Biomedical data sharing to enable Learning Health Systems

Jonathan C. Silverstein, MD, MS, FACS, FACMI
Chief Research Informatics Officer,
Health Sciences and Institute of Precision Medicine
Visiting Professor, Department of Biomedical Informatics
Affiliate Scholar, Pitt Cyber
University of Pittsburgh

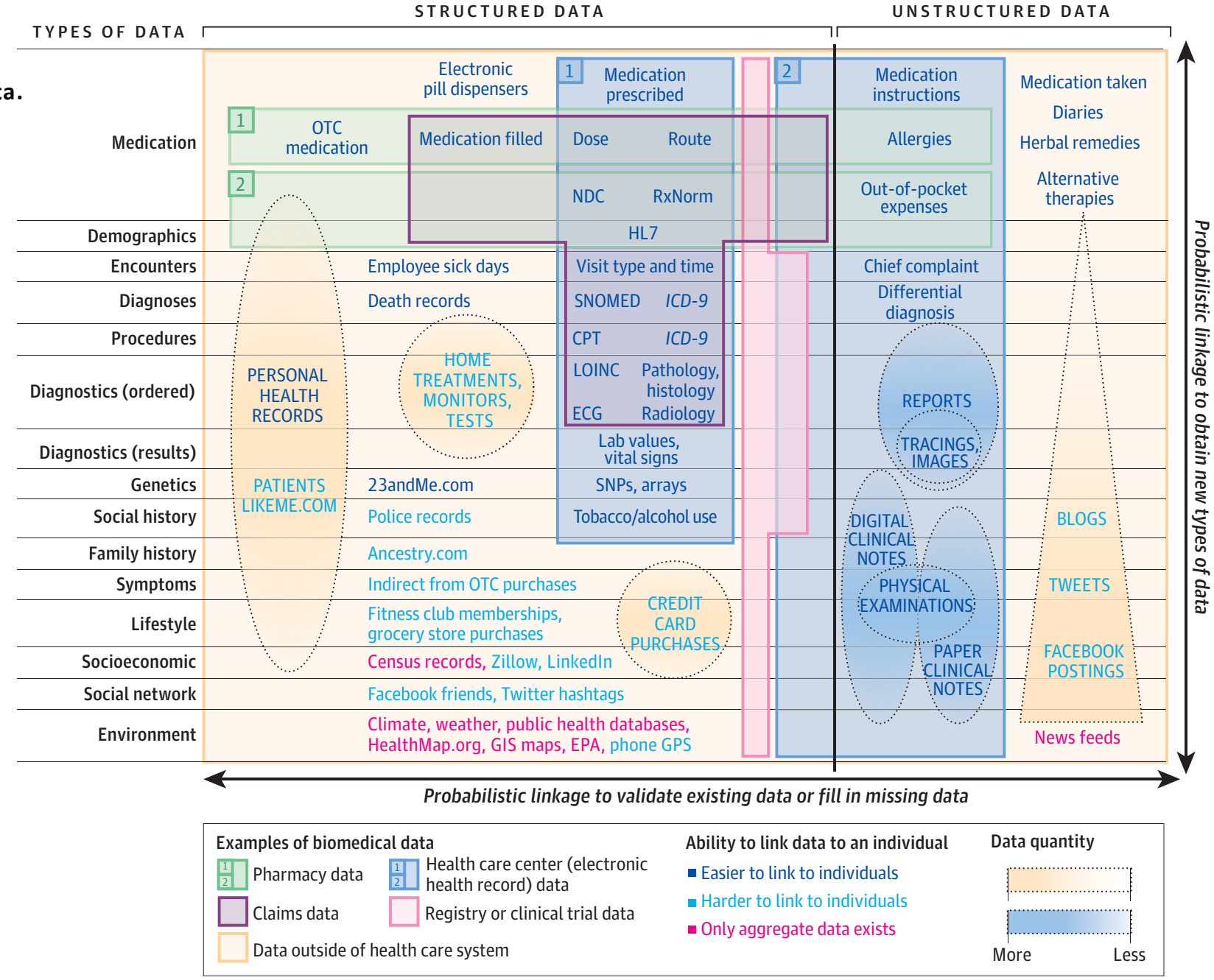
U.S. healthcare challenges in a slide?

- People are dying of preventable causes.
- Cost is out of control.
- Quality can’t be measured.
- Variability is local and widespread.
- New technology is exponentiating.
- Decision-making is maximally distributed.
- Data is not available routinely for learning.

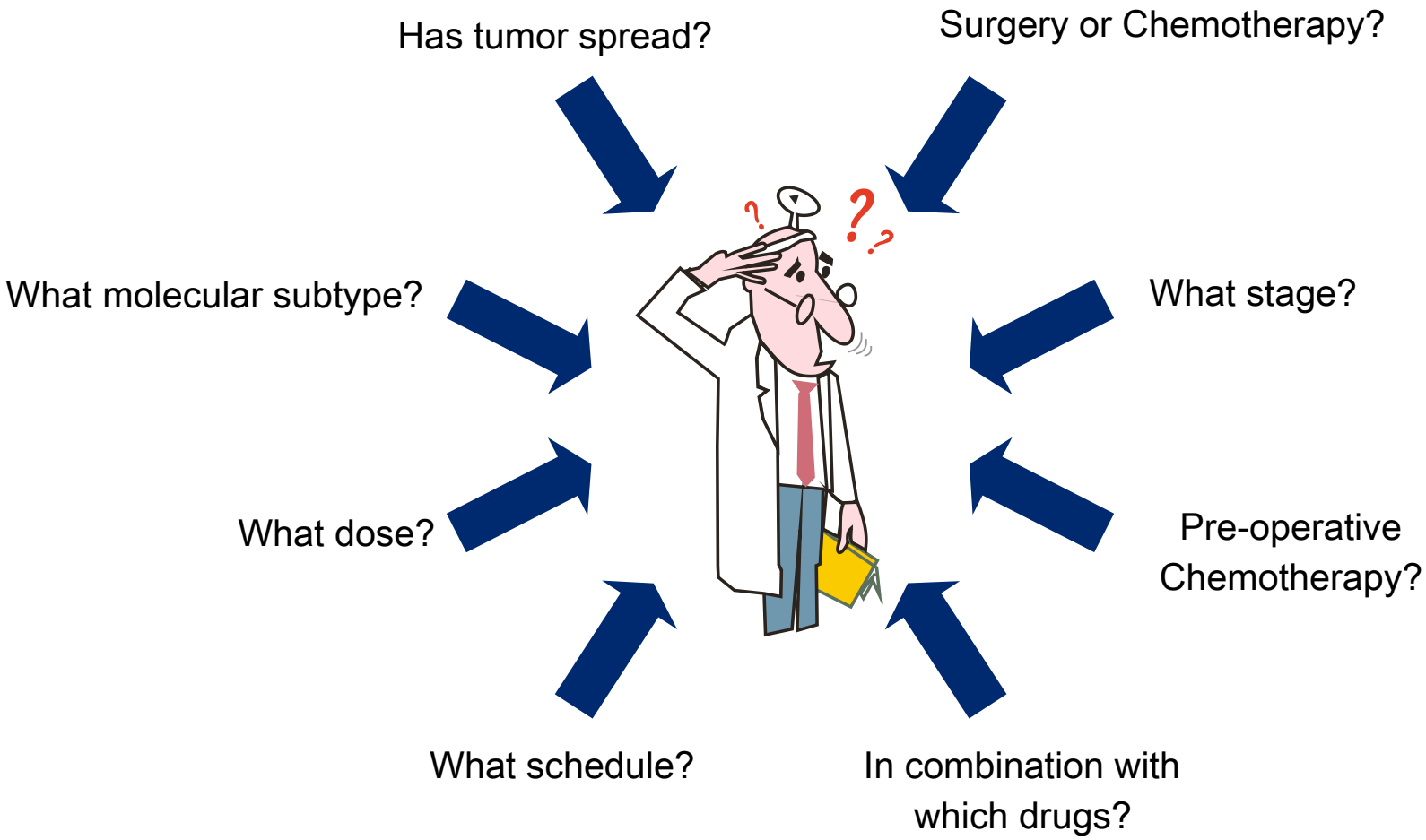
LEARNING HEALTH COMMUNITY

Weber GM, Mandl KD, Kohane IS.
Finding the Missing Link for Big Biomedical Data.
JAMA. 2014 Jun 25;311(24):2479–2480.
PMID: 24854141

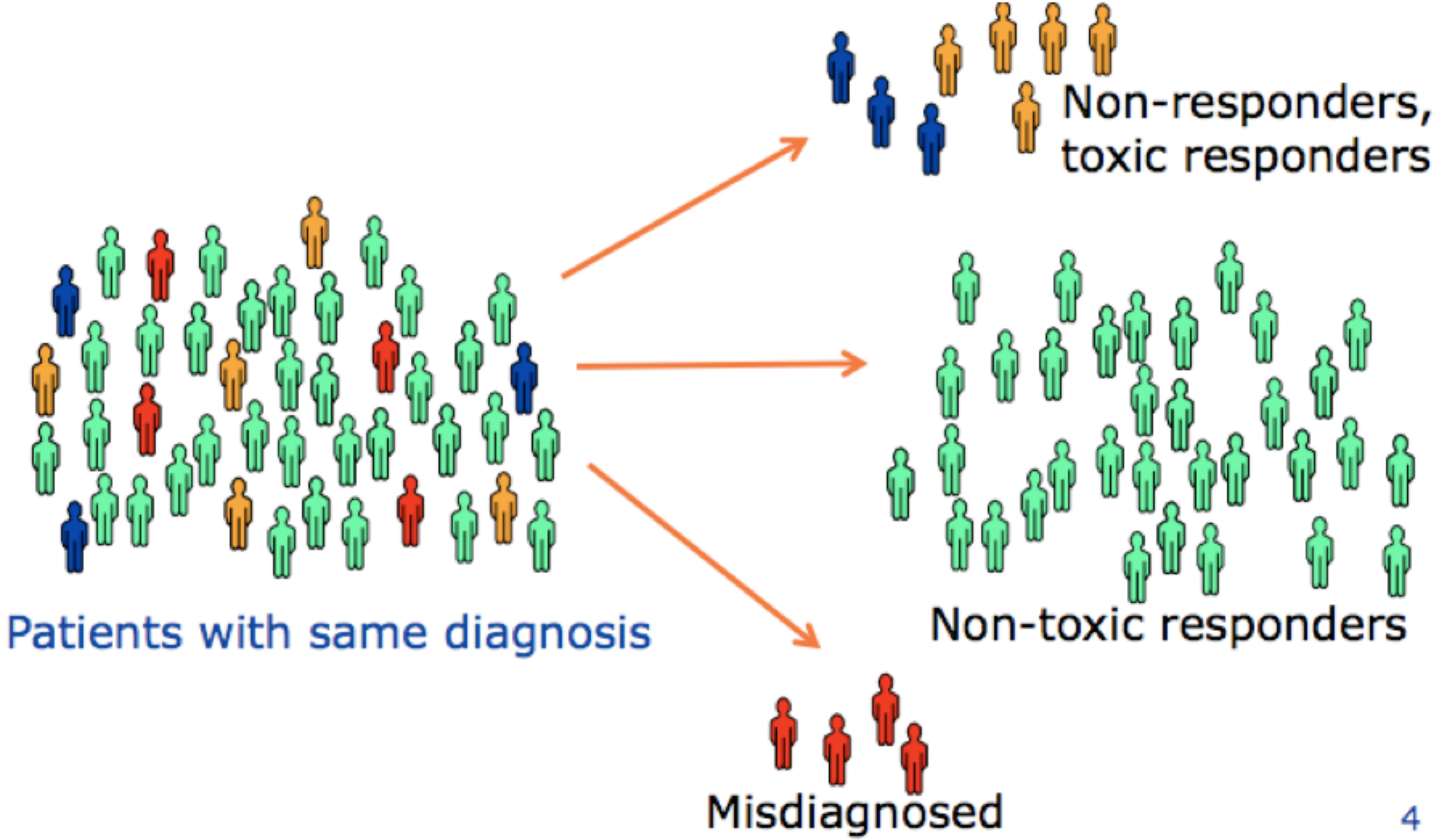
Figure. The Tapestry of Potentially High-Value Information Sources That May be Linked to an Individual for Use in Health Care



Provider, Patient & Payor Faced With Bewildering Choices: The Current Practice of "Qualitative" Medicine



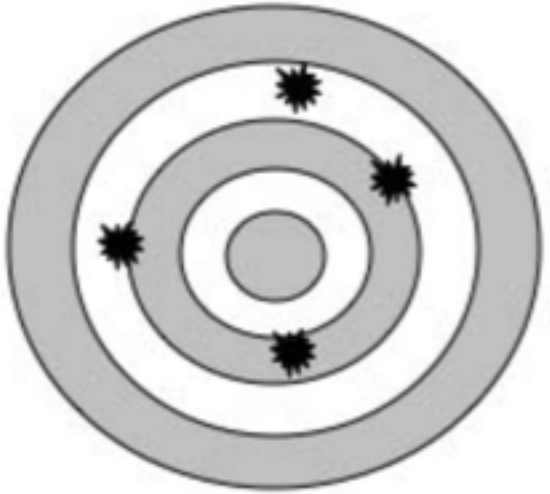
From Patrick Soon-Shiong, MD


Non-responders,
toxic responders
Patients with same diagnosis
Non-toxic responders
Misdiagnosed

NOAA.gov

https://celebrating200years.noaa.gov/magazine/tct/accuracy_vs_precision.html

Not Accurate
Not Precise



Accurate
Not Precise

Not Accurate
Precise

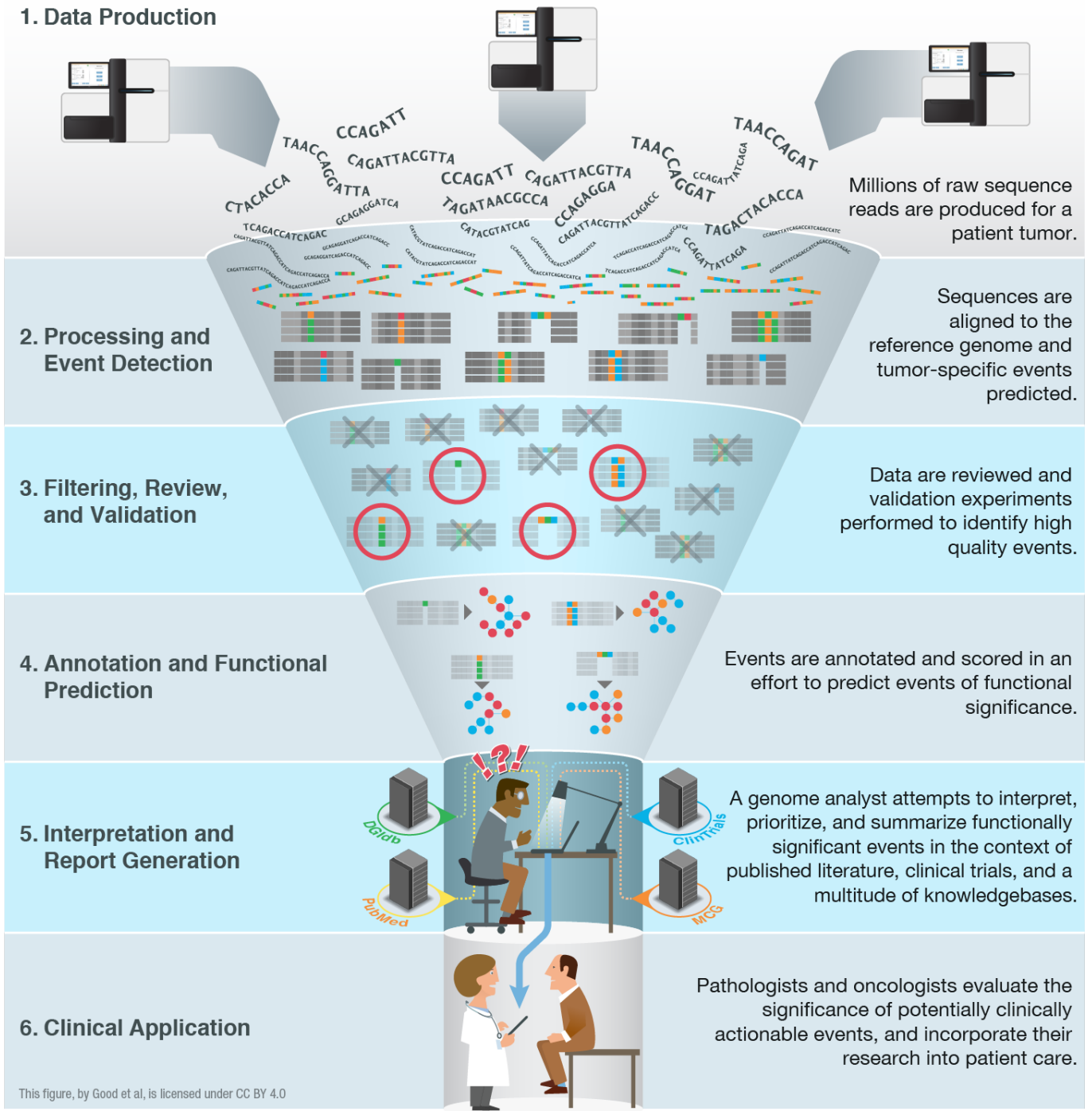
Accurate
Precise

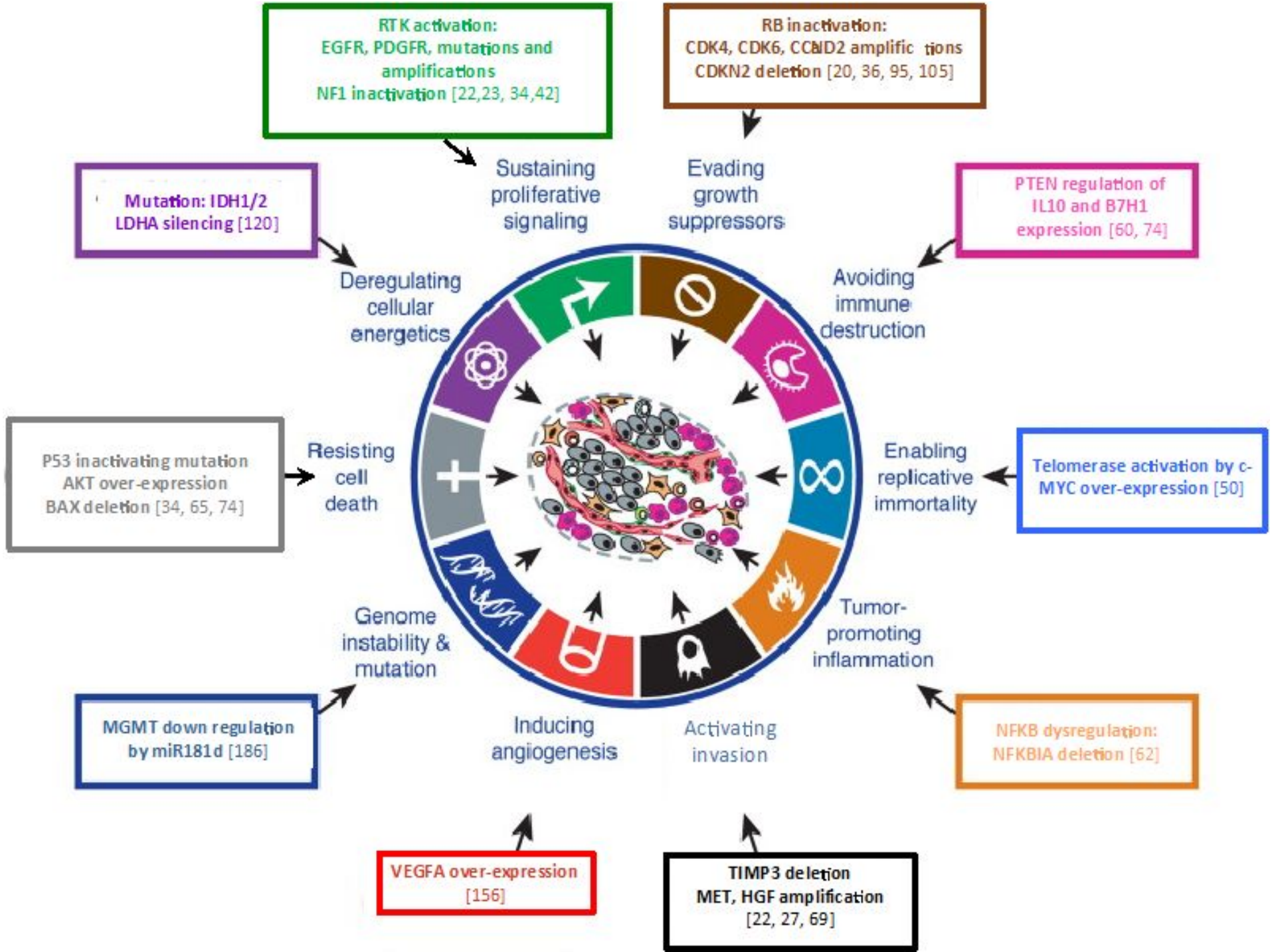
Good BM, Ainscough BJ, McMichael JF, Su AI, Griffith OL.
Organizing knowledge to enable personalization of medicine in cancer.
Genome Biol. 2014
Aug 27;15(8):438.
PMCID: PMC4281950

Easy

Hard

Hardest




RTK activation:
EGFR, PDGFR, mutations and amplifications
NF1 inactivation [22,23, 34,42]
RB inactivation:
CDK4, CDK6, CCND2 amplifications
CDKN2 deletion [20, 36, 95, 105]
Sustaining proliferative signaling
Evading growth suppressors
Mutation: IDH1/2
LDHA silencing [120]
PTEN regulation of IL10 and B7H1 expression [60, 74]
Deregulating cellular energetics
Avoiding immune destruction
P53 inactivating mutation
AKT over-expression
BAX deletion [34, 65, 74]
Resisting cell death
Enabling replicative immortality
Telomerase activation by c-MYC over-expression [50]
Genome instability & mutation
Tumor-promoting inflammation
MGMT down regulation by miR181d [186]
NFKB dysregulation:
NFKBIA deletion [62]
Inducing angiogenesis
Activating invasion
VEGFA over-expression [156]
TIMP3 deletion
MET, HGF amplification
[22, 27, 69]

The Promise of Personalized Medicine

- Accelerate drug development, biomarker discovery, and guide diagnosis, treatment, and prevention
- Detect disease at an earlier stage, when it is easier to treat effectively
- Shift practice from reaction to prevention
- Reduce the overall cost of healthcare

- (credit Rebecca Crowley Jacobson, VP, UPMC Enterprises)

PRESS RELEASES

NCI Press Release

NCI-MATCH precision medicine clinical trial releases new findings, strengthens path forward for targeted cancer therapies

Posted: June 4, 2018

Contact: NCI Press Office
240-760-6600

The National Cancer Institute's Molecular Analysis for Therapy Choice (NCI-MATCH) trial, the largest precision medicine trial of its kind, has achieved a milestone with the release of results from several treatment arms, or sub-studies, of the trial. The new results offer findings of interest for future cancer research that could ultimately play a role in bringing targeted treatments to patients with certain gene abnormalities, regardless of their cancer type.

Credit: National Cancer Institute

Findings from three arms were released at this year's American Society of Clinical Oncology (ASCO) annual meeting in Chicago, adding to findings from one arm released in November 2017. The study was co-developed by NCI, part of the National Institutes of Health, and the ECOG-ACRIN Cancer Research Group, part of the NCI-sponsored National Clinical Trials Network (NCTN). ECOG-ACRIN and NCI are co-leading the trial.

"The outcomes data being released today from this groundbreaking precision medicine trial are an exciting step for NCI-MATCH," said Lyndsay Harris, M.D., of NCI's Cancer Diagnosis Program and NCI study chair. "These findings represent a large collection of data in populations of patients who may not have been studied in conventional clinical trials, and they will have important implications for future precision medicine trials."

Nov 30, 2017

FDA Approves Foundation Medicine's FoundationOne CDx™, the First and Only Comprehensive Genomic Profiling Test for All Solid Tumors Incorporating Multiple Companion Diagnostics

--Landmark approval advances personalized cancer care as an estimated 1 in 3 patients across five common advanced cancers are expected to match with an FDA-approved therapy--

--The Centers for Medicare and Medicaid Services issued a preliminary National Coverage Determination (NCD) for FoundationOne CDx, improving access to molecular information for personalized healthcare--

CAMBRIDGE, Mass.--(BUSINESS WIRE)-- Foundation Medicine, Inc. (NASDAQ:FMI) today announced that the U.S. Food and Drug Administration (FDA) approved FoundationOne CDx™, the company's comprehensive companion diagnostic test for solid tumors. FoundationOne CDx is intended for use by health care professionals to help inform cancer treatment management in accordance with professional guidelines for patients with solid tumors. The first and only FDA-approved test of its kind for all solid tumors, FoundationOne CDx is a diagnostic test that acts as:

This press release features multimedia. View the full release here: http://www.businesswire.com/news/home/20171130006320/en/

Foundation Medicine is headquartered in Cambridge, MA (Photo: Business Wire)

- a comprehensive companion diagnostic to identify patients who may benefit from treatment with specific FDA-approved targeted therapies;
- a comprehensive genomic profiling (CGP) test that includes genomic biomarkers to help inform the use of other targeted oncology therapies, including immunotherapies;
- a tool for physicians that identifies patient opportunities for clinical trial participation; and,
- an FDA-approved platform for companion diagnostic development for biopharma companies developing precision therapeutics.

FoundationOne CDx assesses all classes of genomic alterations in 324 genes known to drive cancer growth, providing potentially actionable information to help guide treatment decisions. It is also indicated as a companion diagnostic for patients with certain types of non-small cell lung cancer (NSCLC), melanoma, colorectal cancer, ovarian cancer or breast cancer to identify those patients who may benefit from treatment with one of 17 on-label targeted therapies, including 12 therapies currently approved as first-line therapy for their respective indications. FoundationOne CDx also reports genomic biomarkers, such as microsatellite instability (MSI) and tumor mutational burden (TMB), that can help inform the use of immunotherapies; genomic alterations in other genes relevant to patient management; and, relevant clinical trial information.

Based on previous CGP testing conducted by Foundation Medicine, it is estimated that approximately 1 in 3 patients across five common advanced cancers are expected to match with an FDA approved therapy.[1] The number of matched on-label therapies indicated on FoundationOne CDx is expected to increase over time as Foundation Medicine and its biopharma partners pursue FDA approval for additional companion diagnostics on the platform. Today, approximately 50% of new cancer drugs in development are projected to have a companion biomarker.[2]

News Releases

CancerLinQ Engages Leading Technology Companies Tempus and Precision HealthAI to Accelerate Data-Driven Insights to Oncologists and the Cancer Care Community

FOR IMMEDIATE RELEASE

December 21, 2017

TEMPUS

Data-driven cancer treatment

Genomic Alterations

Targeted Therapy

Clinical Trials

Immunotherapy Options

Patient Cohort

Patient Cohort aggregates deidentified clinical, genomic, and outcomes data from previously tested patients, allowing you to evaluate treatment regimens for your patients with similar clinical and genomic presentation.

The future includes even more data

- Sequencing of entire exome and entire genome
- Sequencing of individual tumor cells
- Detection of tumor sequence fragments in blood
- Sequencing of multiple areas of a tumor
- Sequencing of metastases and recurrence
- Assembling more integrative analyses across DNA, RNA, protein
- Algorithms to help us untangle the complex molecular changes to find the drugable targets

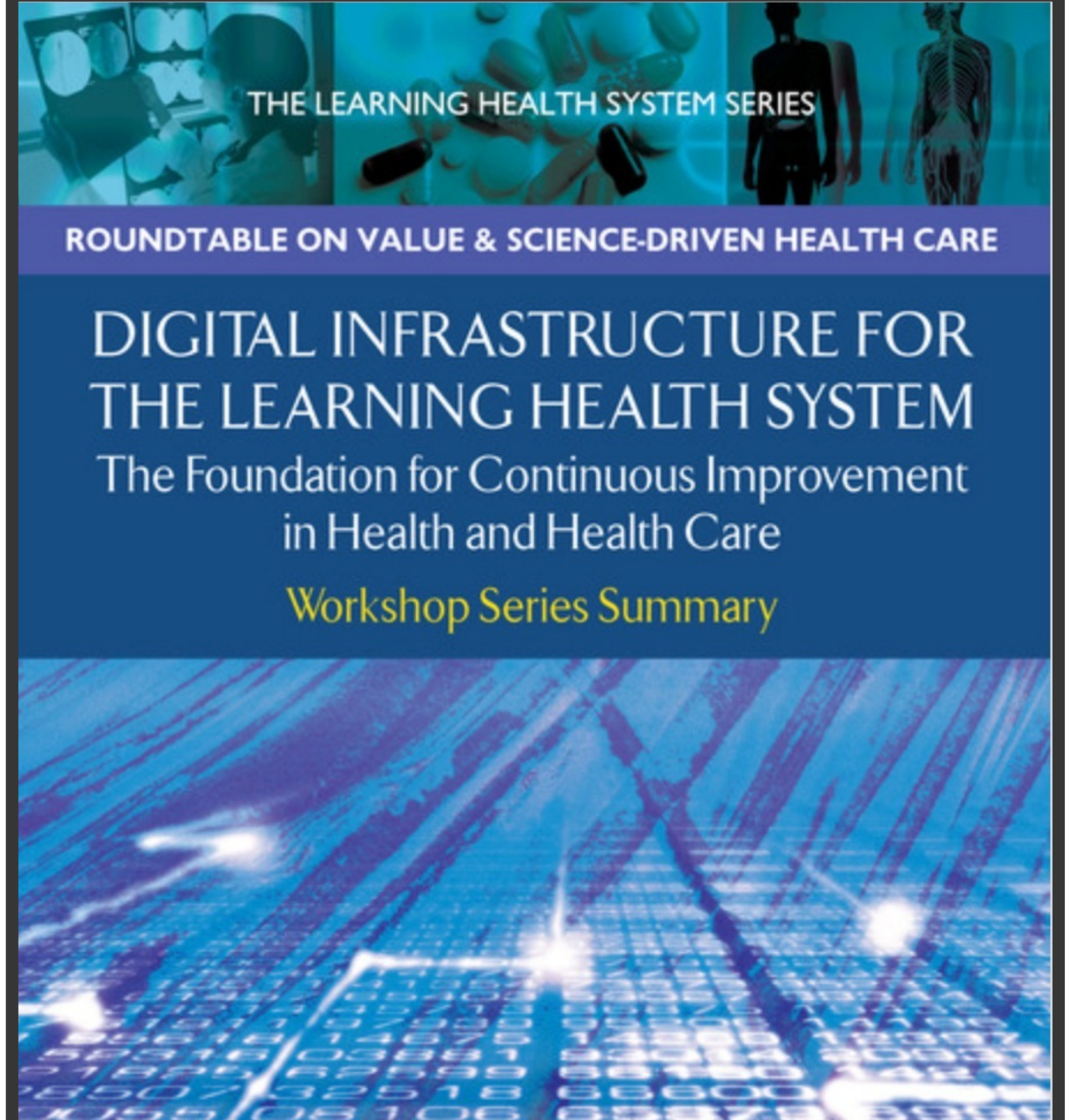
- (credit Rebecca Crowley Jacobson, VP, UPMC Enterprises)

The Future of Medicine

- Evidence based (data driven)
- Practice based (generation of data)
- Targeted and precise
  - Personalization to individual mutations
  - AI/ML to specific vectors/features
- A Learning Health System
  - "…gets the right care to people when they need it and then captures the results for improvement…" Institute of Medicine/National Academy of Medicine

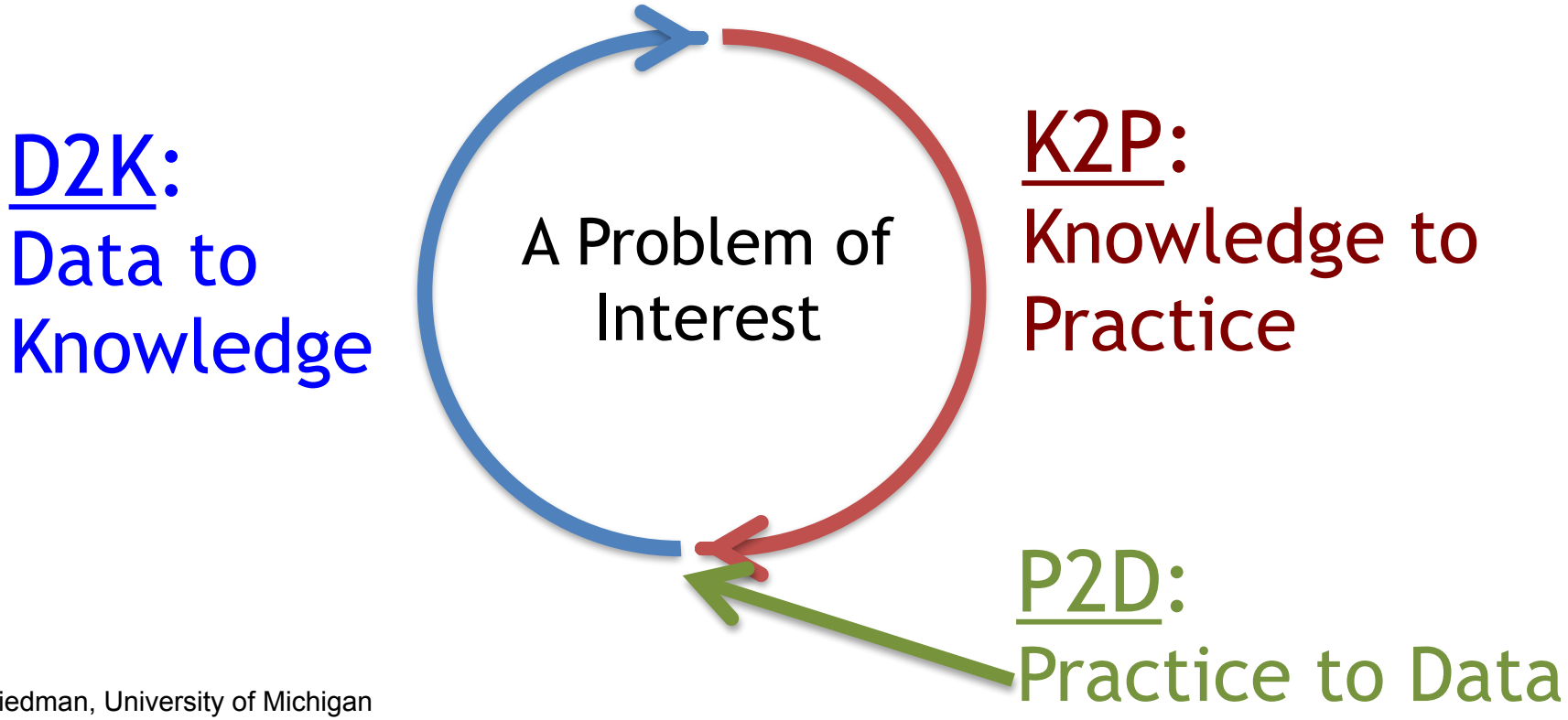
LEARNING HEALTH COMMUNITY

"We seek the development of **a learning health system** that is designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; **to drive the process of discovery as a natural outgrowth of patient care;** and to ensure innovation, quality, safety, and value in health care."



The LHS Links Discovery to Better Health

Better Health = [D2K] [K2P] [P2D]



Checklist View: Properties of a Health System That Can Learn

- ✓ Every patient's characteristics and experiences are available to learn from
- ✓ Best practice knowledge is immediately available to support decisions
- ✓ Improvement is continuous through ongoing study
- ✓ An infrastructure enables this to happen routinely and with economy of scale
- ✓ All of this is part of the culture

Core Values of the Learning Health System

1) Person Focused	2) Privacy
3) Inclusiveness	4) Transparency
5) Accessibility	6) Adaptability
7) Governance	Cooperative & Participatory 8) Leadership
9) Scientific Integrity	10) Value

https://lillypad.lilly.com/entry.php?e=8284

www.LearningHealth.org

116 Endorsements of the LHS *Core Values**

(As of 11/30/2017)

Program in Health Informatics, SONHP

Lewin and Associates LLC

Ursus Technologies (SDVOSB)

The Center for Learning Health Care

Veterans Health Administration
Office of Informatics & Analytics

Division of Health and Social Care Research

SecureHealthHub, LLC

Siemens Health Services

Kani Consulting Group LLC

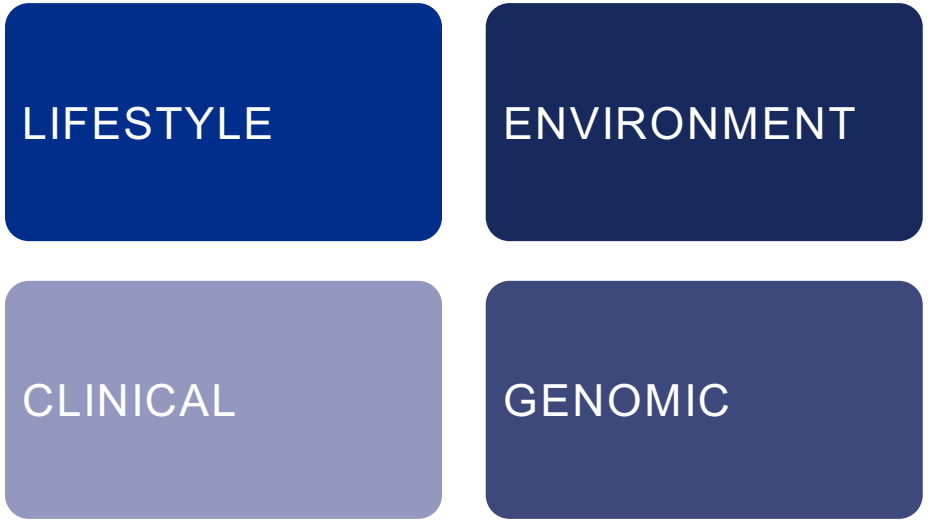
Department of Primary Care and Public Health

Medical Advocacy Mural Project

*To be included on the www.LearningHealth.org website.

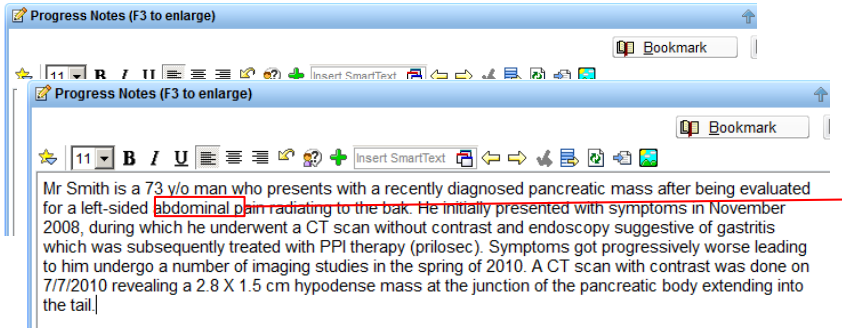
Learning Health Systems **at Scale**: Research Coupled with Impact

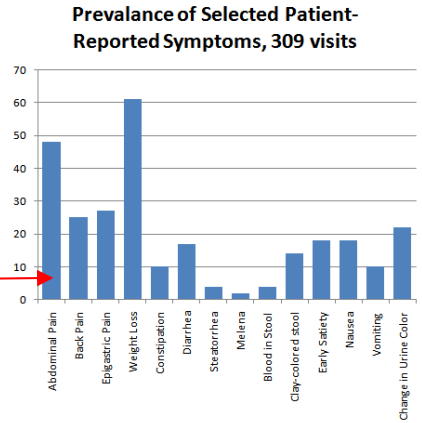
- Scalable expansion of data collection and use
- Collaboration over
  - Millions of patients with
  - New outcomes techniques
- Automated feature modeling
- High-performance analysis
  - clinical, cost, sensor (phone),
  - imaging, and omics
- Tight clinical integration



- From theory and experimentation to observation and simulation, toward a learning health system: "…enables discovery [and innovation] as a natural outgrowth of patient care…" – NAM (formerly IOM)
- Quality Improvement, Decision Support, Population Health, Cost, Safety, Hardening, Personalization and Commercialization

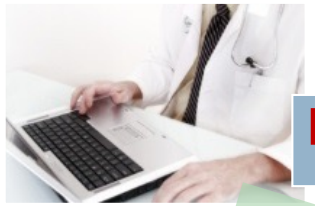
Structured Clinical Documentation: Statement of Problem



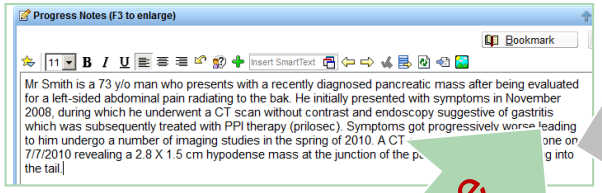


- Clinical documentation is a rich source of information on interactions between the health system and individual patients.
- Question: How can we capture this information **Consistently and Completely** for analysis—especially the interesting parts of progress notes?
- Answer: Tools Balance Expressivity and Workflow

Three Different Approaches



Free Text

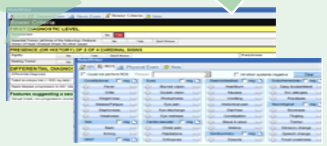


Parse and Abstract Data

Abstract Data

Generate

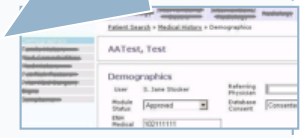
Enter Data



Structured Tools

Natural Language Processing

Manual Chart Review



Web Form

Database

Research Informatics Office (RIO)

- Mission
  - “to support investigators through innovative collection and use of biomedical data”
- Science-as-a-Service
  - Health Record Research Request (R3)
  - PaTH Network (PCORI CDRN)
  - NMVB, TCRN, Cancer Registry, PGRR
  - UPMC, Enterprises, IPM and PSC relationships
  - Delivery, Help Desk: REDCap, Neptune, ACT, AoU, etc.

Neptune
Research Data Warehouse

R3 Honest Broker Service

PaTH Network
patient empowered research

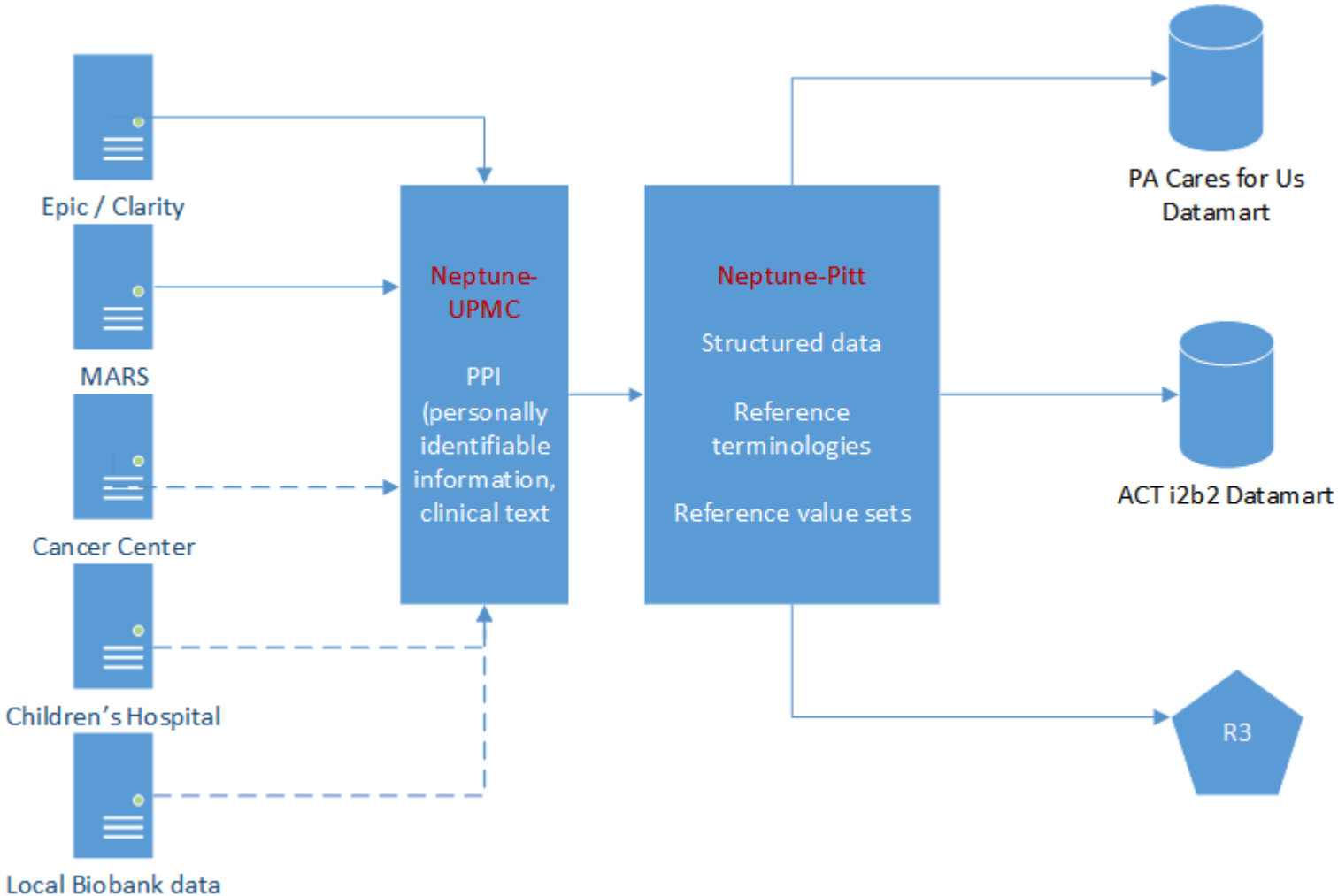
TCRN

All of Us℠
THE FUTURE OF HEALTH BEGINS WITH YOU
The Precision Medicine Initiative®

PA cares for Us
UNIVERSITY OF PITTSBURGH

Neptune

Research Data Warehouse



- Architecture
  - Atomic data warehouse
  - Footprint in both UPMC and Pitt
- Data Domains
  - Personally identifiable data (PPI), demographics
  - Encounters: outpatient, ED, inpatient
  - Diagnoses: billing, encounter based, problem list
  - Procedures: billing
  - Medications: orders/prescriptions, dispensing
  - Laboratory tests: orders, results
  - Social history: tobacco, alcohol
  - Vitals, allergies
  - Clinical text
- Terminologies & Value sets
  - Demographics (race, ethnicity, gender)
  - Encounter types
  - Diagnoses: ICD-9, ICD-10
  - Procedures: ICD-9, ICD-10, CPT-4, HCPCS
  - Medications: RxNorm, NDC
  - Laboratory tests: LOINC

- Period: January 2004 - November 2017
- Update frequency: monthly
- Patients: 6.35M
- Diagnoses: 190M
- Procedures: 91M
- Laboratory test results: 973M
- Medication orders: 62M

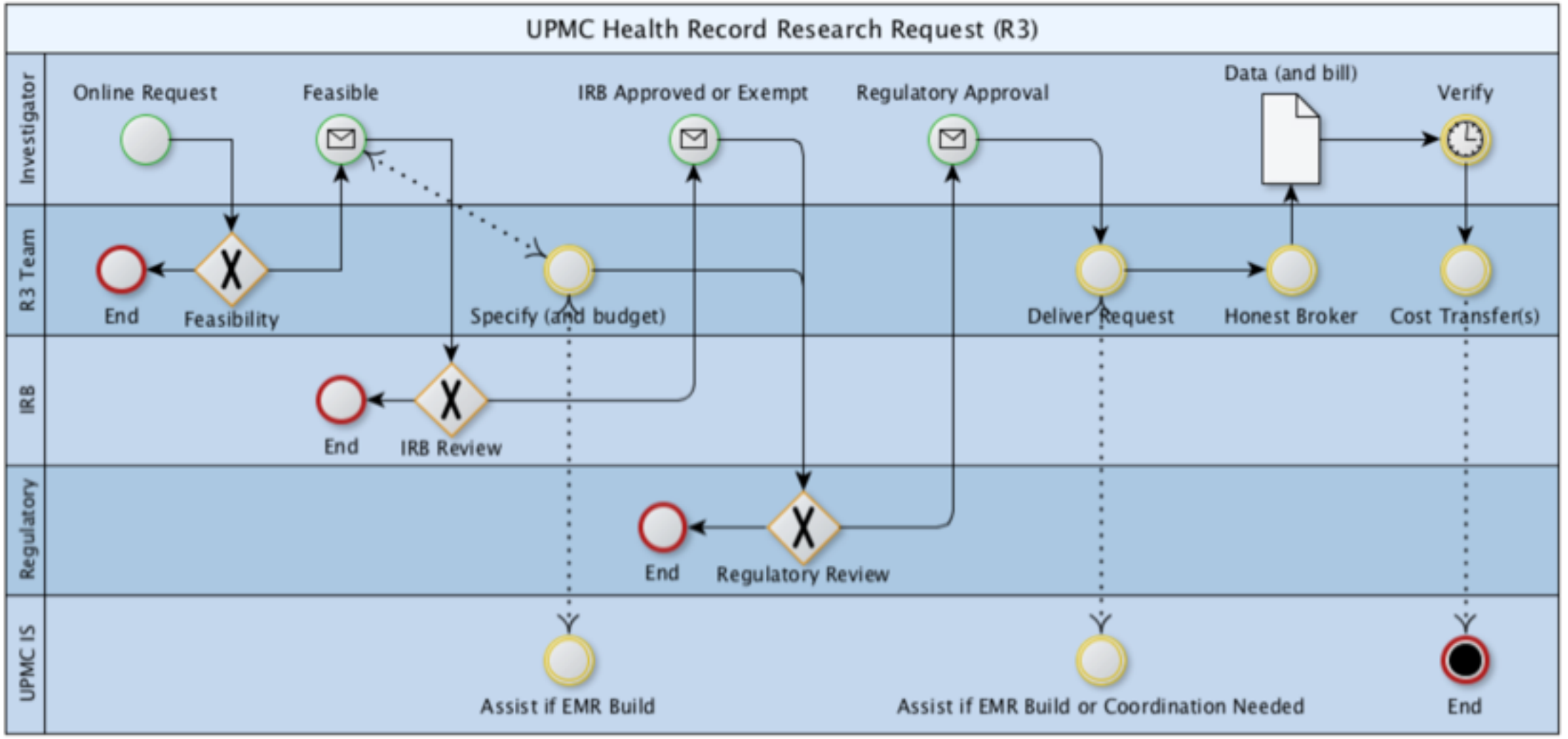
Health Record Research Request (R3)

- University and UPMC desire to make certain de-identified clinical data available…for research.
- Under CRIO, on behalf of UPMC, certain DBMI staff operate via HIPAA BAA as Honest Broker
- R3 is this service or process of provisioning data through Neptune and of authorizing additional sources

University of Pittsburgh
Institutional Review Board

Honest Broker Certification

R3 Workflow


UPMC Health Record Research Request (R3)
Investigator
R3 Team
IRB
Regulatory
UPMC IS
Online Request
Feasible
IRB Approved or Exempt
Regulatory Approval
Data (and bill)
Verify
End
Feasibility
Specify (and budget)
Deliver Request
Honest Broker
Cost Transfer(s)
End
IRB Review
End
Regulatory Review
Assist if EMR Build
Assist if EMR Build or Coordination Needed
End

At its core, TIES is a natural language processing (NLP) pipeline and clinical document search engine. The software de-identifies, annotates, and indexes your clinical documents, making it easier for your researchers to search for and find the documents and cases. TIES also supports tissue ordering and acquisition and integration with tissue banks and honest brokers. It also works across institutions with separate TIES installations as the TCRN.

What is the TCRN?

The TIES Cancer Research Network (TCRN) is a federated tissue and data sharing network that enables researchers to identify cases, access data and request tissue/materials from one or more network partners.

Rebecca Jacobson, Michael Becich, et al, University of Pittsburgh/UPMC

Authority, responsibility, effectiveness

- National Algorithm Safety Board – Ben Shneiderman
  - Was the “right” data used to train? How does it perform? Can AI’s be responsible (who has the liability? the builder, maintainer, implementer?)

- A “Fundamental Theorem” of Biomedical Informatics (Friedman)

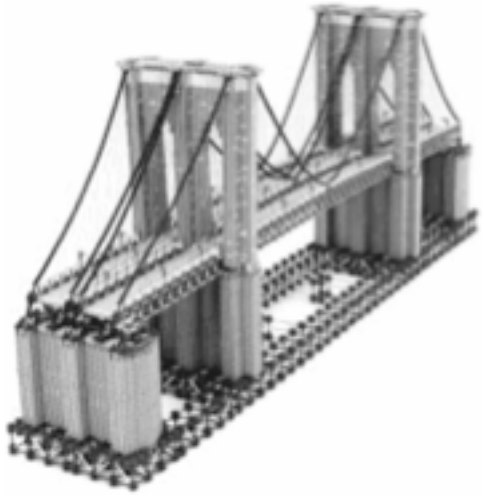
What can we expect?

- “Personalized medicine” or “Precision medicine” will play an increasing role in healthcare, particularly in cancer
- Genome sequencing will become increasingly common; integral to patient care
- Correlation of molecular changes to phenotype will be critical for both research, and also for selecting therapy for patients

- **New technology approaches are required for adaption and scale**

Important characteristics

- We must integrate systems that may not have worked together before
- These are human systems, with differing goals, incentives, capabilities
- All components are dynamic—change is the norm, not the exception
- Processes are evolving rapidly too



We are not building something simple like a bridge or an airline reservation system

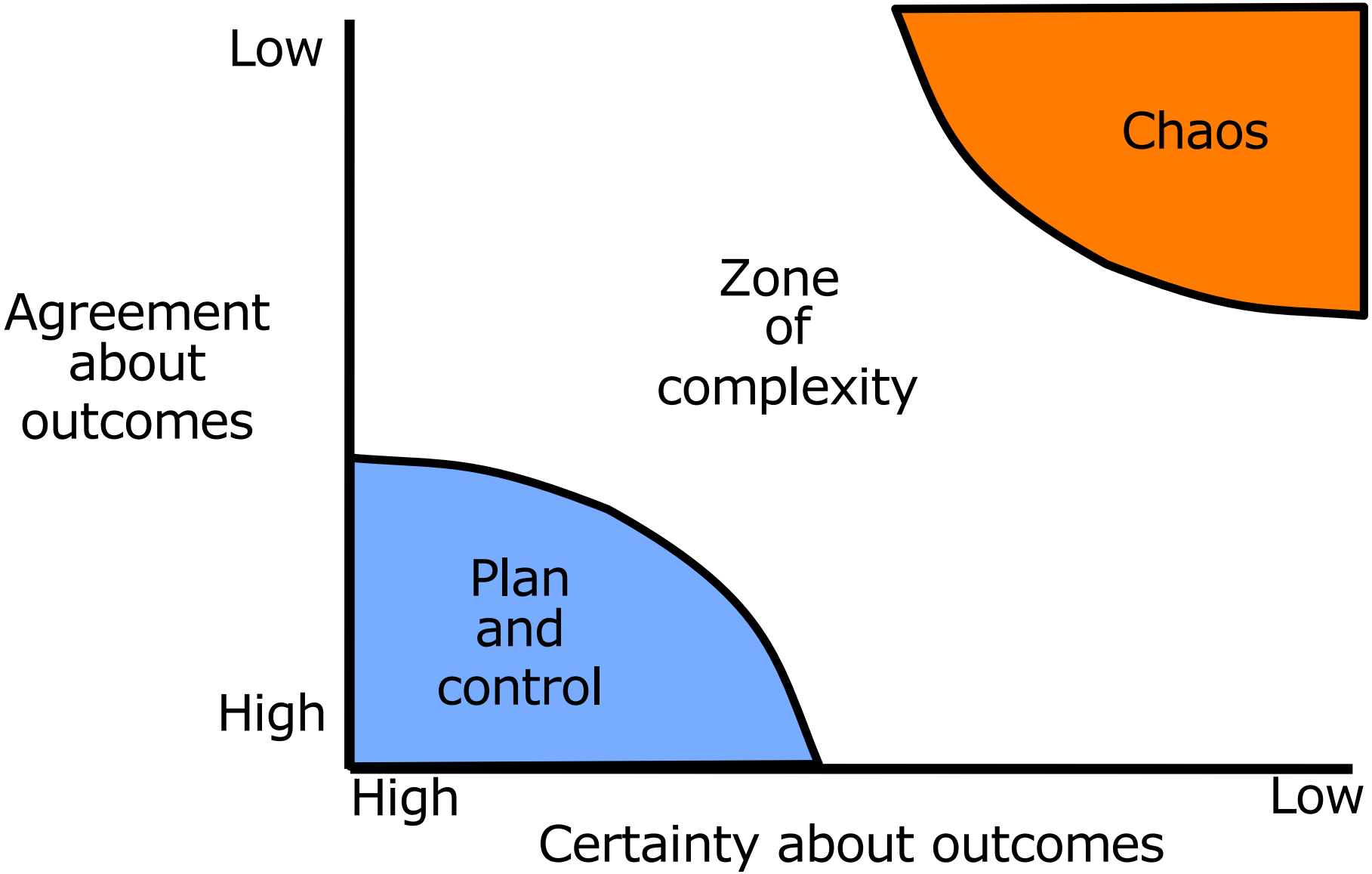
Healthcare is a **complex adaptive system**

A complex adaptive system is a collection of individual agents that have the freedom to act in ways that are not always predictable and whose actions are interconnected such that one agent's actions changes the context for other agents.

Crossing the Quality Chasm, IOM, 2001; pp 312-13

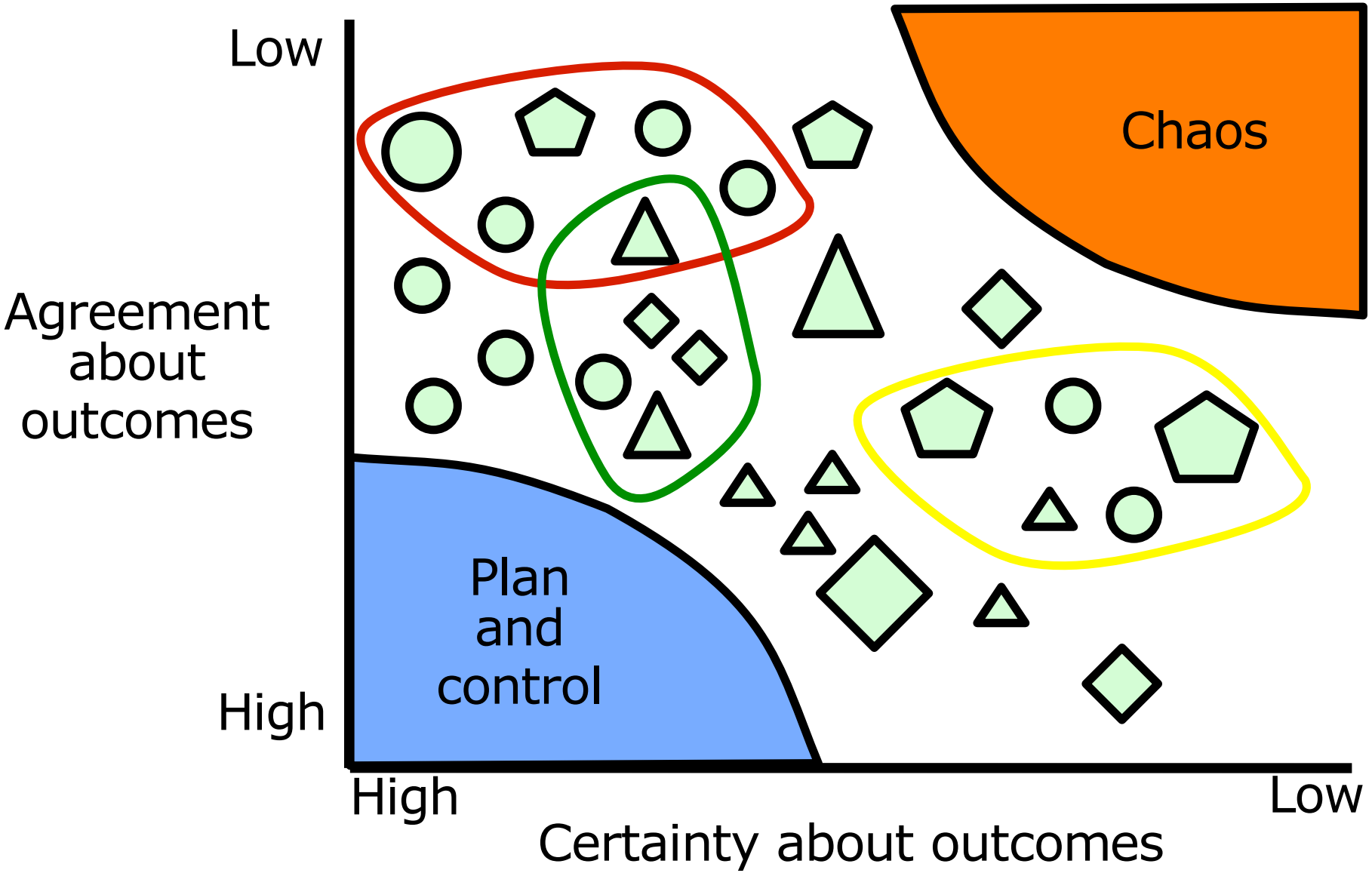
- Non-linear and dynamic
- Agents are independent and intelligent
- Goals and behaviors often in conflict
- Self-organization through adaptation and learning
- No single point(s) of control
- Hierarchical decomposition has limited value

We need to function in the **zone of complexity**



Ralph Stacey, *Complexity and Creativity in Organizations*, 1996

We need to function in the **zone of complexity**



Ralph Stacey, *Complexity and Creativity in Organizations*, 1996

We call these groupings **virtual organizations** (VOs)

A set of individuals and/or institutions engaged in the controlled sharing of resources in pursuit of a common goal

Healthcare = dynamic, overlapping VOs, linking

Patient – primary care

Sub-specialist – hospital

Pharmacy – ?

But U.S. health system is marked by fragmented and inefficient VOs with insufficient mechanisms for controlled sharing

Foster I, Kesselman C, Tuecke S. The anatomy of the grid: Enabling scalable virtual organizations. Int J High Perform Comput Appl. 2001;15(3):200–222.

Service Oriented Science

- **New information architectures enable new approaches to publishing and accessing valuable data and programs.** So-called service-oriented architectures define standard interfaces and protocols that allow developers to encapsulate information tools as services that clients can access without knowledge of, or control over, their internal workings. Thus, **tools formerly accessible only to the specialist can be made available to all**; previously manual data-processing and analysis tasks can be automated by having services access services. Such service-oriented approaches to science are already being applied successfully, in some cases at substantial scales, but much **more effort is required before these approaches are applied routinely across many disciplines.** Grid technologies can accelerate the development and adoption of service-oriented science by **enabling a separation of concerns between discipline-specific content and domain-independent software and hardware infrastructure.**

Foster I. Service-Oriented Science. Science. AAAS; 2005;308(5723):814.

We need hosted federation of services

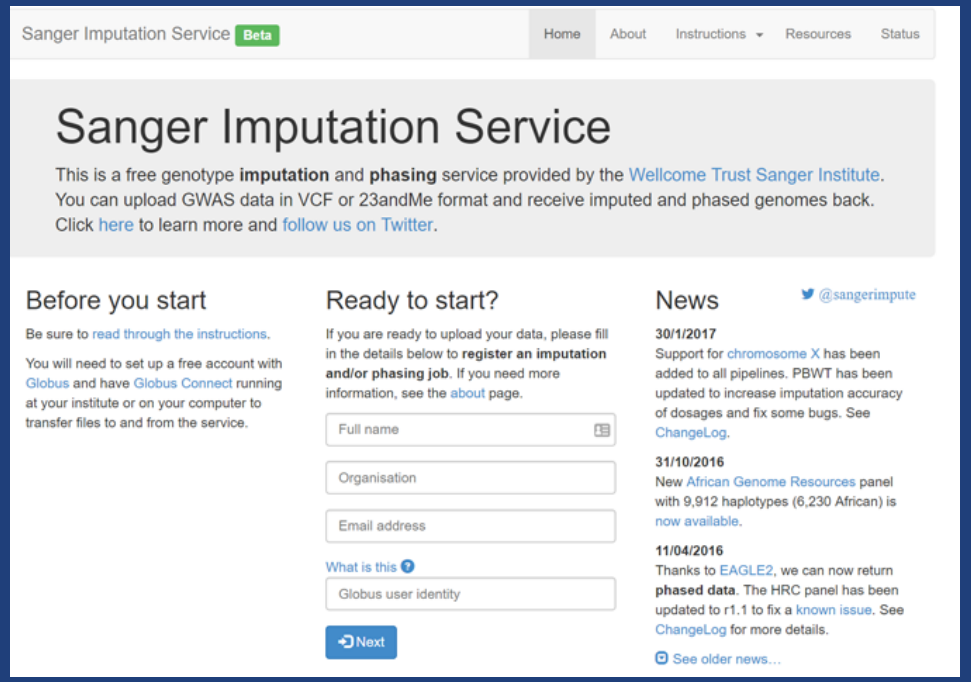
- Attribute-based authorization.
- Distributed identity management.
- End-to-end security.
- Data naming, linking, movement, and integration.
- Flexible, but enforceable policy/sociability.
- Extensibility.
- Redundancy.
- Robust in multiple industries/stability.
- Without central ownership/manageability.

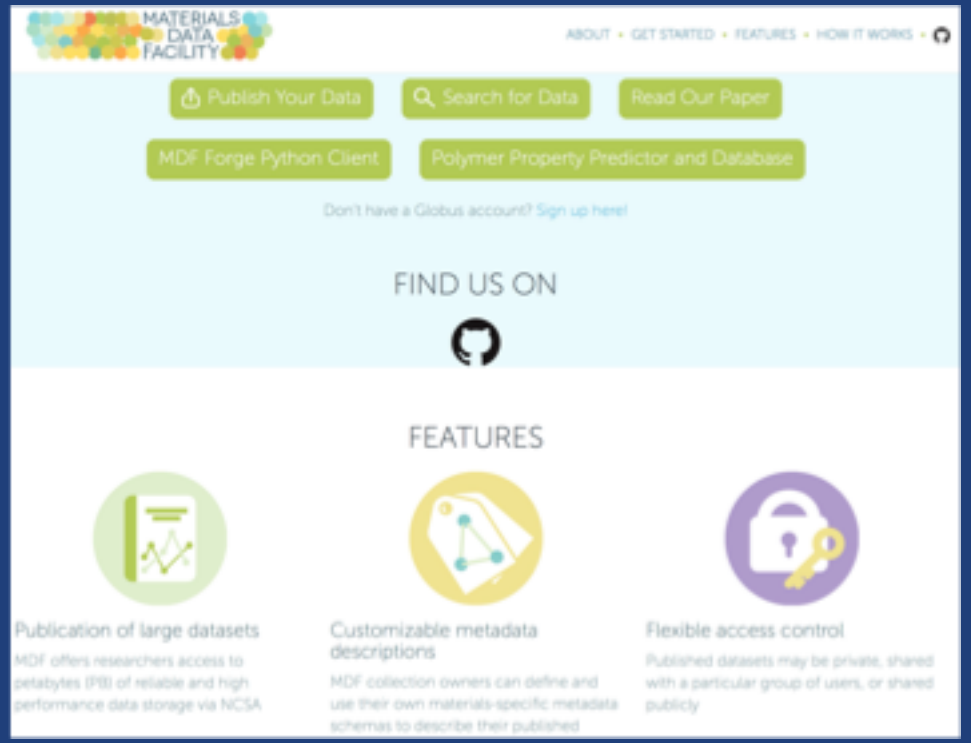
Globus as solution

- Solves issue with third party access to private data
- Complement to other software/systems
- Easy to streamline and scale
- Useful for teams with distributed resources and agents
- Distribution of big data

Many variants possible

- **Manage access to data at multiple locations**
- **Manage access to data on cloud**
- **Upload data for analysis**
- **Data download from scientific instruments**
- **Data publication**
- **Transfer data to computer for analysis**





A key message: Outsource all that you can

- **Outsource responsibility for determining user identities**
- **Outsource control over who can access different data and services within the portal**
- **Outsource responsibility for managing data uploads and downloads between various locations and storage systems**
- **Leverage standard web user interfaces for common user actions**


Securely manage protected data
HIPAA (with BAA)
PII
PHI
CONTROLLED UNCLASSIFIED

Future: Data, Algorithms, Intelligence, oh my!

- All data is attributed (may be private, but not anonymous).
- Algorithms (machine “partners”) have responsible humans behind them.
- Intelligence is manifest as complex adaptive socio-technical learning collectives including machine “partners”.

- BIG QUESTIONS:
- In the Future: will we want to know we’re interacting with an AI (we seem to want to now, but we also want efficiency…)?
- In the Future: will we value or even tolerate anonymity (e.g. blockchain)?