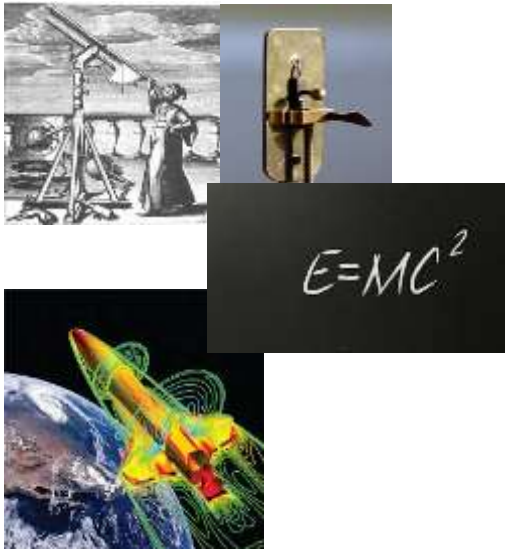
Ken Buetow, Ph.D

Director, Computation Science and Informatics,
Complex Adaptive Systems @ ASU
Professor, School of Life Science

Kenneth.Buetow@ASU .edu

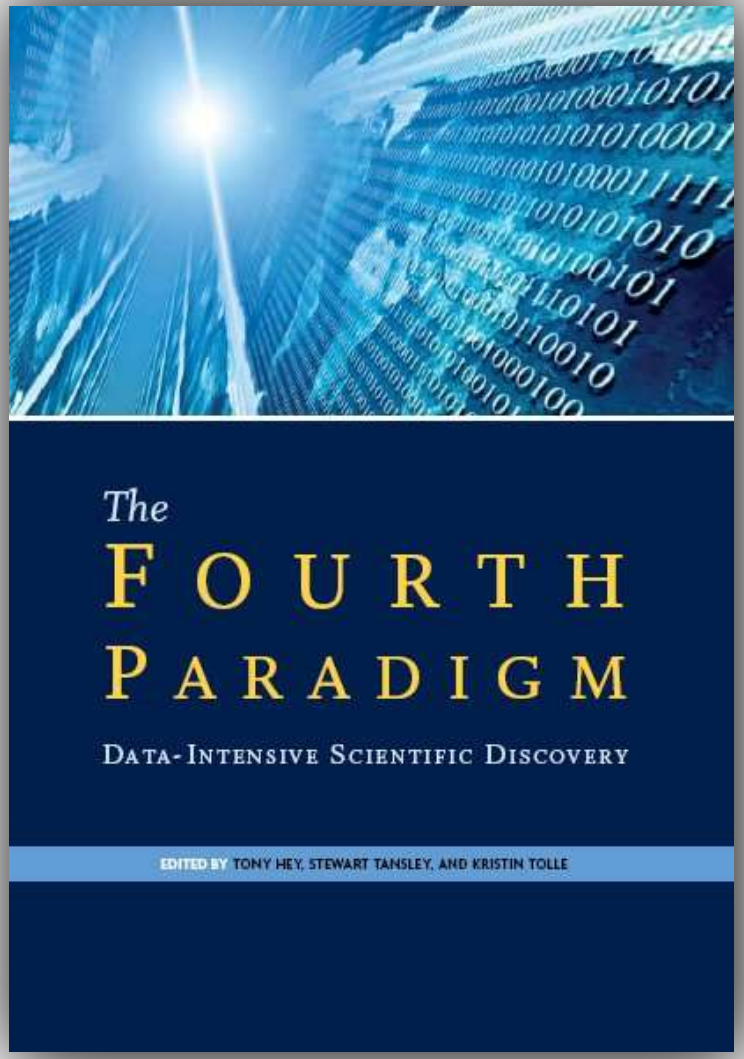
4th Paradigm Science



- 1st – empirical
  - Describing natural phenomena
- 2nd – theoretical
  - Using models and generalizations
- 3rd – computational
  - Simulating complex phenomena
- **4th – data-intensive science**
  - **Data generated by instruments or simulation**
  - **Processed by software**
  - **Information/knowledge stored in computer**
  - **Scientist analyzes using data management/statistical/mathematical/algorithmic approaches**

4th Paradigm Science

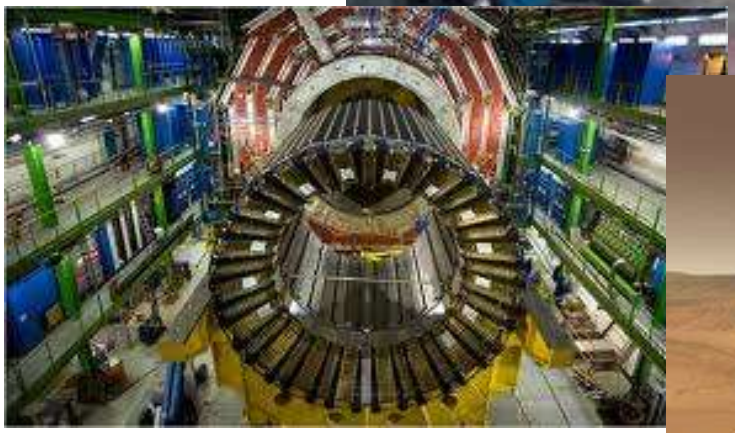
- A new method of pushing forward the frontiers of knowledge, enabled by new technologies for gathering, manipulating, analyzing and displaying data.
- Complementing **data-generating** science with **data-driven** science
- Ecumenical
  - Astronomy
  - Physics
  - Economics
  - Climate
  - Genomics
- Transdisciplinary



Creating a *new* Data Science Instrument:

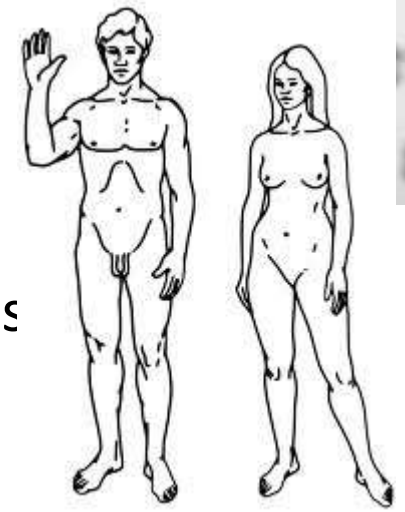
A Next Generation Cyber Capability (NGCC)

Traditional “Data-Generating” Scientific “Instruments”

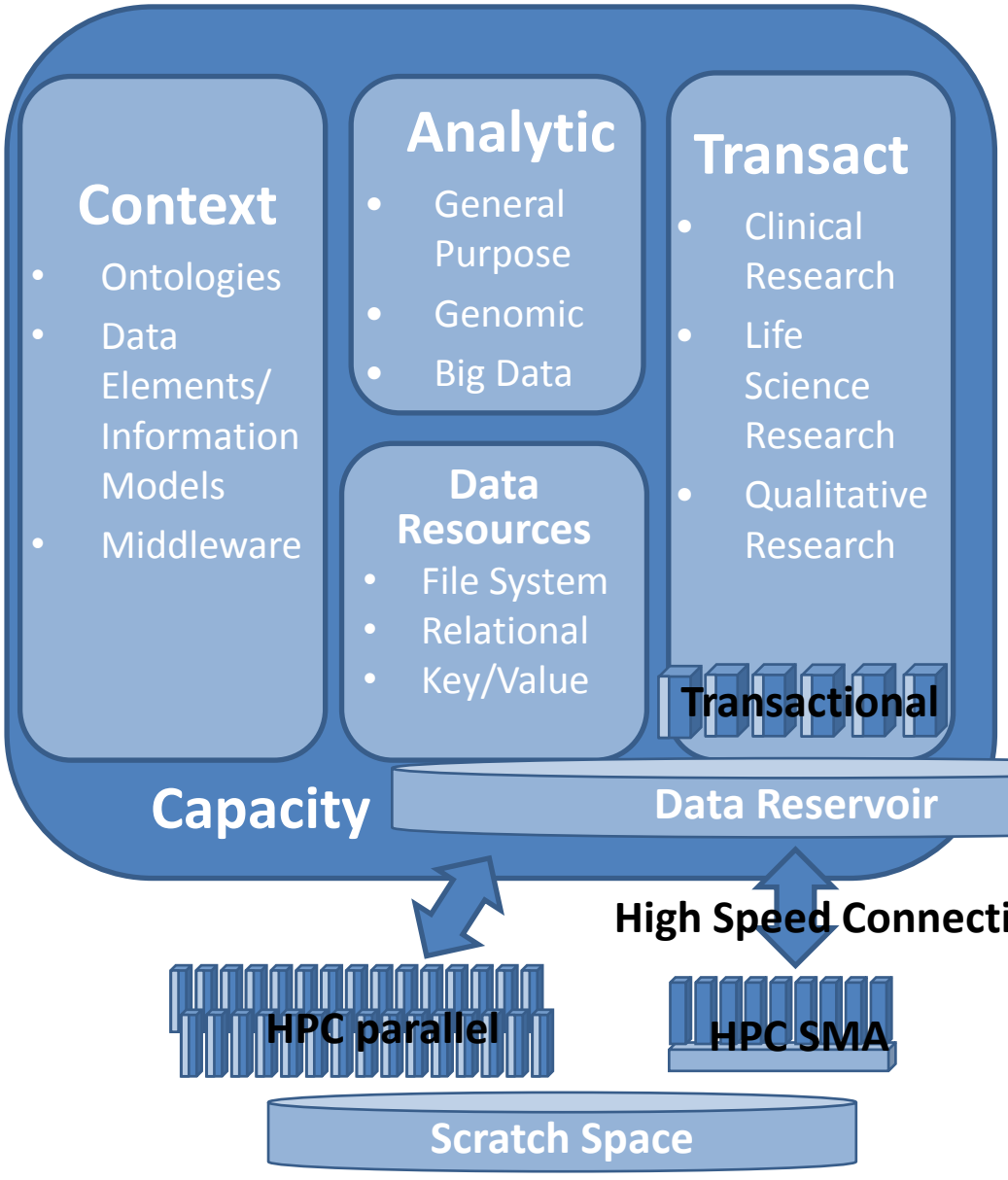


The NGCC **Data Science** "Instrument"- *an elemental whole composed of:*

- Physical Capacity
  - Ultra-high bandwidth Networks
  - Large-scale storage
  - Mulitple "flavors" of computation
- Logical Capabilities
  - Software
  - Metadata
  - Semantics
- Human Resources
  - Transdisciplinary Teams



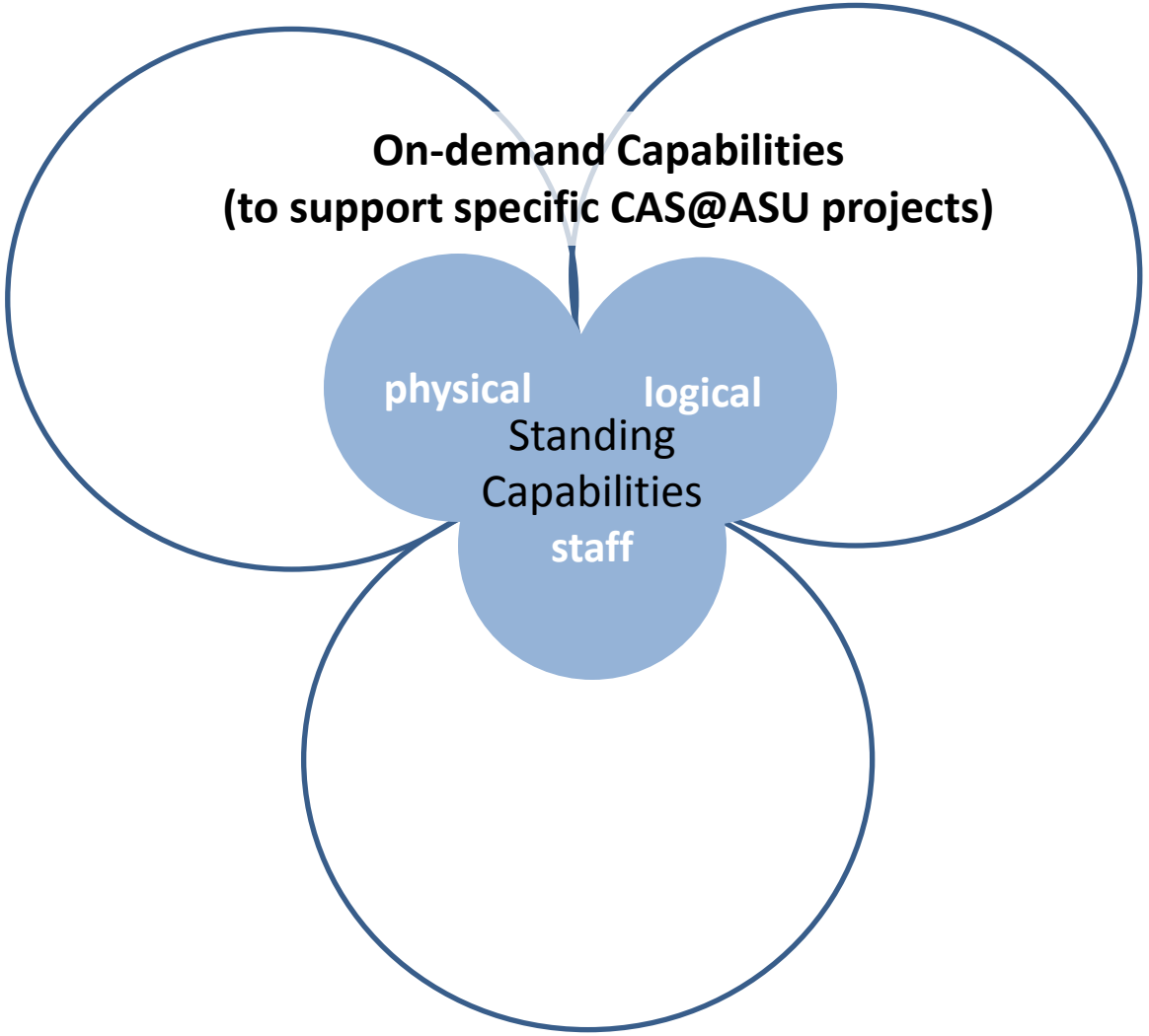
NGCC Data Science "Instrument"



Content

- CRF
- Instrument
- EHR
- Document
- Filing
- Practice Guidelines
- Physician Experience
- Web site
- Wiki
- Media
- Social media

NGCC “Instrument” – Elastic Capabilities

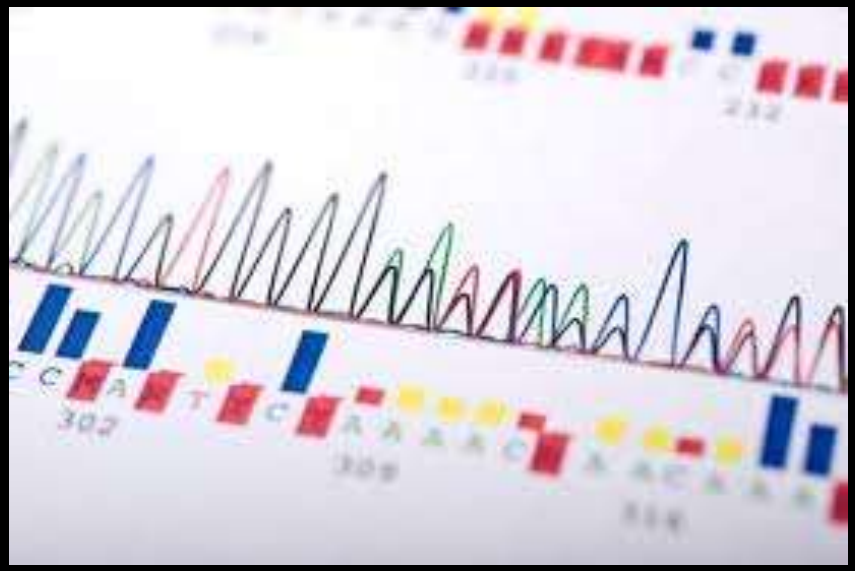


In Silico Medicine

An example

Personalized Medicine

Personalized Medicine

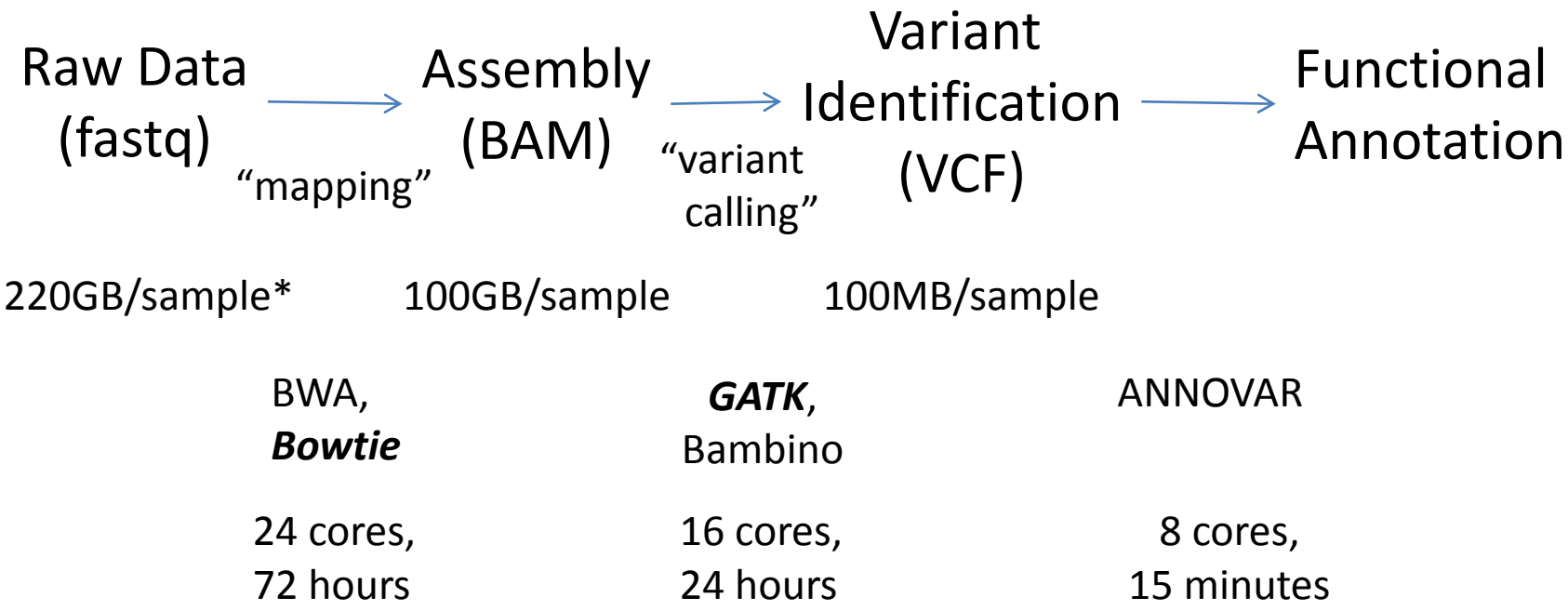


…Revolution

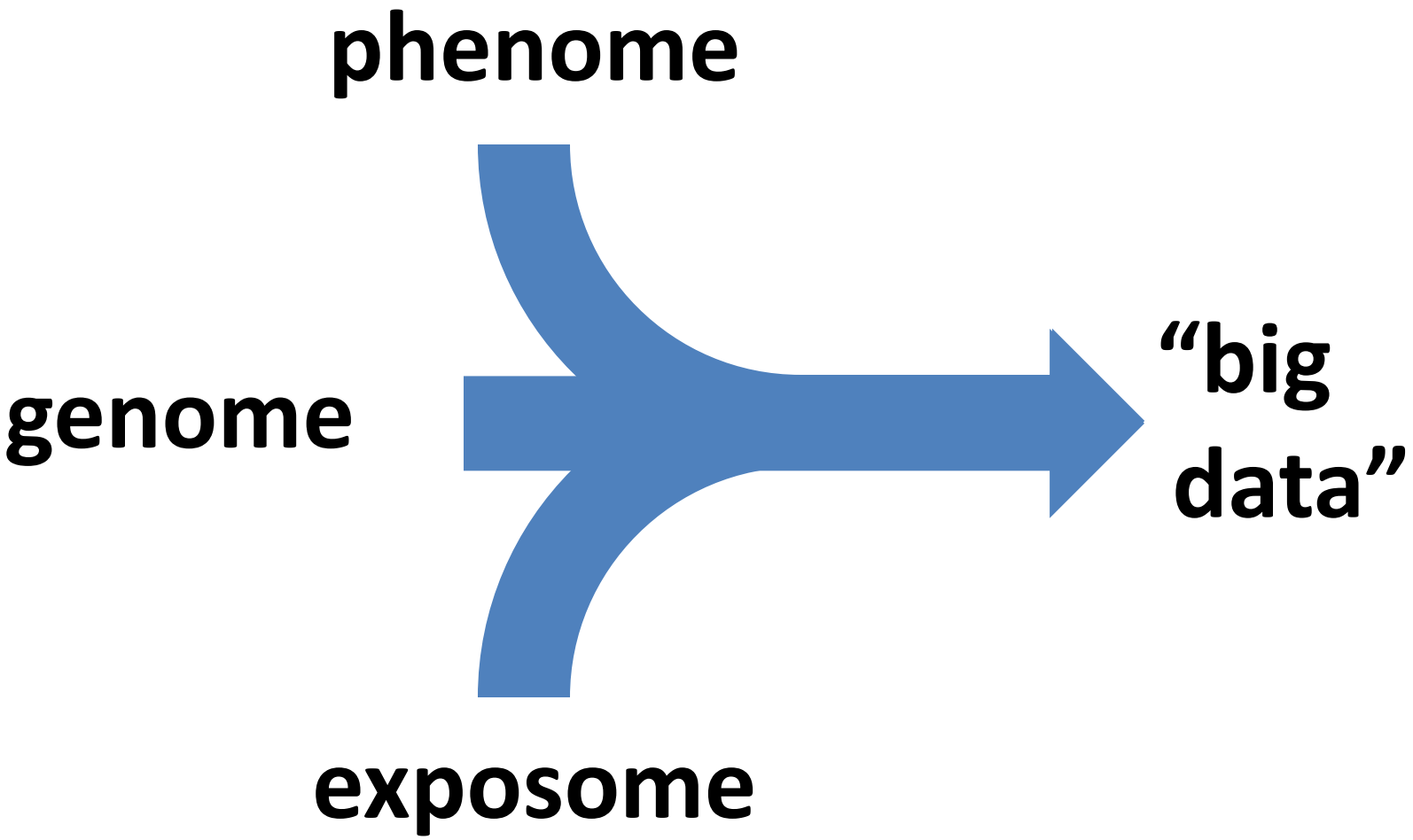
Personalized Medicine

…Renaissance

Whole Genome Analysis

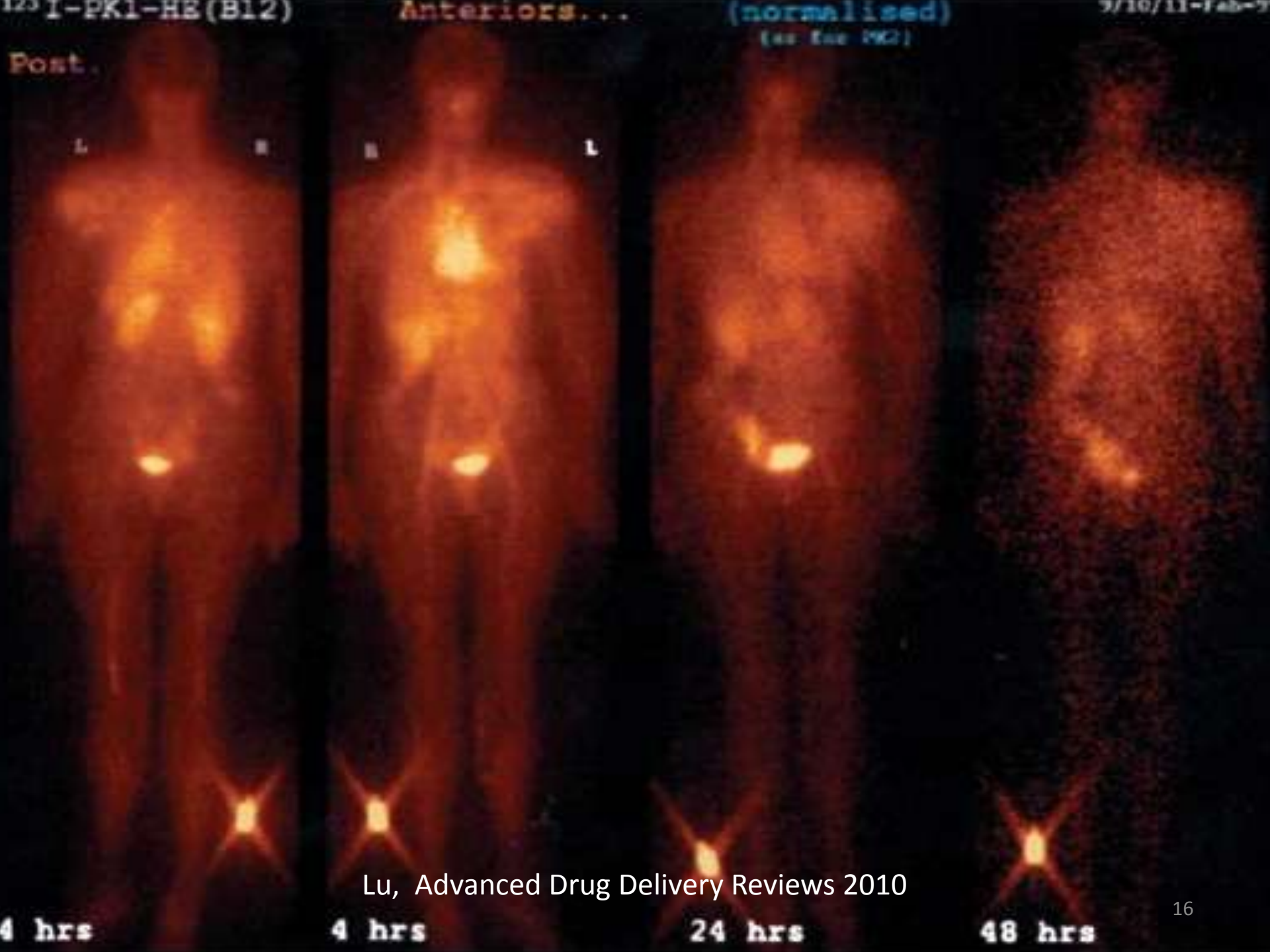


* Assuming 30X coverage


phenome
genome
exposome
“big data”

Phenome Data

- Diverse types
  - Clinical Observation
  - Clinical Laboratory
  - Imaging
  - Registry
  - Biospecimens
  - Reference
- Distributed sources
  - Research Center
  - Care Delivery Setting
    - Hospital
    - Practice
    - Laboratory
  - Registry
  - Industry
  - Consumer



Lu, Advanced Drug Delivery Reviews 2010

Real time consumer data

cardiovascular

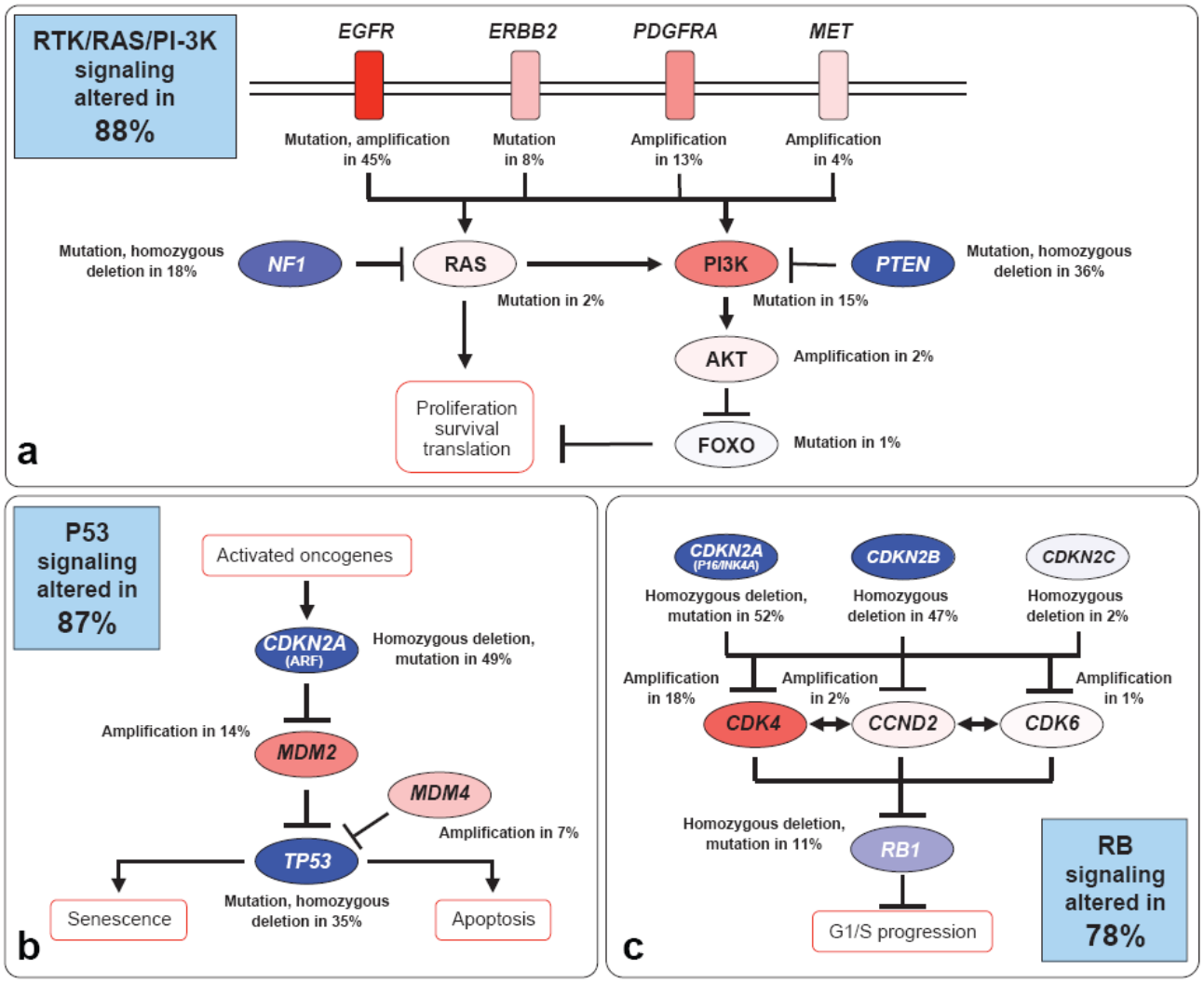
diabetes

fitness

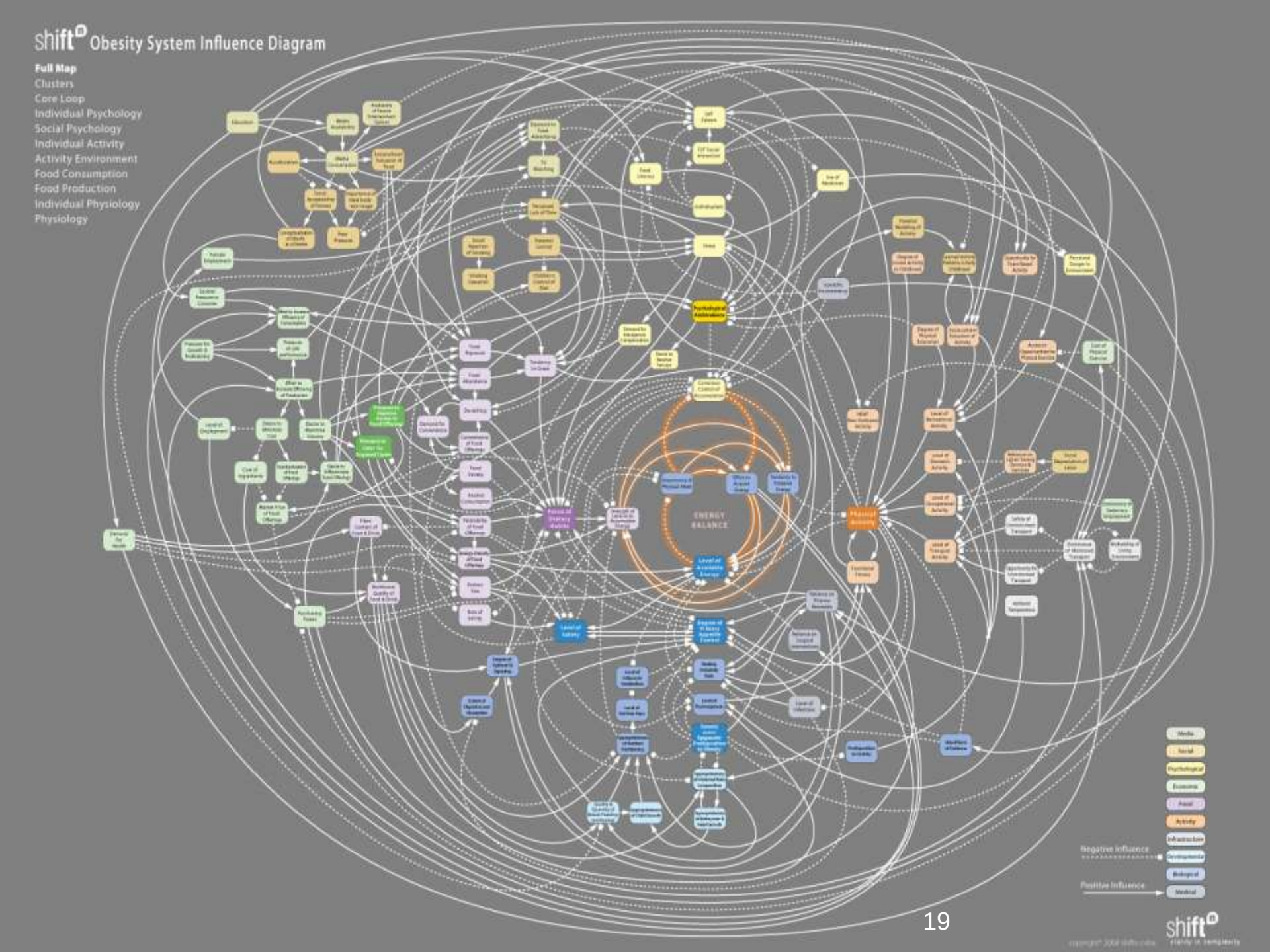
***shipments* of telehealth devices grow to about 2 million by 2013**

Source: http://mobihealthnews.com

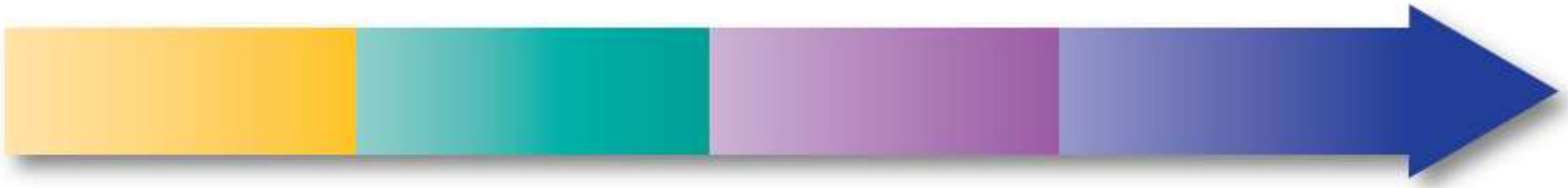
It takes a ***network…***



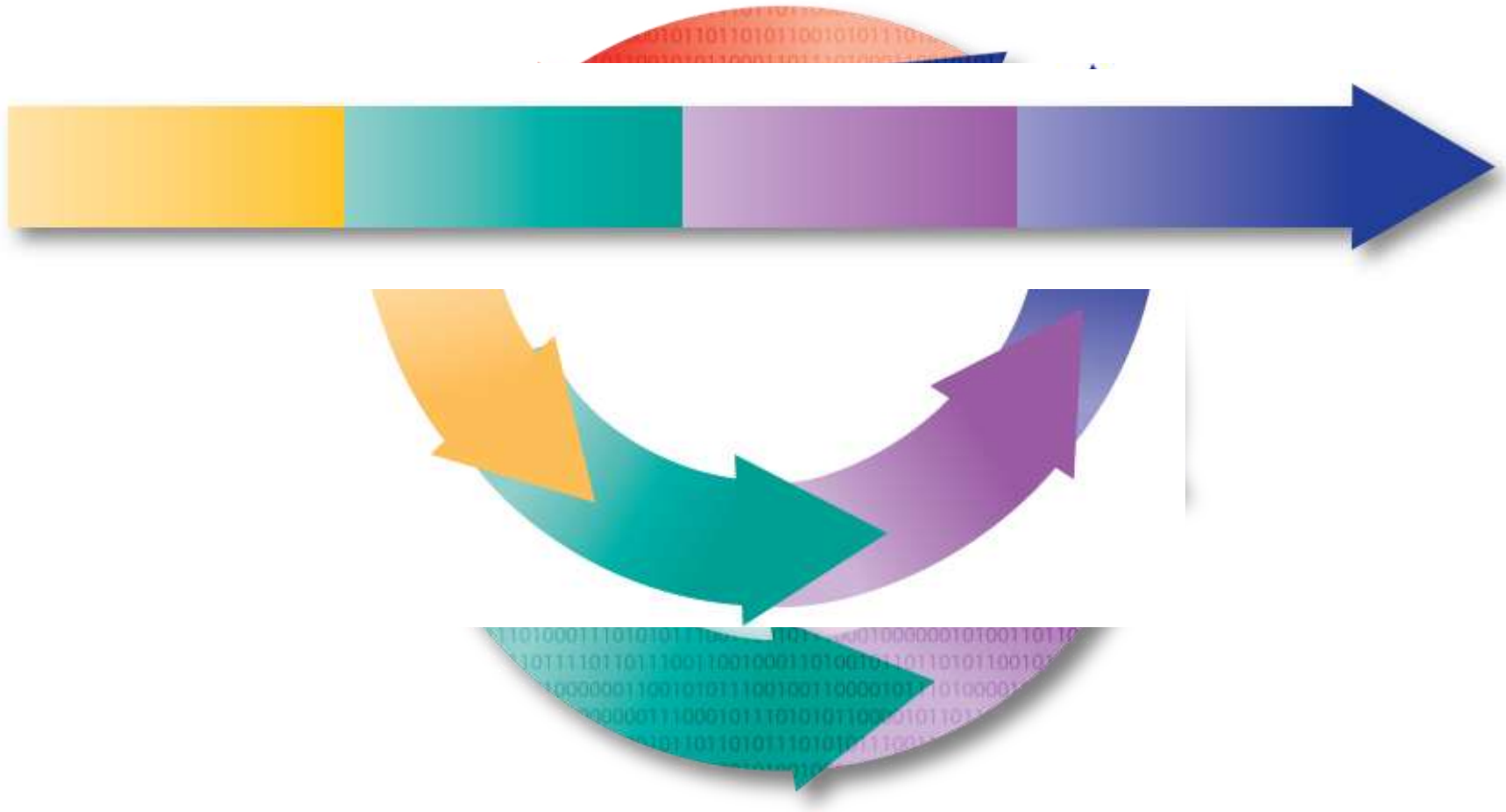
TCGA: *Nature* 2008


shift Obesity System Influence Diagram
Full Map
Clusters
Core Loop
Individual Psychology
Social Psychology
Individual Activity
Activity Environment
Food Consumption
Food Production
Individual Physiology
Physiology
ENERGY BALANCE
Negative Influence
Positive Influence
Media
Social
Psychological
Economic
Food
Activity
Infrastructure
Environmental
Biological
Medical

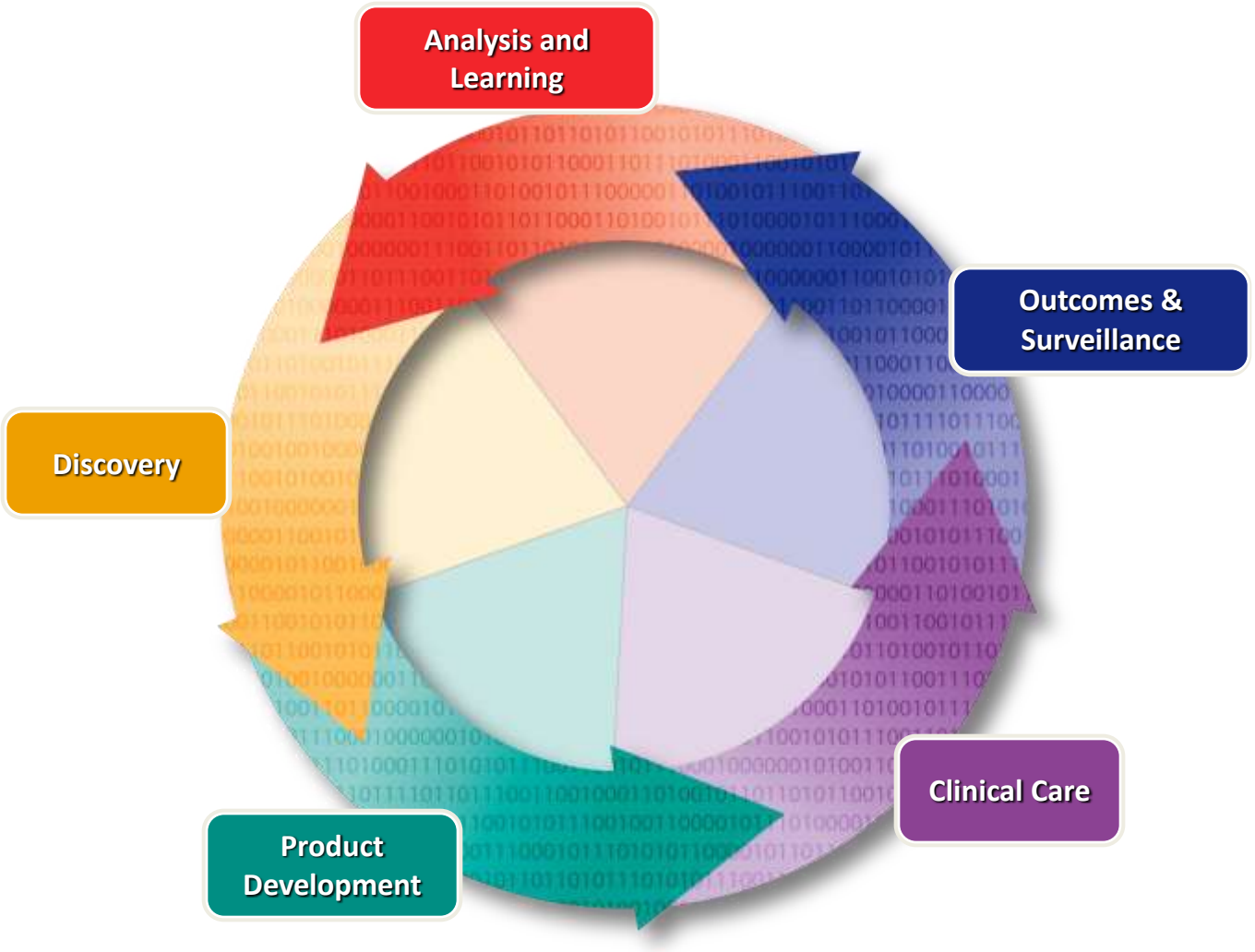
20th Century Science Paradigm



Discovery	Product Development	Clinical Care	Outcomes & Surveillance
• Biological pathways • Target identification and validation	• Candidate selection and Optimization • Pre-clinical testing • Phase I, II, III • New Drug application and Approval	• Product launch • Clinical adoption	• Reporting of serious/fatal ADRs • Re-labeling (or recall) as needed • Additional indications as warranted



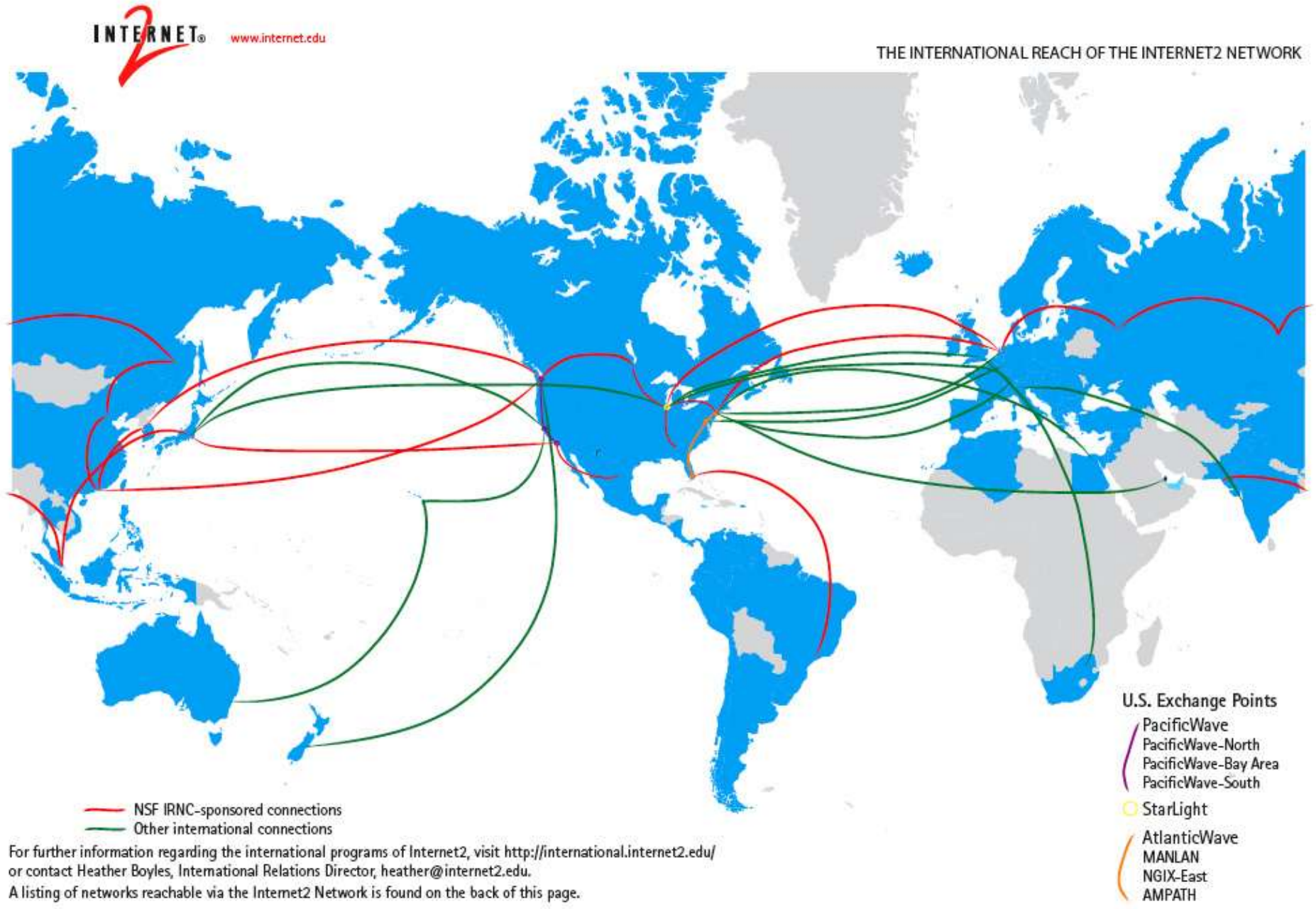
21st Century Data Science Paradigm



Seeking Collaborators and Partners

Come Join the Effort!

Kenneth.Buetow@ASU .edu


INTERNET2
www.internet.edu
THE INTERNATIONAL REACH OF THE INTERNET2 NETWORK
U.S. Exchange Points
PacificWave
PacificWave-North
PacificWave-Bay Area
PacificWave-South
StarLight
AtlanticWave
MANLAN
NGIX-East
AMPATH
NSF IRNC-sponsored connections
Other international connections
For further information regarding the international programs of Internet2, visit http://international.internet2.edu/
or contact Heather Boyles, International Relations Director, heather@internet2.edu.
A listing of networks reachable via the Internet2 Network is found on the back of this page.

Cancer: a Complex Adaptive System

"the whole is more than the sum of the parts"

- **Evolves** over time
- Has **adaptive behaviors** (dynamic as opposed to static)
- Displays **emergent properties** (or unintended consequences)
- Requires Transdisciplinary Study to understand: multidimensional, interacting ***"ecosystem"***
  - Biology, Chemistry, Medicine, Business, *Sociology, Anthropology*
  - Physicians, Nurses, Social Workers, Regulators, Researchers, *Payors, Consumers, Public Health Officials*
  - Industry, Academe, Government, *NGOs*
- Interdependencies
  - Resources
  - ***Information***

NGCC “Instrument” Proof of Concept: utilizing heterogeneous cancer data

Sample Research Questions

1. **For tumor samples, which are from patients who had progressive disease, retrieve all the genes which are in a high-amplification chromosomal region along with a given gene.**
2. **Retrieve the gene expression scores of all the genes that share a high amplification region with a known high-expression probe, along with the information about the sex, chemotherapy treatments and lifespan of the source patient.**
3. **Retrieve the set of genes, which have a documented association with one or more drugs, that are expressed at or over a specified threshold, but are not in a known high-amplification region**
4. **For all genes known to be in high amplification regions, identify those with high expression vs. methylation**

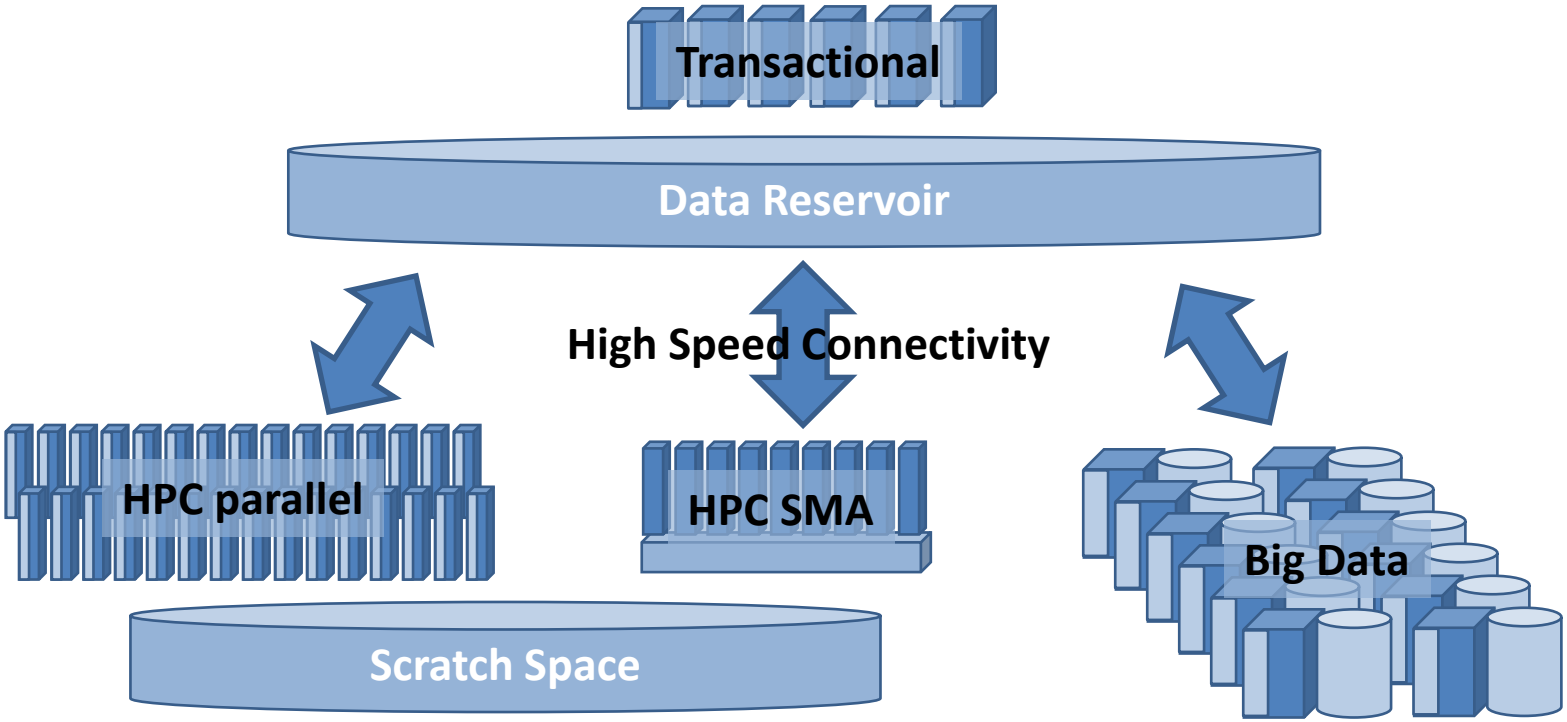
NCI CBIIT/Booz Allen Hamilton collaboration

NGCC “Instrument”: The Future

Potential Questions

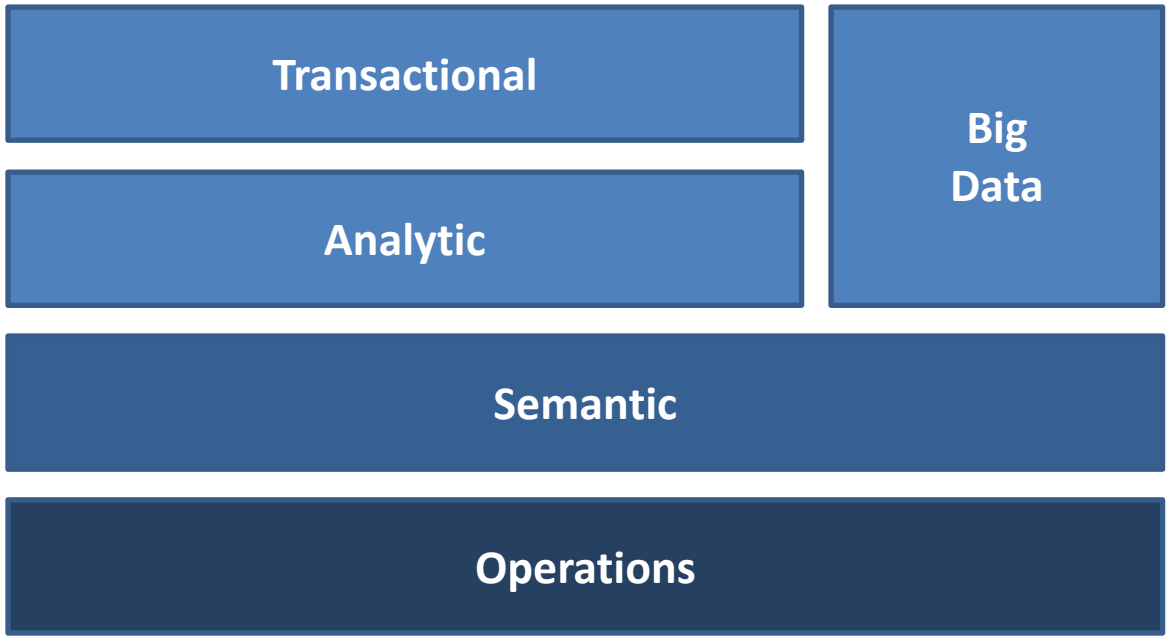
- **Given the clinical, demographic, and life style characteristics of this patient, how does survival differ among alternative interventions?**
- **What are the adverse events associated with alternatives?**
- **What are the quality of life differences associated with the alternatives?**
- **Given all the above, what intervention “might be right for me”**

NGCC “Instrument”: Physical Infrastructure



One size does ***NOT*** fit all…

NGCC “Instrument” Logical Infrastructure



NGCC Data Science "Instrument": Human Capabilities

- ***Physical Infrastructure***
  - Architects and Engineers
- ***Logical Infrastructure***
  - Engineer, Developers, and Administrators, Web Developer, Help Desk Staff
  - Analysts and Students
  - Knowledge Architects and Engineers
- ***Enterprise***
  - Managers and Administrators