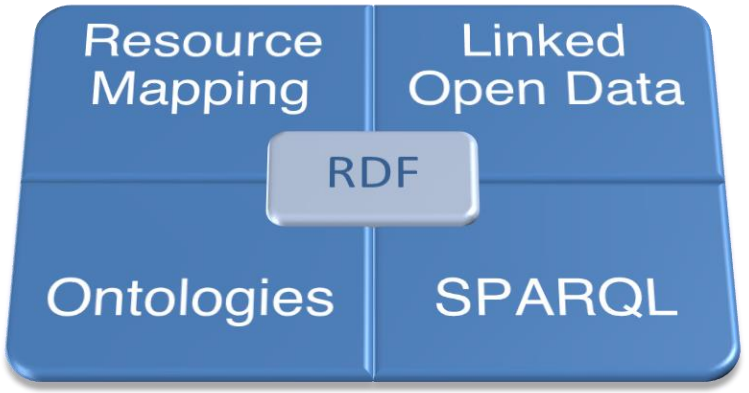
- Resource description framework (RDF)-based integration (W3C standard): agile solutions, rapid and efficient data integration, built for interoperability
- An RDF-based data model (a relationship graph) is more naturally suited to knowledge representation
- semantic data is much easier to connect, to visualize and extend
- RDF and its web ontology language, OWL are excelling in representing data with changing needs, ability to reuse, repurpose, easy to adopt and to maintain
- It offers network exploration, visual query, faceted browsing, machine inference, and pattern recognition in a globally standardized framework
- Recent advances:
  - development of public formal ontologies, harmonization across concepts and vocabularies, provenance and versioning

Meaningful, not arbitrary schema-based connections provide capabilities for interoperability, inference, reasoning and pattern-based queries

Semantic Data Integration

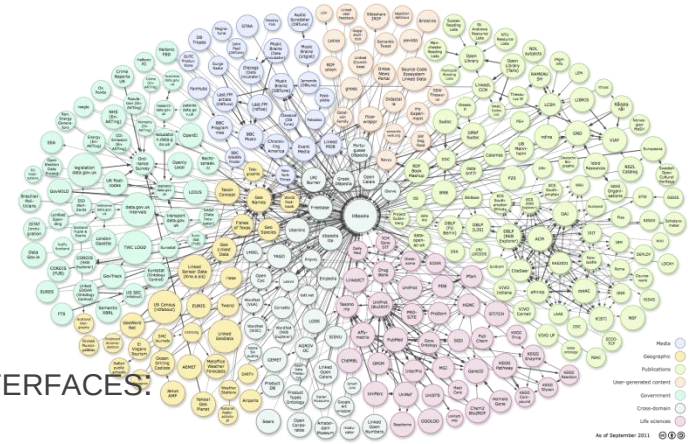
Components

URIs, Resource Description Framework (RDF), Serialization (RDFa, RDF/XML, N3, Turtle, …), SPARQL



- The built-in interoperability of full compliant RDF resources is a MUST for today's life sciences needs to utilize arrays of publicly available linked data
- Joint industry/academia efforts: Translational Medical Ontology and Knowledgebase (TMO), connex between medical informatics and bioinformatics in knowledge building in the clinic
- Harmonization between concepts, vocabularies and clinical standards: Biomedical research integrated domain group (BRIDG - CDISK, HL7 RCRIM, NCI, FDA) – creating a shared view of dynamic and static semantics for the domain of protocol-driven research and associated regulatory artifacts

LLD, LOD Motivation



- **Shareability**

  Make existing data more openly accessible (standard interfaces: SPARQL, resolvable URIs)

- **Integration**

  Create and maintain a list of links between datasets for easy query across them

- **Normalization**

  Match and index an existing RDF data set using common reference ontologies (CROs), so that the dataset can be queried using ontological terms

- **Discoverability**

  Enable a researcher to discover what is available in the Semantic Web about a set of proteins, genes or chemical components, (clinical trials, patients, …) as published results, raw data, tissue libraries, …

- **Federation**

  Integrate information from distributed data sources for queries using SPARQL.

Source: W3C

LLD, LOD Consequences

- The significant increase in the quality of Linked Data (LLD, LOD) brings promising add-ons to qualify experimental findings early on through enrichment with external resources
- But interoperability and provenance remain still impediments for broader applicability as well as changes in licensing for previously 'open' public resources
- Government funding for public linked open data resources is unsure due to global budget restraints - we need to establish new business models between data provider and consumer to warrant continuous availability of such resources; either through private/academic/government partnerships or new concepts based on resource value for organizations

As the socio-economic benefits of maintaining public resources by far outweigh contributions towards their sustainability, such models will benefit all participants greatly.

INFORMATICS

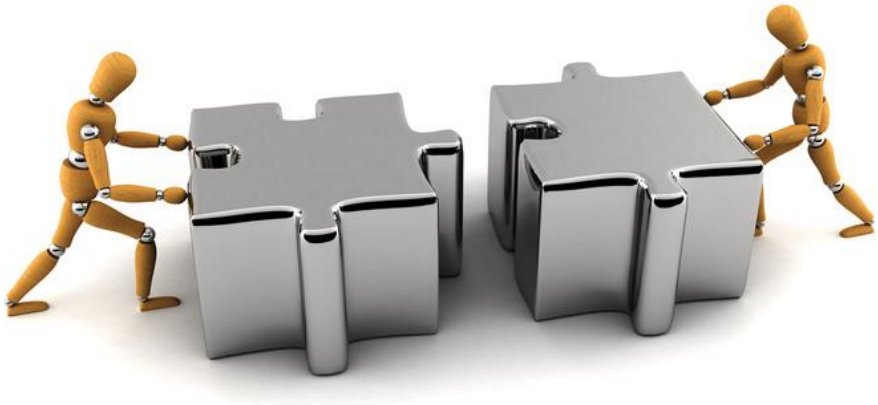
Complexity and Change Require Dynamic, Adaptable Models

- Meaningful use - ambitious incentive to move hospitals and private practices to electronic health records (EHR) is 3 Years old.
- Regulators adopt HIPAA for the 21st century, and similar efforts are underway for telemedicine
- The complexity of functional biology requires network analysis on data in their context - Semantic integration is a logical consequence
- Healthcare finally starts benefitting from big data, cloud services and other disruptive technologies that dramatically changed other vertical industries
- Despite over $10 Billion in healthcare incentives, costs have tripled within the last decade
- We can no longer afford non-interoperable EHR systems and customer-disconnected CRO's - openEHR aims to improve this

Comparative effectiveness starts driving reimbursement …

Dynamic, Adaptable is New Norm

- Interoperable!
  - Common Standard Integration
  - Shareability
  - Discoverability
  - Federation

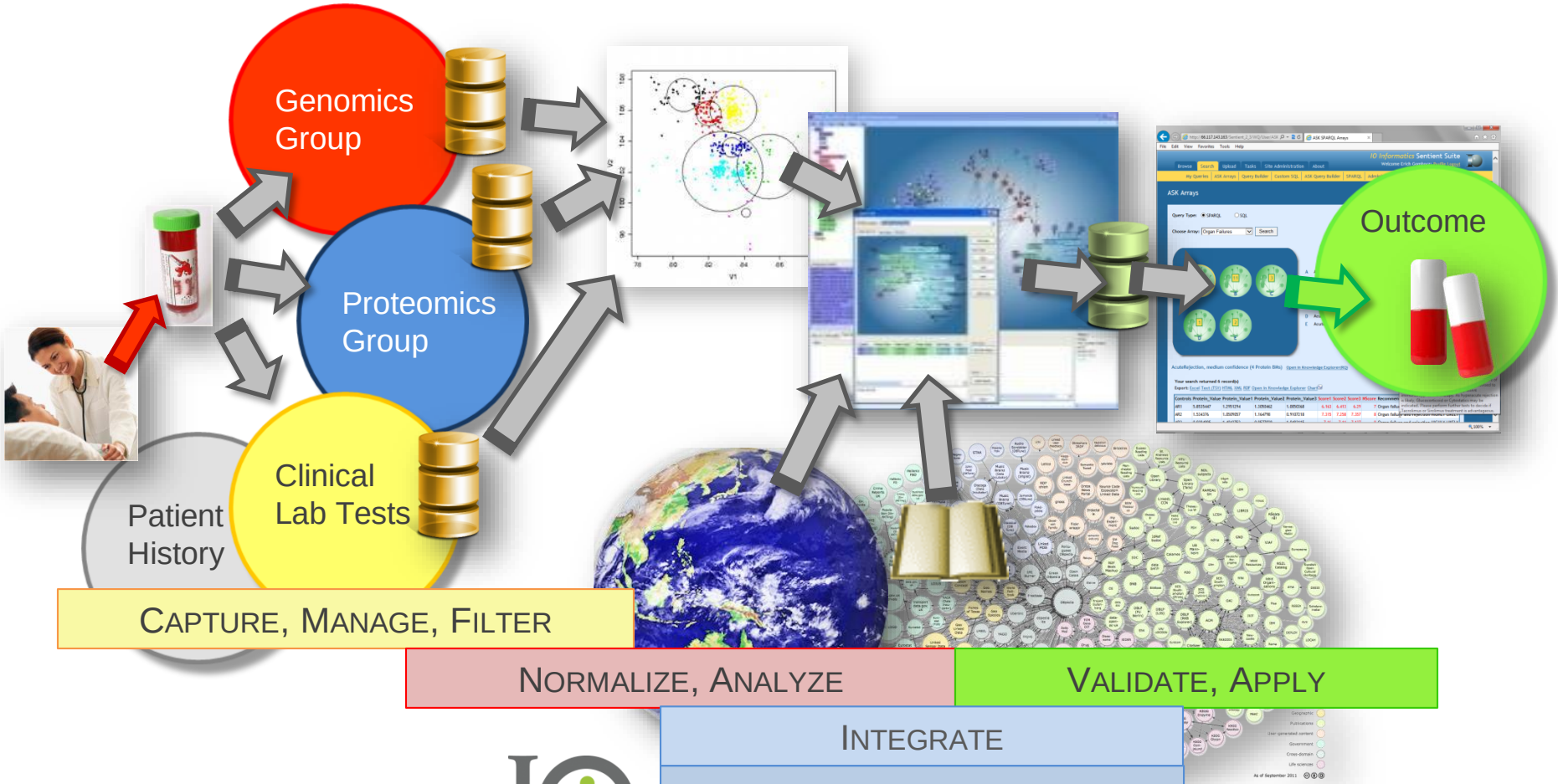


New and better scientific methods and analysis tools require adaptation for changes

INFORMATICS

Integrative Workflow

Outcome-focused, Collaborative Precision Medicine



INFORMATICS

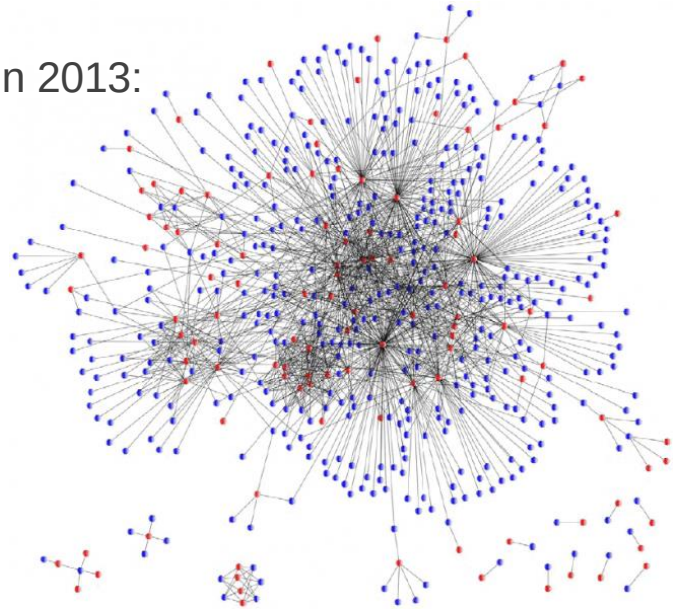
INTEROPERABLE, INTEGRAL SYSTEMS

- Understanding Biology
- *Driven by "functional knowledge"*
- Actionable criteria rather than try-and-retry

Top technology trends impacting information infrastructure in 2013:

- 1 Big Data
- 2 Modern Information Infrastructure
- ✓ 3 Semantic Technologies
- …
- 8 Information Stewardship Applications
- 9 Information Valuation / Infonomics

Source: Gartner (March 2013)



INFORMATICS

Pharma 3.0, Healthcare 3.0

Progression Towards New Life Sciences Models

A noticeable, steady shift in industry from a product-centric business model to a customer/patient centric business model

- Drivers are HEALTH OUTCOMES
- Regaining growth will require the transition of Pharma from acquisition model to innovate partnerships and collaborative data sharing
- Innovation needs to focus more on business model innovation than product innovation
- Reimbursement grounded in real market effectiveness rather than approval of clinical trial data

Shifting to a patient-centric, outcomes-focused and innovation-driven partnership model requires a significant change: reimbursement based on effectiveness in application rather than the product itself

IO INFORMATICS

PHARMA

	Pharma 2.0 Product Portfolio	Pharma 3.0 Health Outcomes
BUSINESS MODEL	PRODUCT-CENTRIC; B2B	CUSTOMER-CENTRIC; B2C
VALUE DRIVERS	REVENUE AND MARGINS	OUTCOMES IN PATIENTS AND HEALTH SYSTEMS
BRAND VALUE	PRODUCT EFFICACY	CUSTOMER EXPERIENCE
GO-TO-MARKET STRATEGY	PITCHING	LISTENING, CO-CREATING
INORGANIC GROWTH	ACQUISITION OF PRODUCT COMPANIES	INNOVATIVE PARTNERSHIP
INNOVATION	PRODUCT INNOVATION	BUSINESS MODEL INNOVATION
INFORMATION	APPROVAL BASED ON CLINICAL DATA	REIMBURSEMENT BASED ON MARKET EFFECTIVENESS

- **SHIFT**

  FROM DIVERSIFIED PRODUCTS TO HEALTH OUTCOMES

Source: Ernst & Young, Global Life Sciences Report (2012)

HEALTHCARE

	Healthcare 2.0 Procedure Portfolio	Healthcare 3.0 Patient Health Outcome
BUSINESS MODEL	PROCEDURE-CENTRIC; B2B	PATIENT-CENTRIC; B2C
VALUE DRIVERS	REVENUE AND MARGINS	OUTCOMES IN PATIENTS AND HEALTH SYSTEMS
BRAND VALUE	PROCEDURE VARIETY	COMPARATIVE EFFECTIVENESS
GO-TO-MARKET STRATEGY	PITCHING	LISTENING, ADJUSTING
INORGANIC GROWTH	EXPANSION OF PROCEDURES	INNOVATIVE PARTNERSHIP
INNOVATION	PROCEDURE INNOVATION	BUSINESS MODEL INNOVATION
INFORMATION	APPROVAL BASED ON PROCEDURE	REIMBURSEMENT BASED ON PATIENT OUTCOME

- **SHIFT**

  FROM PROCEDURE-DRIVEN TO PATIENT'S HEALTH OUTCOME

- Pharmaceutical Industry
- Government
- Clinical Decision Support

4 Use Cases of Adoption

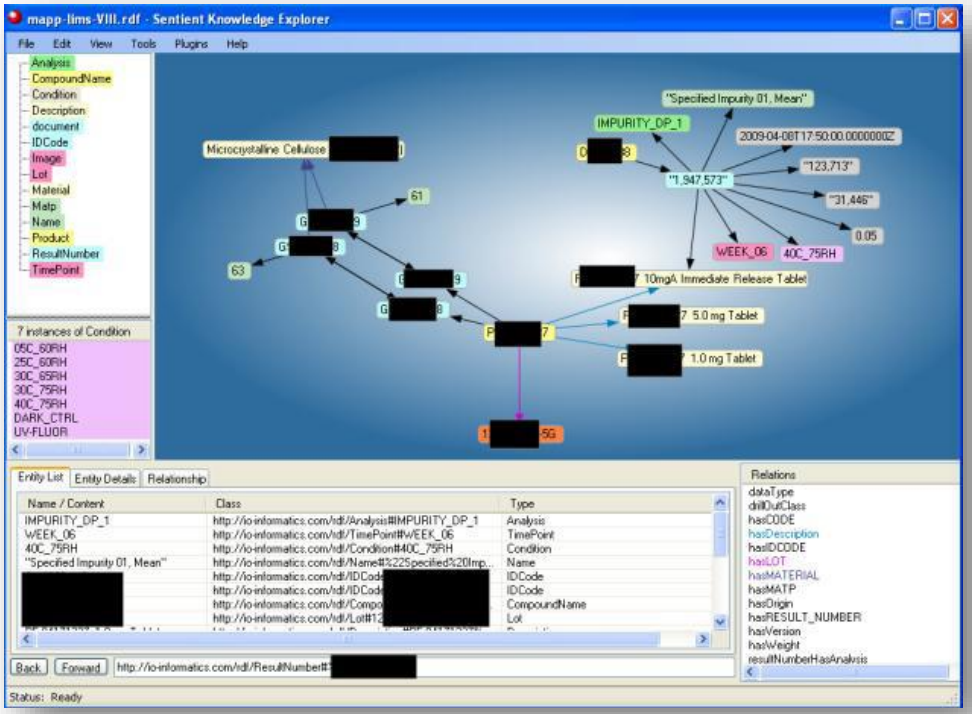
Pharma 1: Drug Purity, Stability

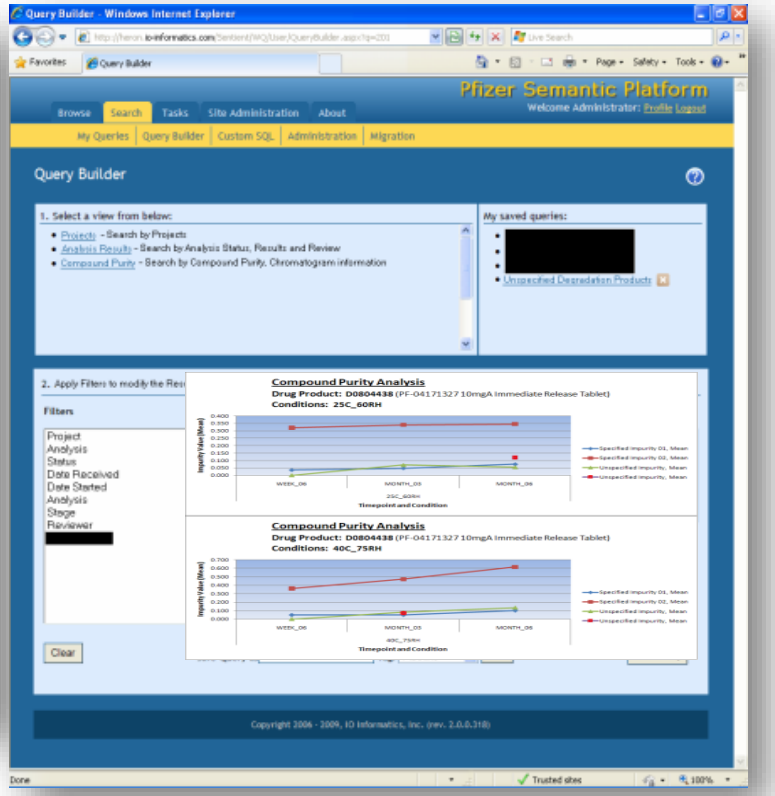
Impact of excipient choice on formulation stability, purity and drug efficacy (Large Pharma)

- Resources
  - Internal CDS, LIMS, Compound DB, Excipient DB, Drug Stability DB
- Problem
  - No common identifiers, time-consuming off-line searches with ambiguous results, delaying responses to the FDA by several weeks
- Solution
  - A semantic platform for compound purity and stability assessment
- RoI / Outcome
  - Cost-efficient, Fast, extensible: accomplished in 6 weeks (1 yr. projected)
  - Inference for quick, unambiguous raw data finding
  - Side benefit: influence of excipient on active ingredient efficacy and overall drug stability

PHARMACEUTICAL MANUFACTURING

FORMULATION INFLUENCE ON DRUG STABILITY





Semantic integration provides immediate report verification and manufacturing based on effect of compound formulation on drug stability, efficacy and purity

INFORMATICS

Pharma 2: Pre-Clinical Toxicity

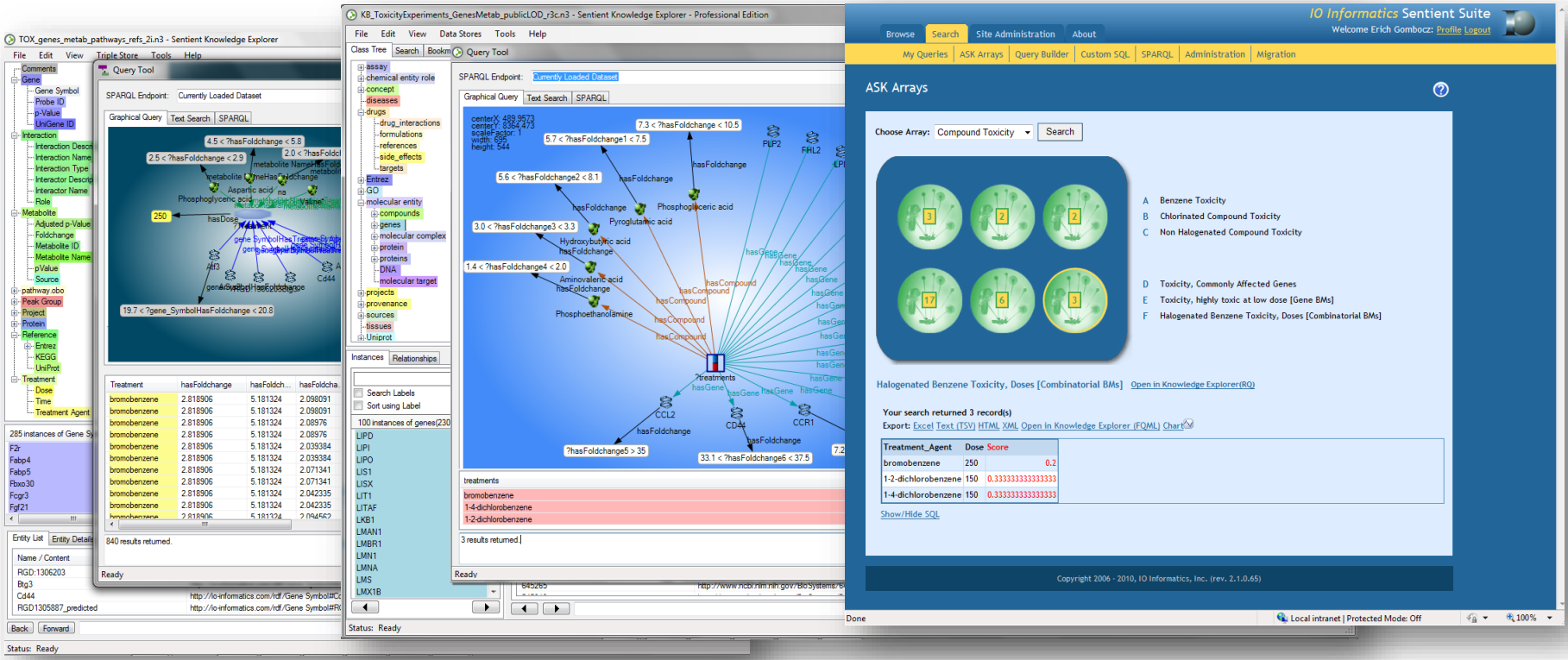
Pre-clinical toxicity assessment and compound toxicity classification (NIST ATP project)

- Resources
  - Animal study results from Genomics (GEP), Metabolomics (MS-MS); External KBs: UniProt, KEGG, Reactome, Diseasome, SIDER, IntAct, BioGrid, HMDB
- Problem
  - Pharmacogenomic correlations across multi-modal studies are not necessarily reflecting biological mechanisms
- Solution
  - Biomarker classifier patterns and their ranges for pre-clinical toxicity screening of compounds
- RoI / Outcome
  - Biomarker-based screening for toxicity with classification of toxicity type

INFORMATICS

Toxicity Classification

Identification of Types of Toxicity (NIST ATP)



Functionally qualified multi-OMIC biomarkers to identify and classify toxicity

Integrated Approach: Biology Confirmed

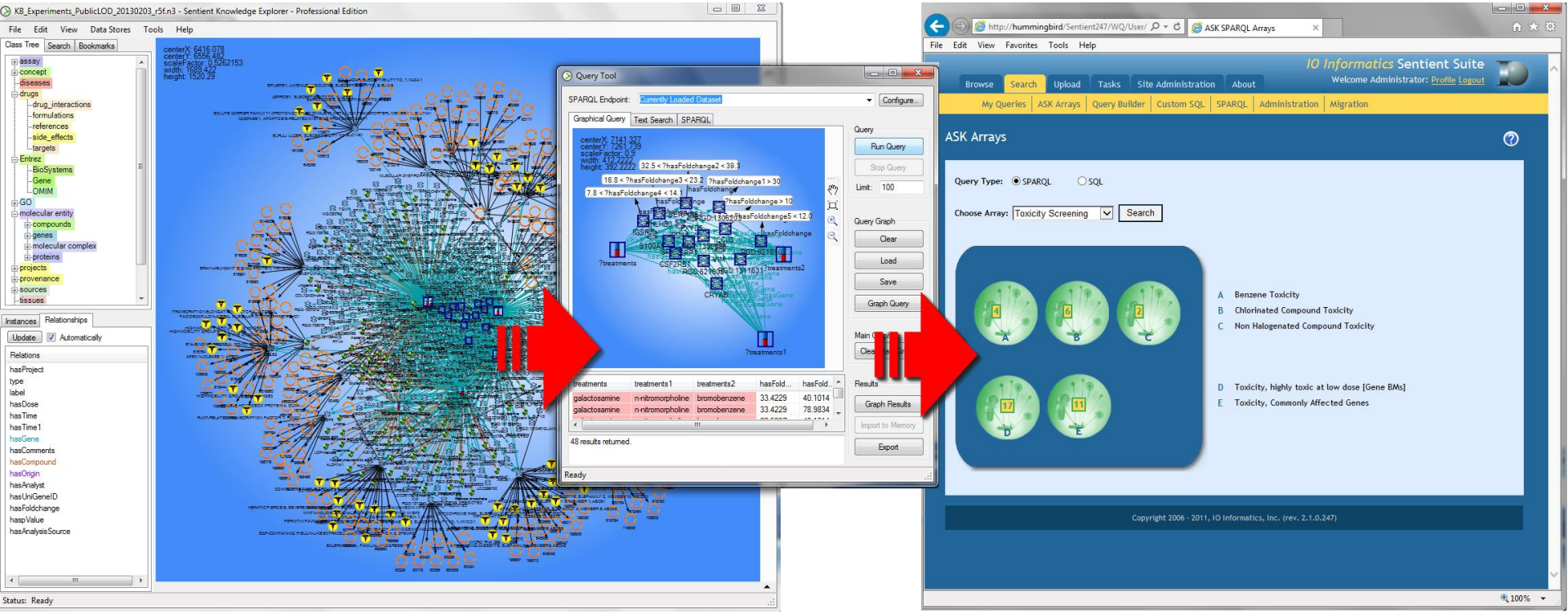
Toxicity Biomarker, Benzene-type toxicity

Marker Class	Instance	UniProt AC	Pathway Gene	Protein	Biology
genes	CYP2C40	P11510	cp2cc	Cytochrome P450 2C40	heme binding, iron ion binding, aromatase activity
genes	AKR7A3	P38918	akr7a3	Aflatoxin B1 aldehyde reductase member 3	detoxification
genes	GPX2	P83645	gpx2	Glutathione peroxidase 2	response to oxidative stress, negative regulation of inflammatory response
genes	MYC	P09416	myc	Myc proto-oncogene protein (Transcription factor p64)	regulation of gene transcription, non-specific DNA binding, activates transcription of growth-related genes
genes	MT1A	P02803, Q91ZP8	mt1a	Metallothionein-1	metal ion binding
genes	HMOX1	P06762	hmox1	Heme oxigenase 1	heme catabolic process, negative regulation of DNA binding
genes	FGF21	Q8VI80	fgf21	Fibroblast growth factor 21 (Protein Fgf21)	positive regulation of ERK1 and ERK2 cascade, MAPKKK cascade and cell proliferation
genes	AKR1B8	Q91W30	akr1b8	Aldose reductase-like protein	oxidoreductase activity
genes	TRIB3	Q9WTQ6	trib3	Tribbles homolog 3	disrupts insulin signaling by binding directly to Akt kinases, expression induced during programmed cell death
genes	YC2	P46418	gsta5	Glutathione S-transferase alpha-5 (EC 2.5.1.18)	response to drug, xenobiotic catabolic process
genes	ABCB1, RGD:619951	P43245	abcb1	Multidrug resistance protein 1 (EC=3.6.3.44)	response to organic cyclic compound, tumor necrosis factor, arsenic-containing substance or ionizing radiation
genes	RGD:1310991	Q5U2P3	Zfand2a	AN1-type zinc finger protein 2A	zinc ion binding
genes	GSTP1, GSTP2	P04906	gstp1	Glutathione S-transferase P (EC 2.5.1.18)	response to toxin, xenobiotic metabolic process, response to reactive oxygen species, response to ethanol
genes	RGD:708417	Q62789	ugt2p7	UDP-glucuronosyltransferase 2B7 (UDPGT 2B7) (EC 2.4.1.17)	major importance in conjugation and subsequent elimination of toxic xenobiotics and endogenous compounds
genes	GCLC	P19468	gclc	Glutamate--cysteine ligase catalytic subunit (EC=6.3.2.2)	response to oxidative stress
genes	TXNRD1	O89049	txnrd1	Thioredoxin reductase 1, cytoplasmic (EC=1.8.1.9)	benzene-containing compound metabolic process, cell redox homeostasis, response to drug
genes	NQO1	P05982	nqo1	NAD(P)H dehydrogenase [quinone] 1 (EC 1.6.5.2)	response to oxidative stress, response to ethanol, superoxide dismutase activity
genes	DDIT4L	Q8VD50	ddit4l	DNA damage-inducible transcript 4-like protein	negative regulation of signal transduction, Inhibits cell growth by regulating TOR signaling pathway
metabolites	Pyroglutamic acid	Q9ER34	aco2	Aconitate hydratase, mitochondrial	citrate metabolism, isocitrate metabolism, tricarboxylic acid cycle
metabolites	Choline	Q64057	aldh7a1	Alpha-aminoadipic semialdehyde dehydrogenase (EC 1.2.1.31)	betaine biosynthesis via choline pathway, response to DNA damage stimulus

Pharmacogenomic correlations of genomic and metabolomic biomarkers qualified with their biological functions

TOXICITY CLASSIFICATION

ENRICHMENT AND VALIDATION OF EXPERIMENTS VIA PUBLIC RESOURCES



Biological mechanism-qualified biomarker pattern permit confident decisions on compound toxicity risk

INFORMATICS

Government 1: FDA CVM

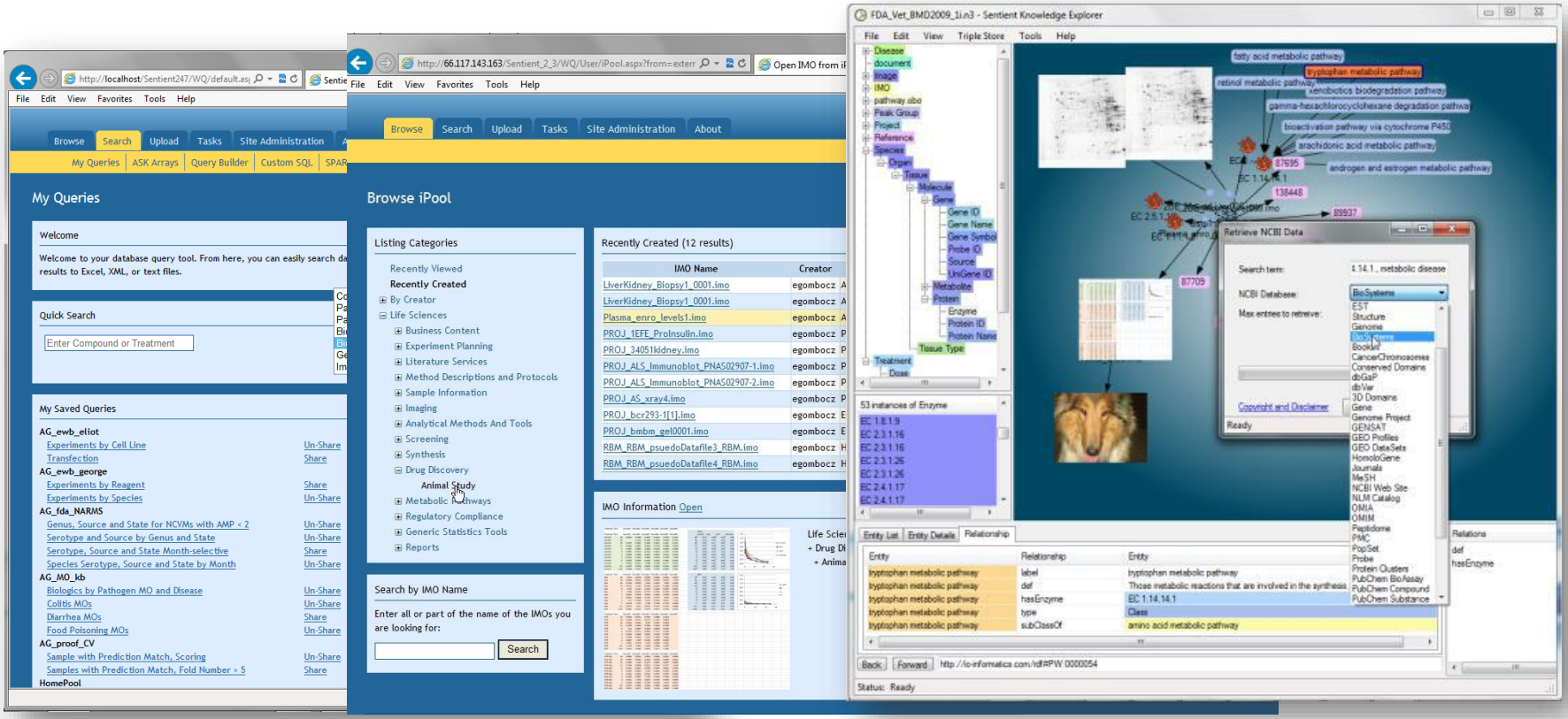
Cross-species disease / adverse event biomarkers to reduce need for animal testing

- Resources
  - Internal: Assays, Proteomics (Gels), Genomics (GEP), Animal observations, Animal biopsies, Imaging endpoints; external KBs: NCBI (Gene, OMIM, Taxonomy, Biosystems), KEGG, UniProt, PubMed
- Problem
  - Disease and adverse effect studies require expensive animal studies which in many cases are not conclusive to be applied to humans
- Solution
  - Find cross-species biomarkers for diseases and adverse effects
- RoI / Outcome
  - Phase 1: Reduced large animal testing, move to smaller animals (shorter cycle}
  - Phase 2: Use of human cell cultures; less (or no) animal testing

INFORMATICS

CROSS-SPECIES BIOMARKERS

REDUCING ANIMAL TESTING



Semantic integration (large animals, small animals, cell culture) for discovery of cross-species biomarkers applicable to human adverse events and diseases

INFORMATICS

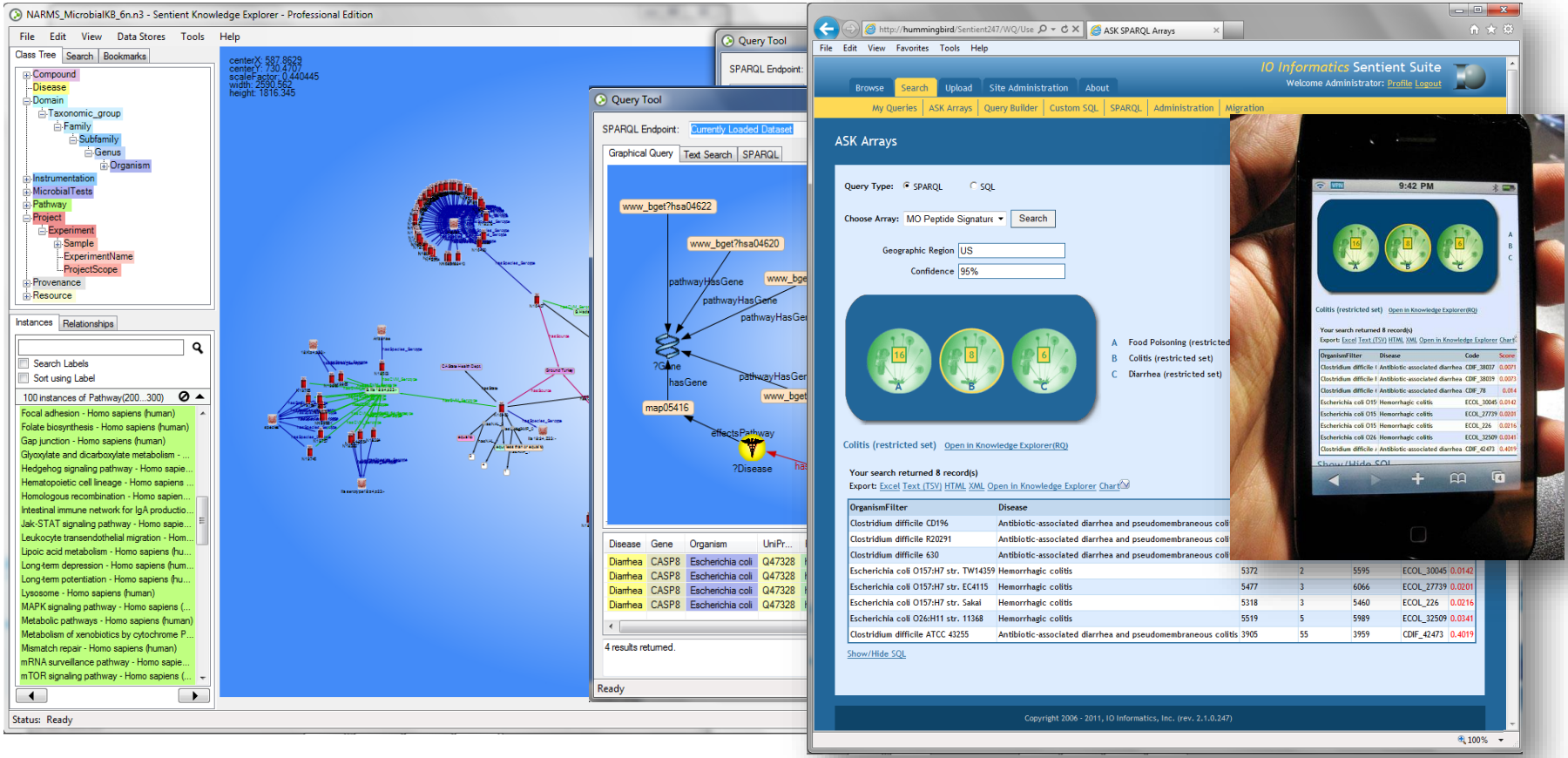
Government 2: NARMS Pathogen

Microbial pathogen knowledgebase to identify biological thread risk for rapid response

- Resources
  - Internal: Microbial Assays, MS Sequencing; External KBs: KBs: ICTV, $MIST_2$, BioCyc, PATRIC, NCBI Taxonomy
- Problem:
  - several public database resources exist, but their schemas are not built for interoperability. Pathogens need to be identified quickly and precisely which entails integral access to as many resources as possible to confidently identify the risk,
- Solution
  - Integral semantic knowledgebase for pathogens with pattern-based alerting
- RoI
  - Rapid risk identification for characterization of biological threads or outbreak of infectious diseases

MICROBIAL KNOWLEDGEBASE

INTEGRATED RESOURCES TO IDENTIFY PATHOGENS

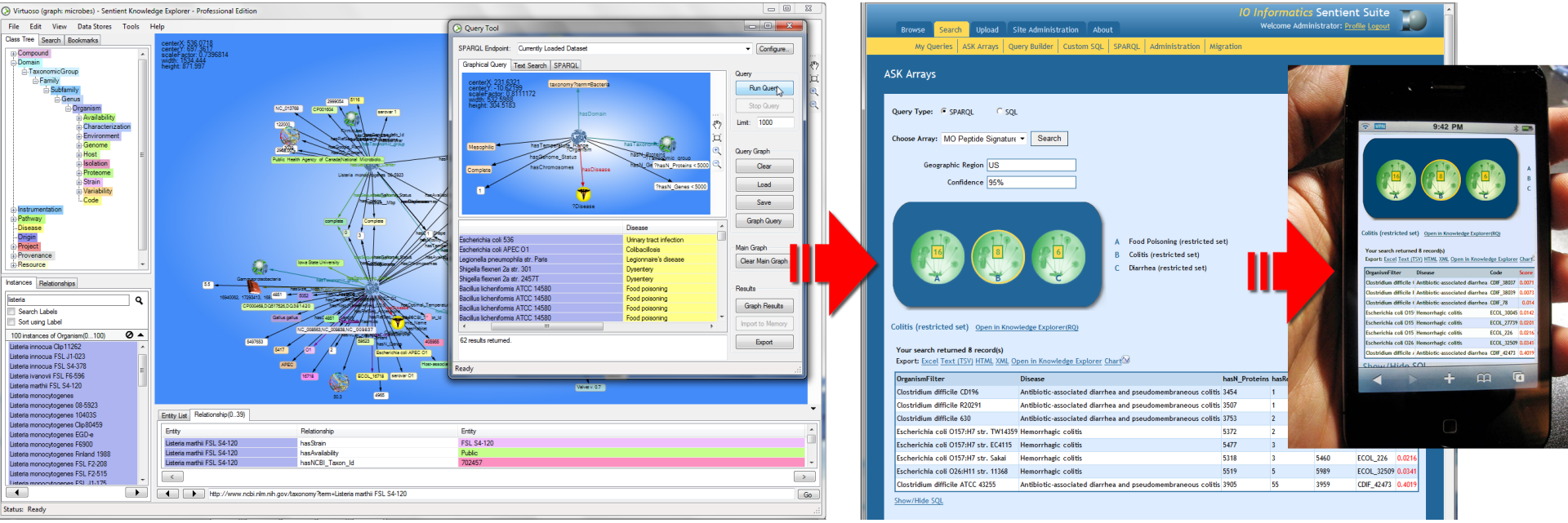


Actionable semantic integration across multiple resources for pathogen detection

INFORMATICS

MICROBIAL KNOWLEDGEBASE

CONTEXTUALIZED EXPERIMENTS FOR FAST, CONFIDENT DECISIONS



Rapid identification and alerting to classify and avert biological threads

INFORMATICS

Clinical1: BIT Organ Rejection

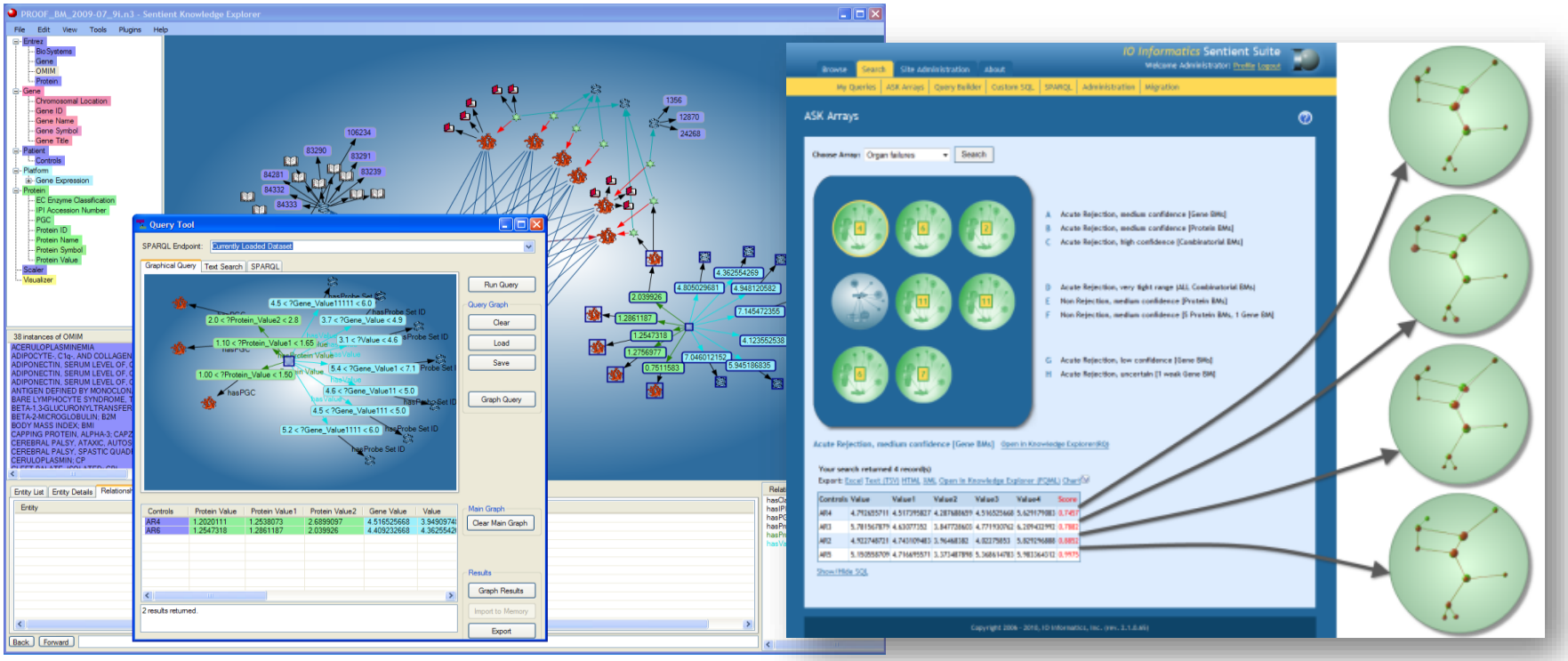
Biomarkers in transplantation for patient screening to avoid costly and invasive biopsies

- Resources
  - Genomics (GEP), Proteomics, Clinical Assays
- Problem
  - Immune reaction after organ transplantation can result in rejection. current methods are based on recurring invasive biopsies, are not always conclusive and strong immune suppression therapy increases other risks, so pre-symptomatic rejection indicators are needed
- Solution
  - Develop combinatorial biomarker test and knowledgebase to screen for patients at risk of organ rejection
- RoI / Outcome
  - An Applied Semantic Knowledgebase (ASK) provides biomarker-pattern based decision support with physician recommendations based on scored risk

INFORMATICS

CAPTURE & QUALIFY RISK PATTERNS

COMBINATORIAL BIOMARKERS FOR ORGAN FAILURE



Integrated Knowledgebase for combinatorial marker-based screening of transplant patients for likelihood of organ failure

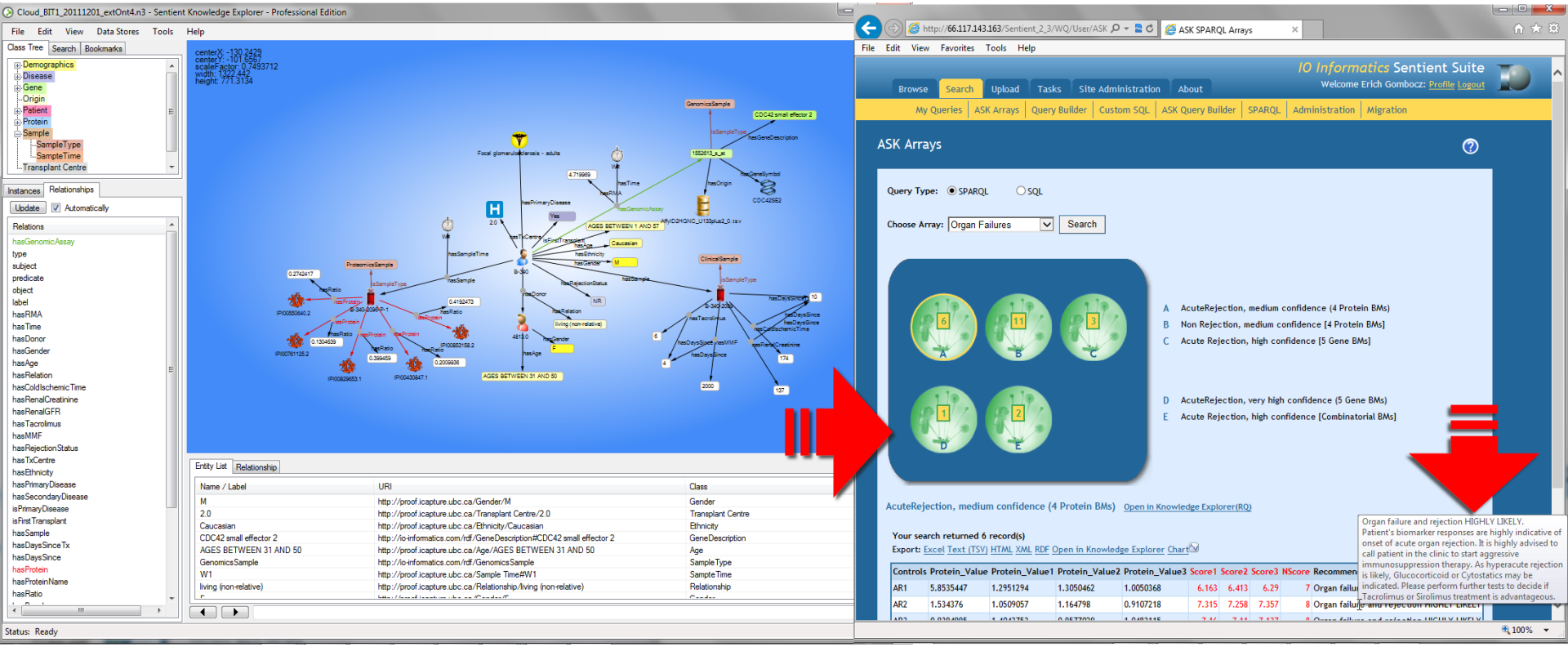
Decision Support

ASK for Precision Medicine

Score-based recommendation for immune suppression therapy at likelihood of rejection, and physician alerting

BIOMARKER-BASED REJECTION RISK

'BEST PRACTICES AWARD' - IMPROVED OUTCOMES IN TRANSPLANTATION



Socio-***economics: risk assessment saves patient's lives and improves their*** quality of life at drastically reduced hospital costs

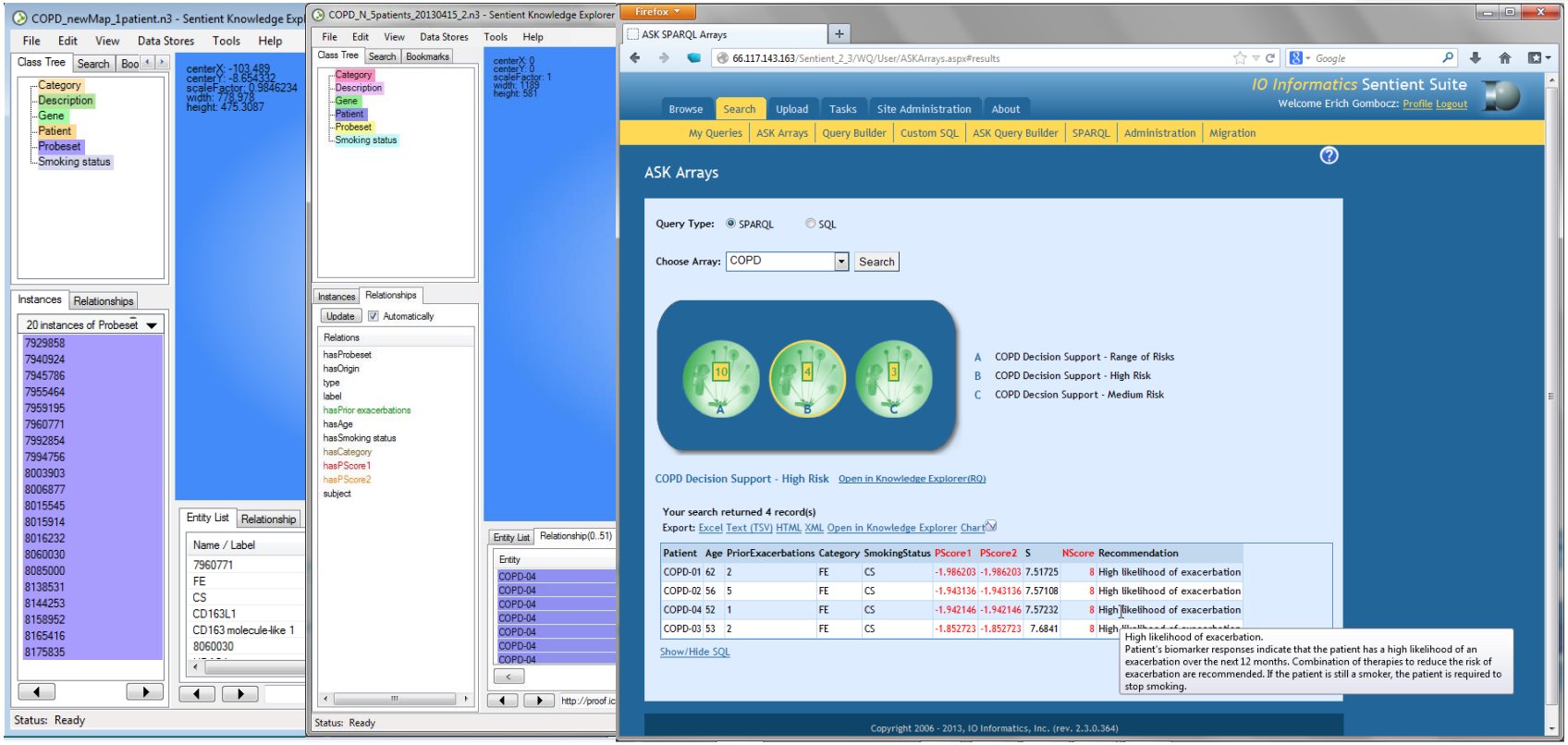
Clinical 2: COPD Prediction

Prediction of COPD events provides a tremendous improvement for patient care through reducing emergency care and hospitalization

- Resources
  - Genomics (GEP), Proteomics, Patient histories, Lifestyle indicators
- Problem
  - COPD exacerbations lead to permanent lung damage. Emergency treatment for acute COPD events is limited, so a decision support system predicting the risk of exacerbations would improve patients outcomes. In Canada, COPD management costs are > $ 1 Billion/year
- Solution
  - Develop combinatorial biomarker test and knowledgebase to screen for patients at risk of an upcoming exacerbation before its manifestation
- RoI / Outcome
  - A biomarker-pattern based predictive decision support system with physician recommendations based on scored risk

INFORMATICS

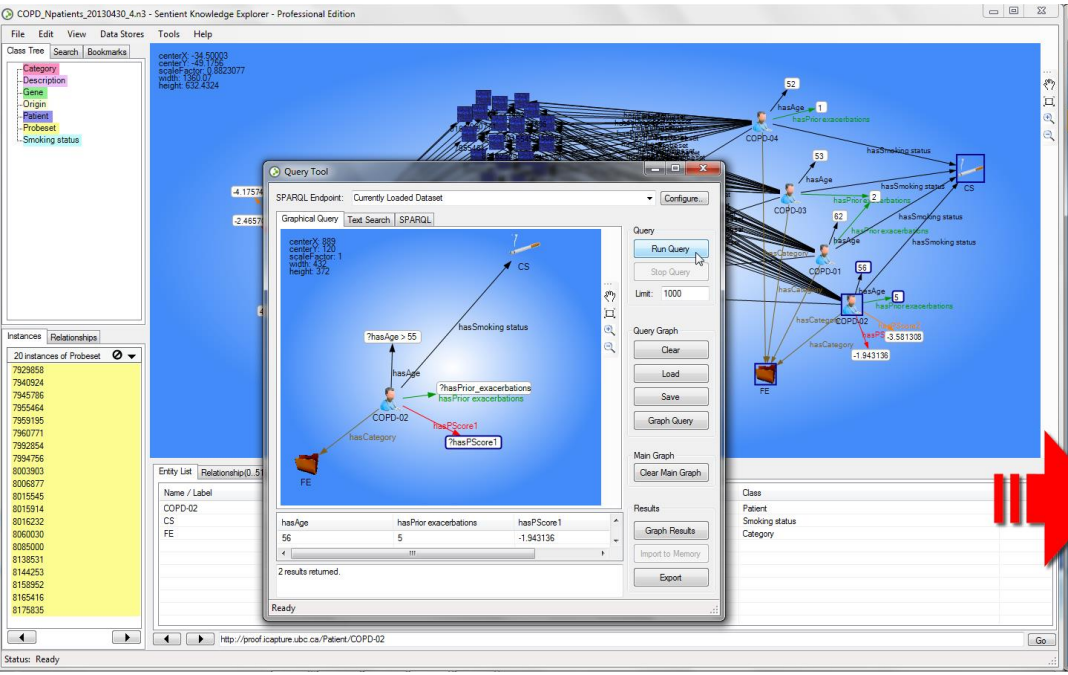
COPD Decision Support

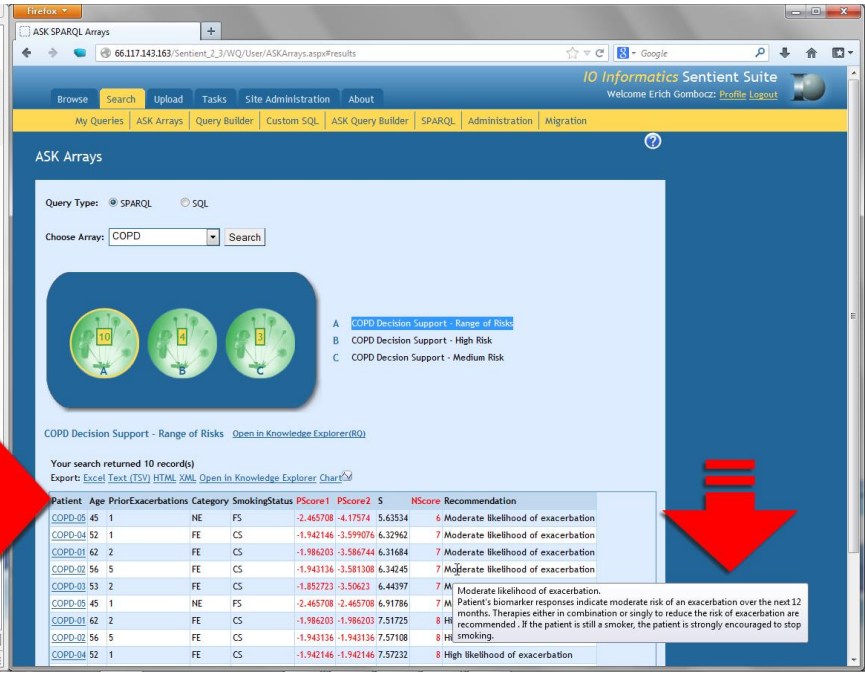


Score-based prediction of exacerbations in COPD with physician recommendation for patients at risk

COPD PHYSICIAN ALERTING

PATTERN-BASED SCREENING TO PREDICT NEXT EXACERBATION





Pattern-based result recommendations include advice for physician and patient about risk aversion on upcoming COPD events

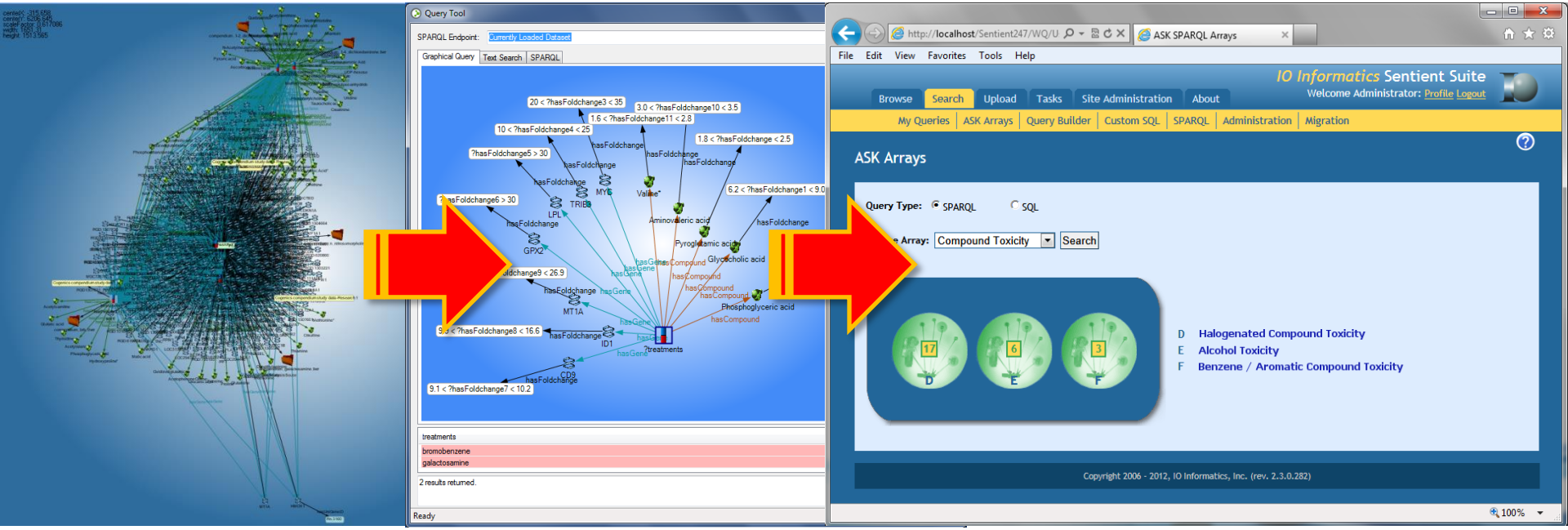
INFORMATICS

What the use cases show

- There are common patterns in all use cases
  - Contextualized, interoperable data
- There are common benefits
  - Patients
    - effective treatment, better outcomes
  - Healthcare Provider
    - drastic cost savings: less emergencies, better preventive care
  - Pharma companies
    - more effective, less adverse effect-prone drugs

Outcome-oriented, integrated knowledge-driven decisions are the uniting cost-saving force in transition of business models in pharma and healthcare

Semantically Integrated Life Sciences Data Provide Actionable Knowledge



- Data Driven, Outcome-Focused
- Socio-Economics, Quality of Life
- Actions Today and Tomorrow

5 Future Outlook

INFORMATICS

Data driven, Outcome-focused

- 1. Fast, extensible knowledge building
- 2. Applied Knowledge is Cost Saver

  - Better knowledge provides the foundation for more confident decisions and better outcomes
  - Fast implementable AND cost-efficient solutions in life sciences will propel the migration to outcome-based reimbursement models using collaborative, open, interoperable data

Business sustainability and the bottom-line are a powerful motivation against model change hesitancy

Socio-Economics 1: Cost savings

- Rapid solutions, widespread applicability
- Avoidance of unnecessary procedures
- Predictive, preventive care: Improved outcomes, less hospitalization
- Patient-centric: best treatment, right drug combinations, right dose
- OECD (2010): public expenditure on health in % of total spending
  - US 48.2%, Netherland 87.7%, Denmark 85.1%, UK 83.2%

- Heart disease and stroke costs the Canadian economy more than $20.9 billion per year in physician services, hospital costs, lost wages and decreased productivity

Actionable knowledge for near-real time alerting of physicians about patients at risk in life-threatening conditions is testimonial to the real value of interoperable, agile data in life science

INFORMATICS

Socio-Economics 2: Quality of Life

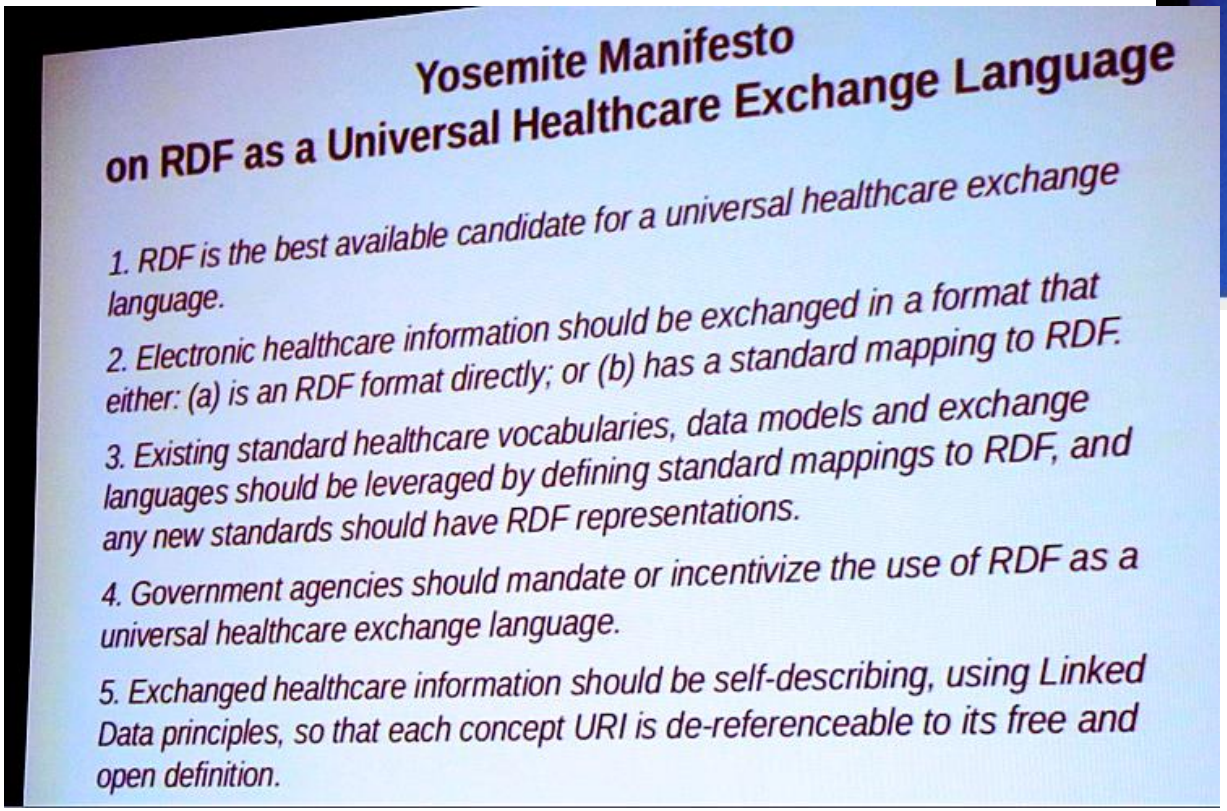
- Applied biological knowledge leads to less invasive procedures - biomarkers instead of biopsies
- Predictive intervention leads to less chronical diseases or lifetime treatments – catch organ problems before they fail; less dialysis, more early detection
- use of semantic ally integrated networks from internal, experimental, clinical, observational and demographic data and public sources to enrich, qualify, validate – even plan additional - experiments has evolved to mainstream acceptability
  - Biomarker-based screening for kidney disease to avoid biweekly dialysis or transplantation
  - Heart transplant patient monitoring with biomarkers instead of costly and unpleasant monthly biopsies
  - Prediction of exacerbations in COPD to prevent permanent lung damage

Applied knowledge at the reward of huge socio-economic benefits, improved prevention, care and quality of life

INFORMATICS

The Yosemite Manifesto

- RDF for universal healthcare exchange



RDF, using linked data principles, provides an interoperable, universal exchange knowledge representation

RDF meets PCAST universal healthcare exchange language requirements

Actions Today and Tomorrow

- We see already today adaption towards open-minded strategic approaches to build integrated, interoperable (and open?) life science knowledge system capable of remarkable results at significantly lower costs
- There still remains a lot to do … to ensure that the life sciences industries are effective and applied to help through early intervention, better prognosis and integrated patient-centric, knowledge-based treatment

- We need to act <u>now</u>
  - ✓ We have the framework and the tools
  - ✓ We have examples
  - We have no time – sustainability of the life sciences / healthcare are at risk!

INFORMATICS

ACKNOWLEDGEMENTS

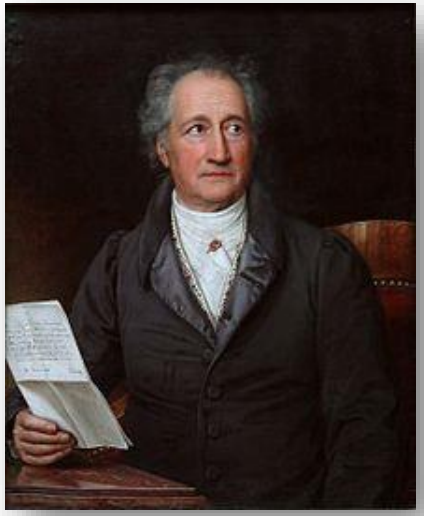
- TOXICITY PROJECT: Pat Hurban, Alan J. Higgins, Imran Shah, Hongkang Mei, Ed K. Lobenhofer (Cogenics, Morrisville, NC), Fulton T. Crews (Bowles Center for Alcohol Studies / UNC, Chapel Hill, NC)
- MICROBIAL PATHOGEN PROJECT: Sherry Ayers (FDA NARMS, Silver Spring, MD)
- SPECIES-INDEPENDENT BIOMARKERS: Haile F. Yancy, Michael J. Myers, Rudell Screven (FDA VET / CVM, Laurel, MD)
- BIOMARKERS IN TRANSPLANTATION AND COPD: Bruce Mc Manus, Raymond T. Ng, Scott Tebbutt (Centre for the Prevention of Organ Failures / PROOF, Vancouver, BC, Canada)
- RDF / OWL DATABASE RESOURCES AND ONTOLOGIES: Jerven T. Bolleman (Swiss Institute Bioinformatics / SIB / UniProt Consortium, Geneva, Switzerland), Michel Dumontier (Bio2RDF II, Carleton University, Ottawa, Canada), Mark A. Musen, Patricia L. Whetzel (BMIR / NCBO Stanford, CA)
- W3C HCLS LLD / PHARMACOGENOMICS SIG: Scott Marshall, Michel Dumontier, Alasdair J G Gray
- IO INFORMATICS: Andrea Splendiani, Charles N. Mead, Jason A. Eshleman, Robert A. Stanley
- WORKING GROUPS: Best Practices in Data Sharing, Informatics for Personalized Medicine

- Grant Support for Toxicity Studies: NIST ATP #70NANB2H3009, NIAAA #HHSN281200510008C

INFORMATICS

Changing the Model in Pharma and Healthcare – Can We Afford to Wait Any Longer?

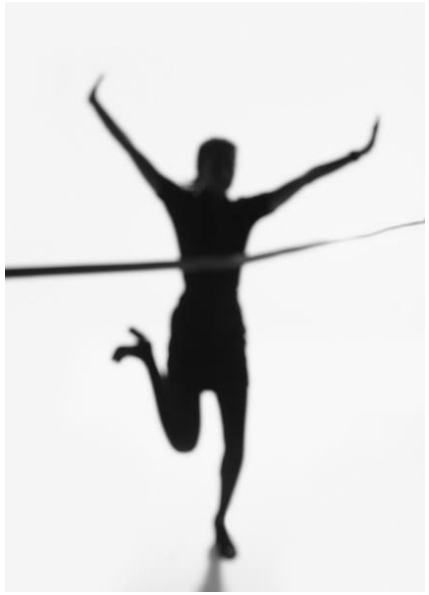
We Cannot!



Johann Wolfgang von Goethe (1782)

"Knowing is not enough; we must apply.
Willing is not enough; we must do."

INFORMATICS



QUESTIONS?

egombocz@io-informatics.com
http;//www.io-informatics.com

THANK YOU !

INFORMATICS


INFORMATICS

REFERENCES

SELECTED REFERENCES (SEE LNBI FOR COMPLETE LIST)

1. Technology survey on LIMS and ELN in Life Sciences. Project Share Biotech, University of Nantes, France (2011)
2. OECD health data systems 2012 (2012)
3. Richter, B.G., Sexton, D.P.: Managing and Analyzing Next-Generation Sequence Data. PLoS Comput Biol 5(6): e1000369. doi:10.1371/journal.pcbi.1000369 (2009)
4. 10 wearable health tech devices to watch. Information Week (2012)
5. Zax, D.: MIT Technology Review, Nov.2011: Is Personal Data the New Currency? (2011)
6. Ohlhorst, F.J.: Big Data Analytics: Turning Big Data into Big Money, Wiley ISBN: 978-1-118-14759-7 (2012)
7. Davis, K.: The $1,000 Genome. The Revolution in DNA Sequencing and the New Era of Personalized Medicine; Free Press, ISBN 9781416569596 (2010)
8. Mardis, E.: The $1,000 genome, the $100,000 analysis? Genome Medicine 2010, 2:84
9. Herper, M.: The truly staggering costs of inventing new drugs. Forbes (2012)
10. PBS, 2012: Healthcare costs: How the US compares with other countries (Oct. 2012)
11. Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation (2004)
12. OWL 2 Web Ontology Language - Document Overview (Second Edition) .W3C Recommendation (2012)
13. Salvadores, M., Horridge, M., Alexander, P.R., Fergerson, R.W., Musen, M.A., Noy, N.F.: Using SPARQL to Query BioPortal Ontologies and Metadata. International Semantic Web Conference. Boston US. LNCS 7650, pp. 180195 (2012)

INFORMATICS

REFERENCES

14. Luciano, J.S., Andersson, B., Batchelor, C., Bodenreider, O., Clark, T., Denney, C.K., Domarew, C., Gambet, T., Harland, L., Jentzsch, A., Kashyap, V., Kos, P., Kozlovsky, J., Lebo, T., Marshall, S.M., McCusker, J.P., McGuinness, D.L., Ogbuji, C., Pichler, E., Powers, R.L., Prud'hommeaux, E., Samwald, M., Schriml, L., Tonellato, P.J., Whetzel, P.L., Zhao, J., Stephens, S., Dumontier, M..: The Translational Medicine Ontology and Knowledge Base: driving personalized medicine by bridging the gap between bench and bedside. J.Biomed.Semantics 2011; 2 (Suppl 2):S1 (2011)
15. Connecting medical informatics and bioinformatics: Advances in knowledge acquisition and management . Mayo Clinic's Enterprise Data Trust, JAMIA PMC3000789 (2010)
16. Richter, B.G., Sexton, D.P.: Managing and Analyzing Next-Generation Sequence Data. PLoS Comput Biol 5(6): e1000369. doi:10.1371/journal.pcbi.1000369 (2009)
17. Bizer C,Heath T, Berners-Lee T, Hausenblas M.: LDOW2012 Linked Data on the Web. WWW Workshop on Linked Data on the Web, 2012 Apr.16, Lyon, France (2012).
18. Callahan, A., Cruz-Toledo, J., Ansell, P., Klassen, D., Tumarello, G., Dumontier, M.: Improved dataset coverage and interoperability with Bio2RDF Release 2. SWAT4LS Workshop. Paris, France (2012).
19. Plasterer, T.N., Stanley, R., Gombocz, E.: Correlation Network Analysis and Knowledge Integration In: Applied Statistics for Network Biology: Methods in Systems Biology M. Dehmer, F. Emmert-Streib, A. Graber, A. Salvador (Eds.) Wiley-VCH, Weinheim, ISBN: 978-3-527-32750-8 (2011)
20. Ernst & Young: The third place: healthcare everywhere – Global Life Sciences Report (2012)

INFORMATICS

REFERENCES

21. Gombocz, E.A., Higgins, A.J., Hurban, P., Lobenhofer, E.K., Crews, F.T., Stanley, R.A., Rockey, C., Nishimura, T.: Does network analysis of integrated data help understanding how alcohol affects biological functions? - Results of a semantic approach to biomarker discovery 2008 Biomarker Discovery Summit 2008, Philadelphia, PA (2008)
22. Higgins, A.J., Gombocz, E.A., Stanley, R.A.: From Correlation to Biological Understanding: Multi-modal semantic networks for biomarker discovery and qualification. Bio-IT World 2008, Boston, MA (2008)
23. Gombocz, E., Stanley, R.: Predictive Toxicology: Applied Semantics with a major impact on drug safety
24. FDA Drug Safety Workshop: Pharmacological Mechanism-Based Drug Safety Assessment and Prediction. FDA, White Oak Campus, Silver Spring, MD (2011)
25. Gombocz, E., Candlin, J.: A Novel Approach to Recognize Peptide Functions in Microorganisms: Establishing Systems Biology-based Relationship Networks to Better Understand Disease Causes and Prevention 8th Annual Conference US Human Proteome Organization: The Future of Proteomics (HUPO 2012), San Francisco, CA, (2012)
26. Gombocz, E., Candlin, J.: How semantic technology helps fighting infectious diseases: Biological systems approach to understand microbial pathogens. Semantic Technology & Business Conference (SemTech 2012), San Francisco, (2012)
27. Gombocz, E., Candlin, J., Stanley, R., Chiang, D.: Semantically Enhancing Protein Identification: Systems Biology Knowledgebase for Infectious Disease Screening..Bio-IT World 2012, Boston, MA (2012)

INFORMATICS

REFERENCES

28. A new way to predict and diagnose organ rejection. Transplant Foundation Research of British Columbia (2010)
29. Bio-IT World Best Practices Award 2010 in Personalized & Translational Medicine: PROOF / iCAPTURE Centre of Excellence; Semantic Data Integration, Knowledge Building and Sharing Applied to Biomarker Discovery and Patient Screening for Pre-symptomatic Heart, Lung or Kidney Failure in Transplantation Medicine (2010)
30. Ng, R.T., Gombocz, E.: Biomarker Development to Improve Decision Support for the Treatment of Organ Failures: How Far Are We Today? ADAPT 2010, Arlington, VA (2010)
31. Stanley, R., McManus, B., Ng, R., Gombocz, E., Eshleman, J., Rockey, C.: W3C Semantic Web Use Cases and Case Studies Case Study: Applied Semantic Knowledgebase for Detection of Patients at Risk of Organ Failure through Immune Rejection. Joint Case Study of IO Informatics and University British Columbia (UBC), NCE CECR PROOF Centre of Excellence, James Hogg iCAPTURE Centre, Vancouver, BC, Canada (2011)
32. Mathers, C.D., Loncar, D.: Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. Nov 2006;3(11):e442.(2006)
33. Mittmann, N., Kuramoto, L., Seung, S.J., Haddon, J.M., Bradley-Kennedy, C., Fitzgerald, J.M.: The cost of moderate and severe COPD exacerbations to the Canadian healthcare system. Respir. Med. Mar 2008;102(3):413-421.(2008)
34. Gombocz, E.A.: On the road to production: Semantic integration cases indicate successful adoption to improve knowledge-based decisions in Pharma and healthcare NCBO Webinar Series, Stanford, CA. Recording (2013)

INFORMATICS