NCI-DOE Collaboration 2020 Ideas Lab: Toward Building a Cancer Patient "Digital Twin"

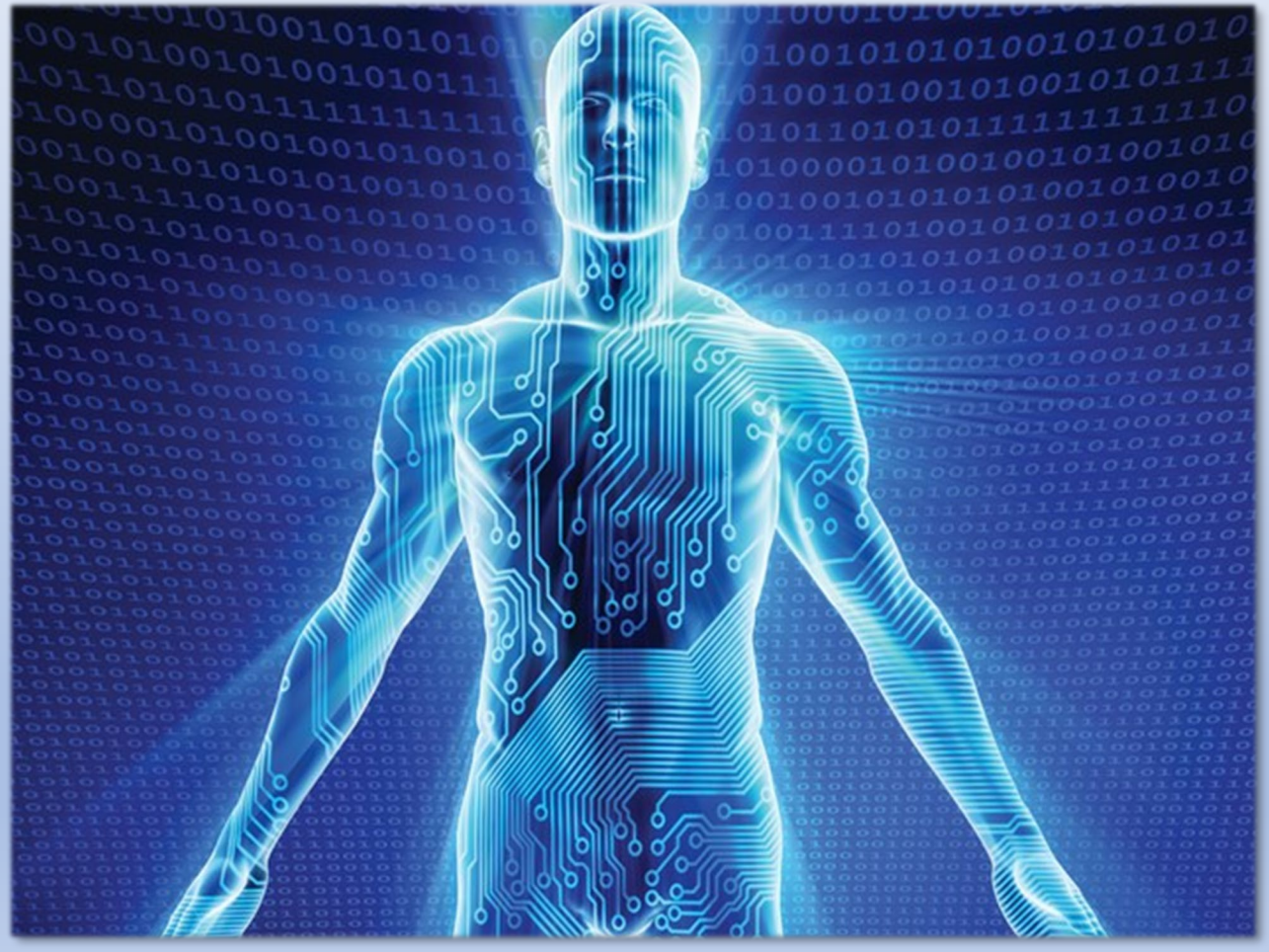
Call to Action

Michael Cooke, Office of Science, U.S. Department of Energy
Emily Greenspan, National Cancer Institute
Eric Stahlberg, Frederick National Laboratory for Cancer Research

July 6-10, 2020
Virtual Meeting

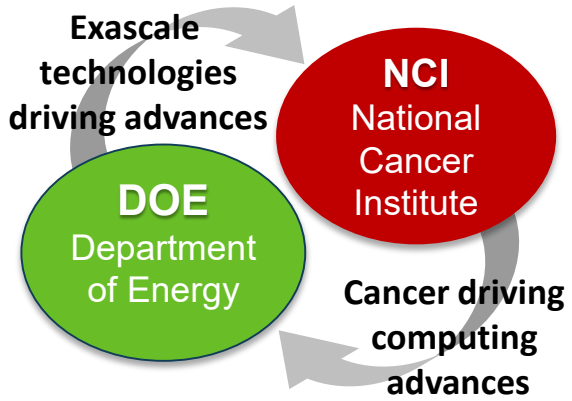
U.S. DEPARTMENT OF ENERGY
NIH NATIONAL CANCER INSTITUTE
Frederick National Laboratory for Cancer Research
sponsored by the National Cancer Institute

Imagine…what would the world look like if there was a digital twin for every cancer patient?



NCI-DOE Collaboration: Joint Design of Advanced Computing Solutions for Cancer (JDACS4C)

DOE-NCI partnership to advance exascale development through cancer research

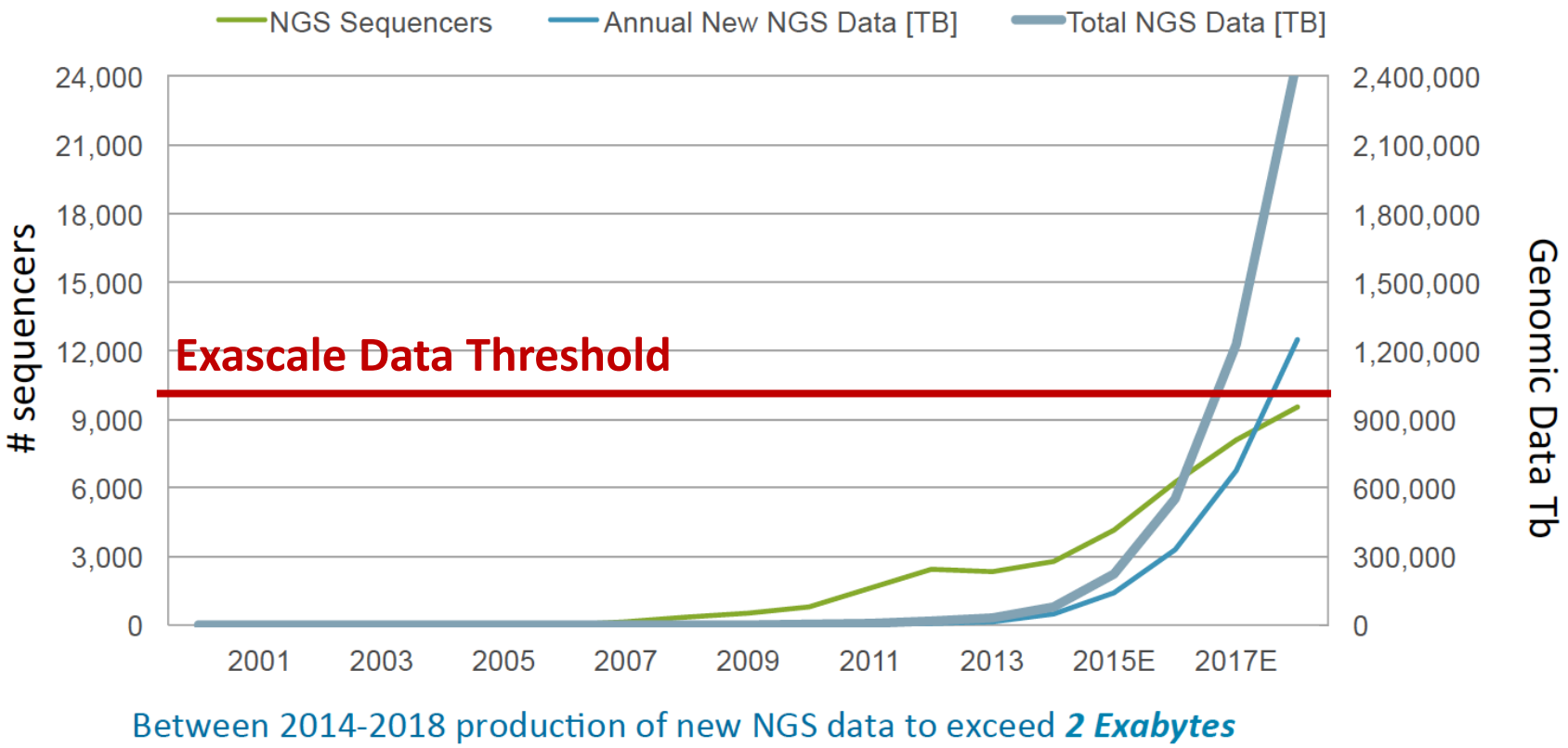


Initiatives Supported
NSCI and PMI

NIH NATIONAL CANCER INSTITUTE
Argonne NATIONAL LABORATORY
OAK RIDGE National Laboratory
Lawrence Livermore National Laboratory
Los Alamos NATIONAL LABORATORY EST. 1943
Frederick National Laboratory for Cancer Research
sponsored by the National Cancer Institute

U.S. DEPARTMENT OF ENERGY | NIH NATIONAL CANCER INSTITUTE

Growing Volume and Complexity of Cancer Data



NGS: Next Generation Sequencing
NGS sequencers include machines from Illumina, Life Technologies, and Pacific Biosciences. Human genome data based on estimates of whole human genomes sequenced
Sources: Financial reports of Illumina, Life Technologies, Pacific Biosciences; revenue guidances; JP Morgan; The Economist; Seven Bridges Analysis.

Healthcare expected to reach over **2,000** exabytes of data by 2020

Gardner, TS, *Trends Biotechnol*., 2013

NCI Precision Oncology: 2006-2020

Interventions to prevent, diagnose, treat cancer based on molecu
mechanistic insight for a particular individu

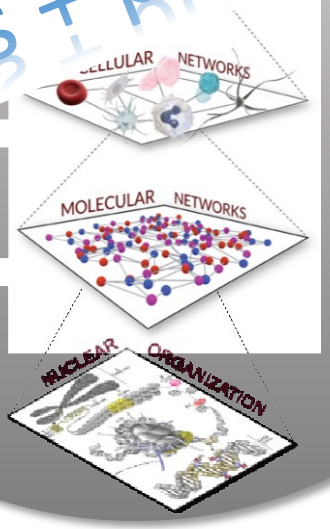
The Cancer Genome Atlas — Understanding genomics to improve cancer care

CLINICAL
TUMOR

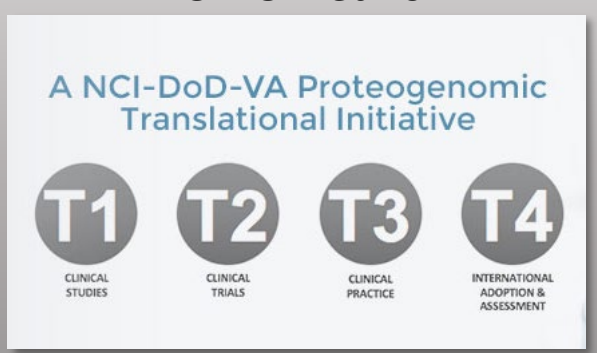
MATCH

DESCRIPTIVE ANALYTICS

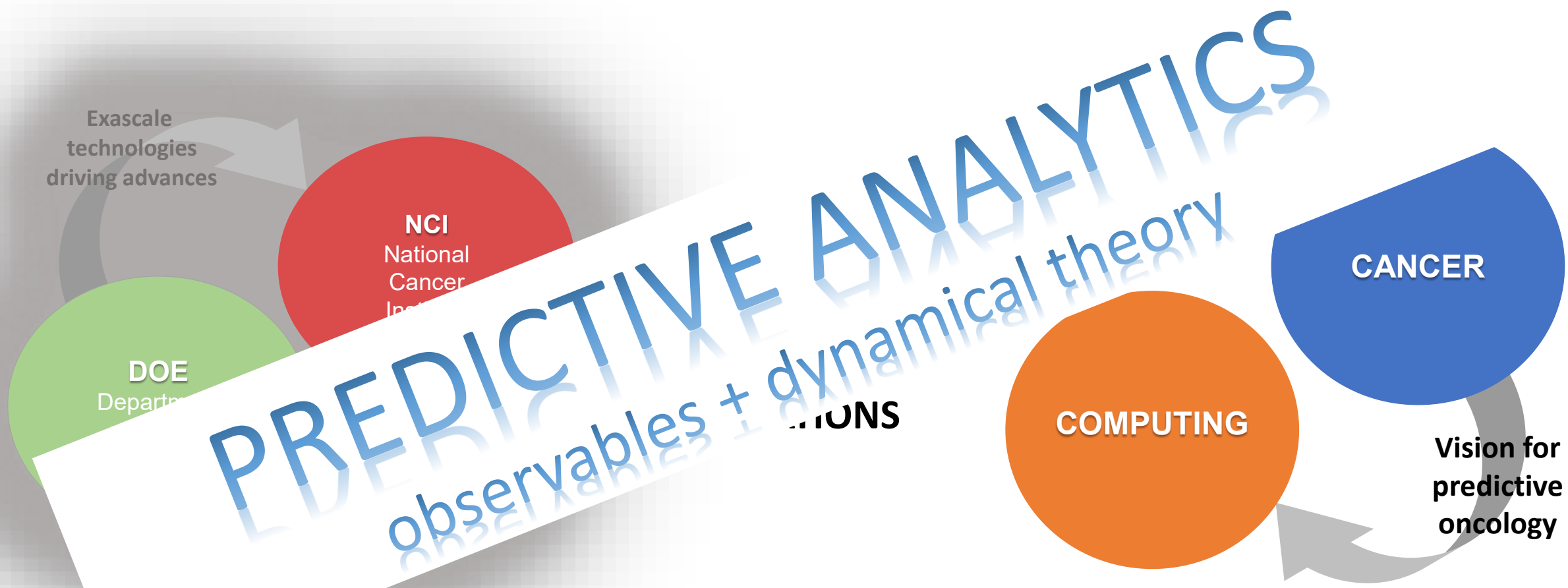
observables + prior measurements



APOLLO Network



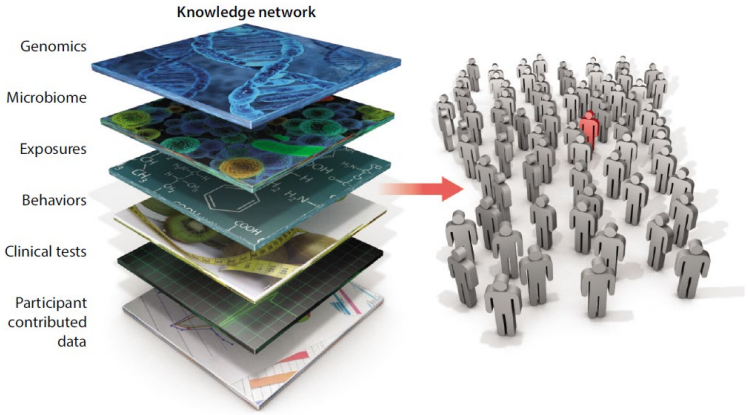
Vision: Predictive Oncology Ecosystem



Towards a National Learning Healthcare System for Cancer

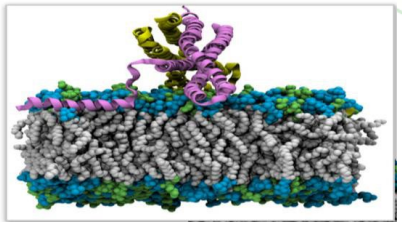
Provide a seamless data environment for patients, providers, and researchers

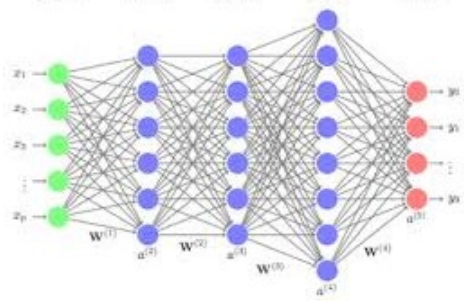
Data



Process, visualize, and analyze data using open APIs and collaboration platforms

Analysis





Develop a data-science aware workforce capable of using a digital ecosystem

Workforce

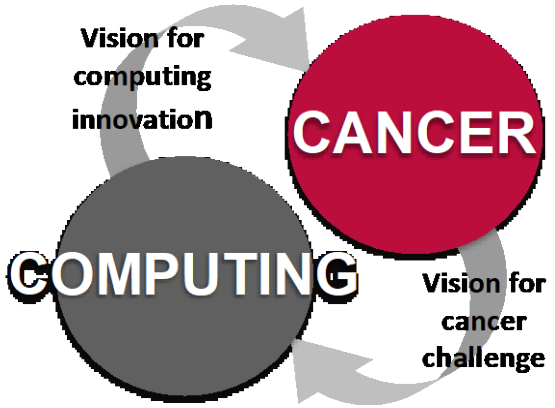
S. Hawgood, et al., Sci Transl Med. 2015; ER Hsu, et al., Clin Pharmacol Ther. 2017

Multidisciplinary Engagement Across Cancer Research, HPC and AI

Envisioning Computational Innovations for Cancer Challenges (ECICC) Community

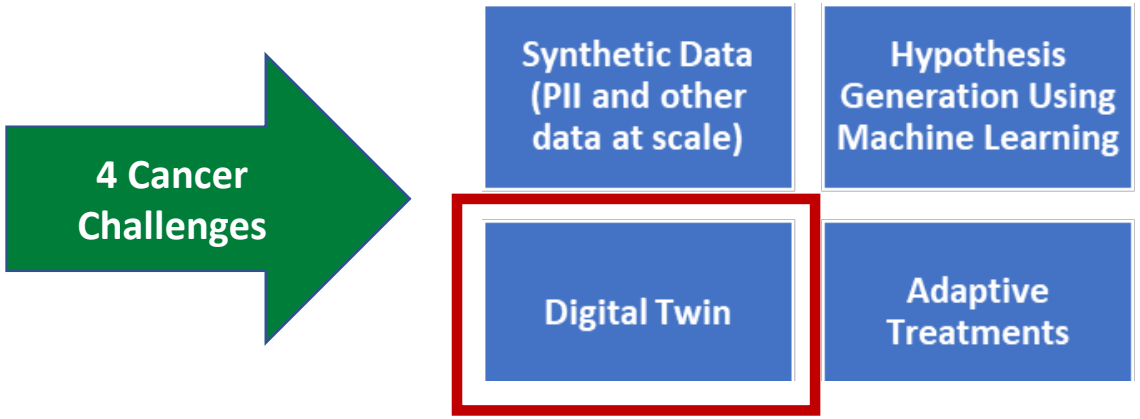
PURPOSE: Build a community, Multidisciplinary engagement, and collaboration among cancer, data, and computational scientists to create transformative impact

ORIGIN: Outgrowth of **NCI DOE Collaboration** Joint Design of Advanced Computing Solutions for Cancer (JDACS4C)



ECICC SCOPING MEETING – March 25-27, 2019

- Multidisciplinary meeting
- Held at DOE's Lawrence Livermore National Lab
- *74 computational scientists & cancer researchers from all career stages participated*
- *Identified over 200 cancer challenge ideas*



What is a Digital Twin?

Dynamic, ultrahigh fidelity simulation of a physical system that can be used to make predictions through virtual experiments

1000 billion DOF physics-based Digital Twin

Loads
Environmental factors
Usage factors

1000 "flight" hours

Large-scale HPC

***Reliability predictions** for all primary structural components and remaining useful life*

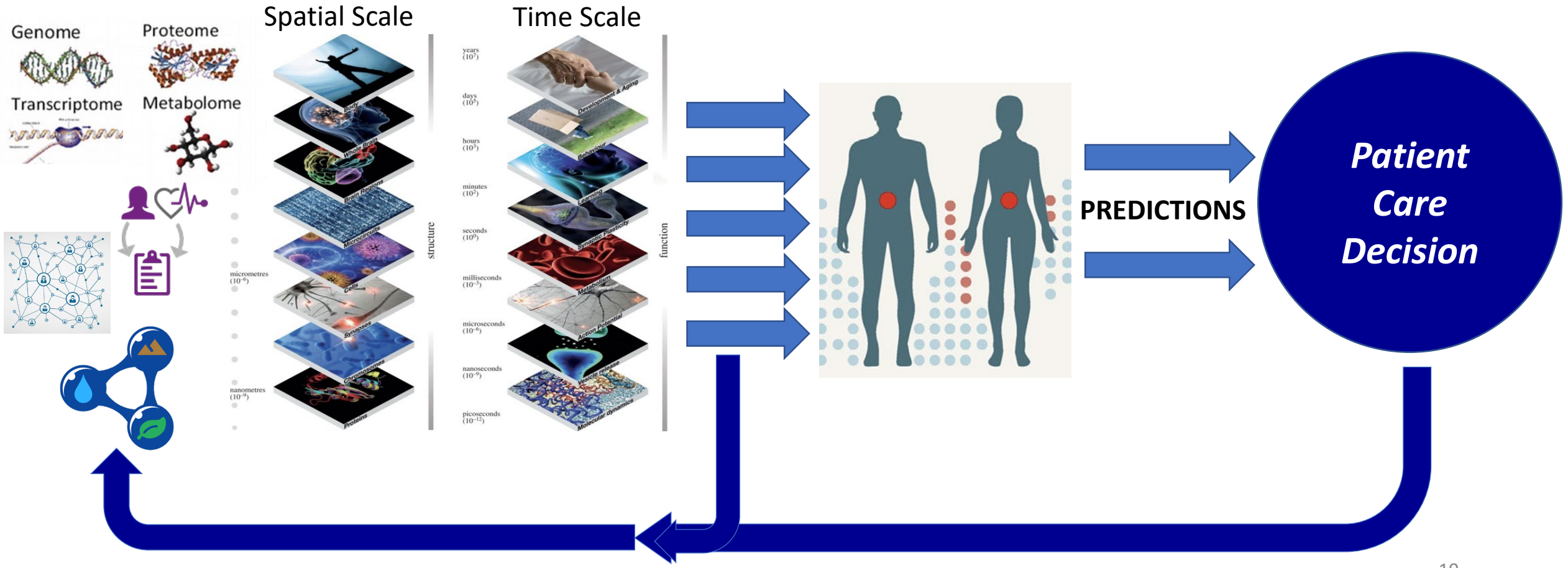
What Makes This Possible?

- Real-time surveillance by 100s of sensors
- Known estimates of flight durations, trajectories and maneuvers
- Known models for entire range of physics acting on the structure over time
- Access to historical and fleet data
- High Performance Computing (HPC)

Tuegel EJ, et al., 2011; Glaessgen EH, et al., 2012

The Challenge: Digital Twin for Predictive Oncology

Patient-tailored models incorporating multi-omic, clinical, environmental and social data that can evaluate and predict the most effective prevention and therapeutic plans



Areas for innovation: Towards a Cancer Patient Digital Twin

- Spatial and time scales of both healthy and disease states
- Data
  - Amount
  - Capture
  - Bridging qualitative and quantitative
  - ML ‘readiness’
- Computational learning frameworks
- Computational and mathematical models for dynamic, multiscale systems in biology

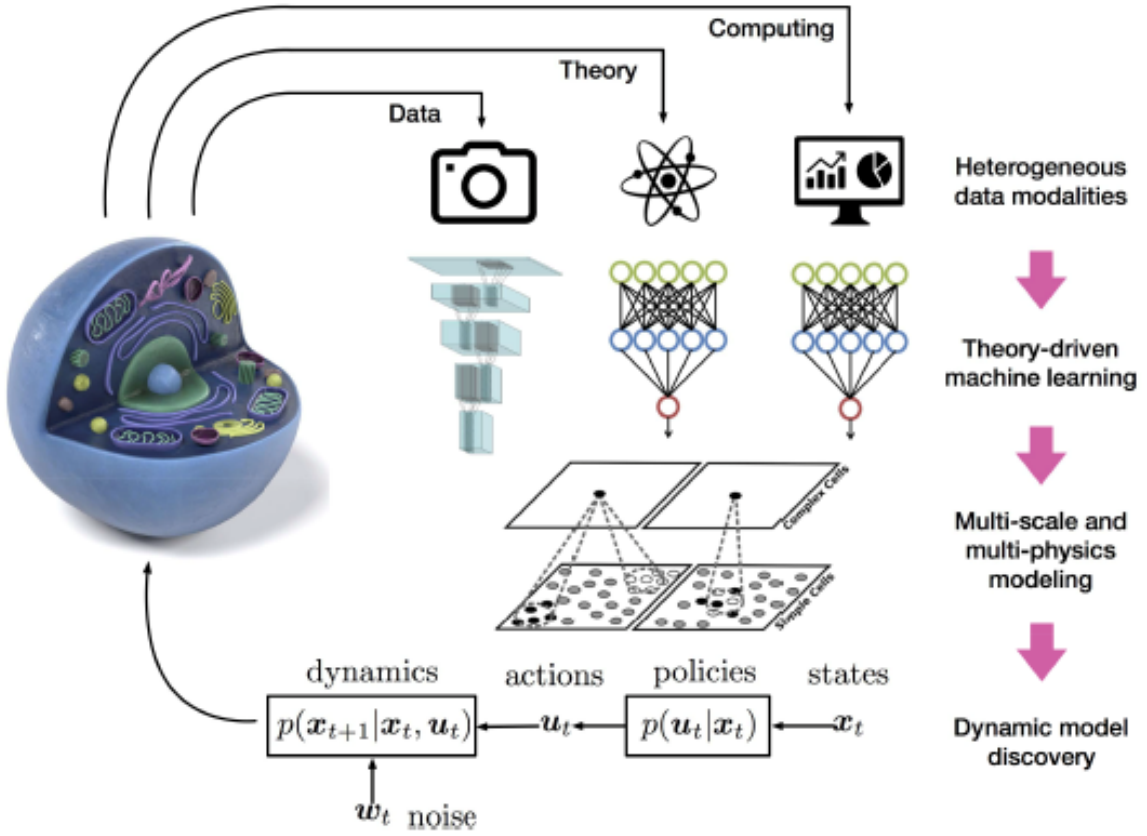
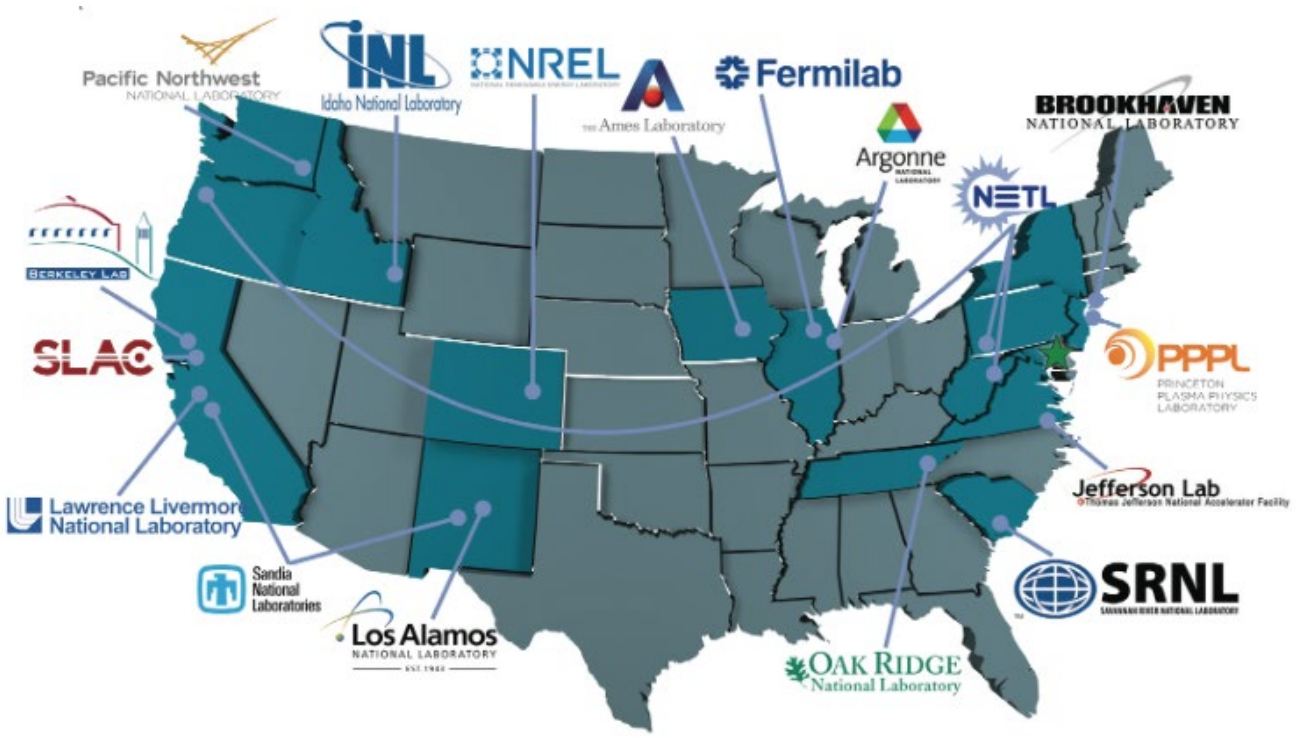


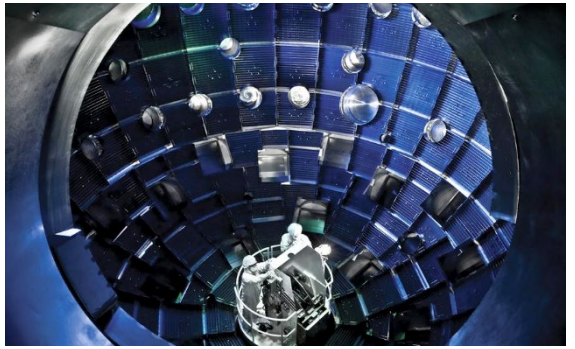
Image: Peng, GCY, et al., Archives of Comp Methods Eng, 2020

DOE Mission

To ensure America's security and prosperity by addressing its energy, environmental and nuclear challenges **through transformative science and technology solutions.**



17 National Laboratories



World-leading experimental facilities

Most Energetic Laser Facility Ever Built
National Ignition Facility, Lawrence Livermore Nat'l Lab

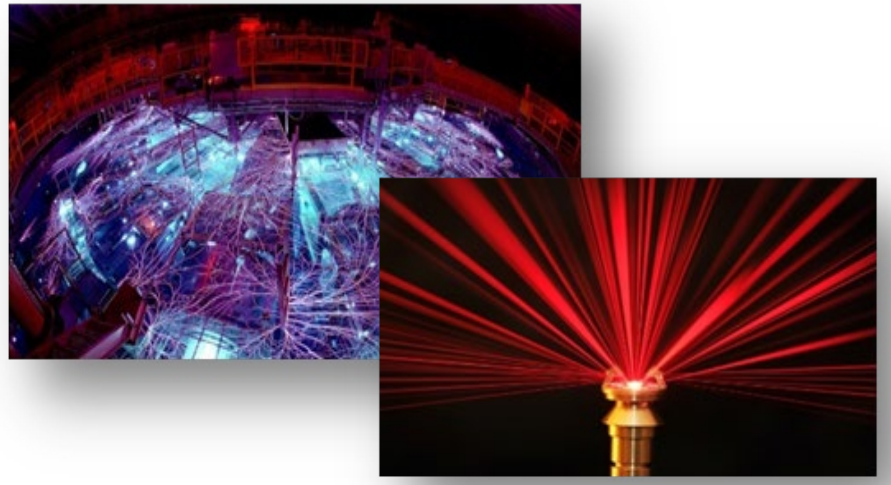
Stewardship; Big Science, Large Teams

Superconducting accelerator will enable the world's most intense high-energy neutrino beam
PIP-II Accelerator, Fermi National Accelerator Lab

DOE Mission and Computing

Nuclear Stockpile Stewardship

To sustain a safe, secure, and effective nuclear deterrent through the application of science, technology, engineering, and manufacturing.

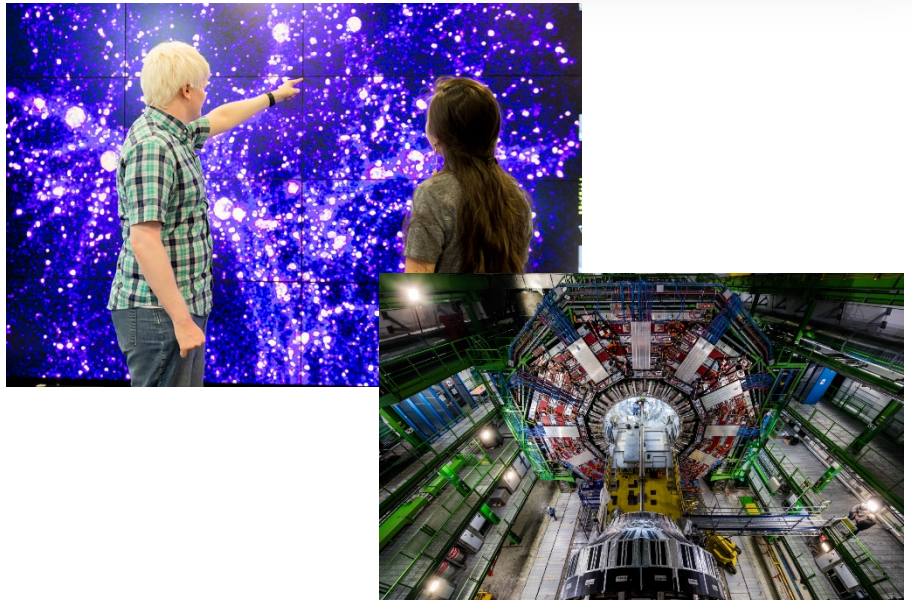


DOE Science Mission

Discover, develop, and deploy computational and networking capability to analyze, model, simulate and predict complex phenomena important to the DOE and the advancement of science.

- Mission Computing
- Leadership Computing

  among the most advanced HPC in the world in terms of performance in solving scientific and engineering problems



DOE is a World Leader in Computing

Summit, Oak Ridge National Laboratory
Leadership Computing Mission
#1 Top 500 from 6/2018 - 11/2019

Cori, Lawrence Berkeley Laboratory
DOE Science Mission

1. Japan
2. **United States - DOE**
3. **United States - DOE**
4. China
5. China
6. Italy
7. United States
8. United States
9. Italy
10. Switzerland
11. **United States – DOE**

DOE is also 14, 16, and 17 in top 20…

Sierra, Lawrence Livermore National Laboratory
Nuclear Stockpile Stewardship

Trinity, Los Alamos and Sandia National Laboratories
Nuclear Stockpile Stewardship

Department of Energy Roadmap to Exascale Systems

Pre-Exascale Systems [Aggregate Linpack (Rmax) = 323 PF!]

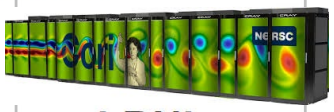
First U.S. Exascale Systems

2012	2016	2018	2020	2021-2023

ORNL
Cray/AMD/NVIDIA

ANL
IBM BG/Q

ANL
Cray/Intel KNL



LBNL
Cray/Intel Xeon/KNL

ORNL
IBM/NVIDIA

ANL
Cray/Intel/TBD

LBNL
Cray/AMD/NVIDIA

LLNL
IBM BG/Q

LANL/SNL
Cray/Intel Xeon/KNL

LLNL
IBM/NVIDIA

ORNL
TBD

ANL
Intel/Cray

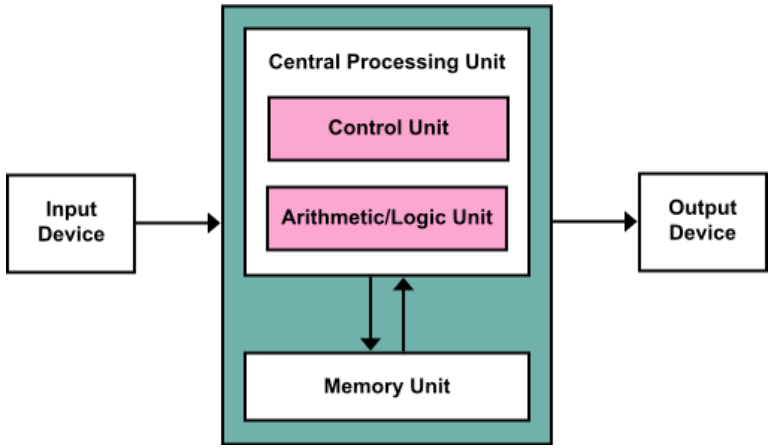
LANL/SNL
TBD

LLNL
TBD

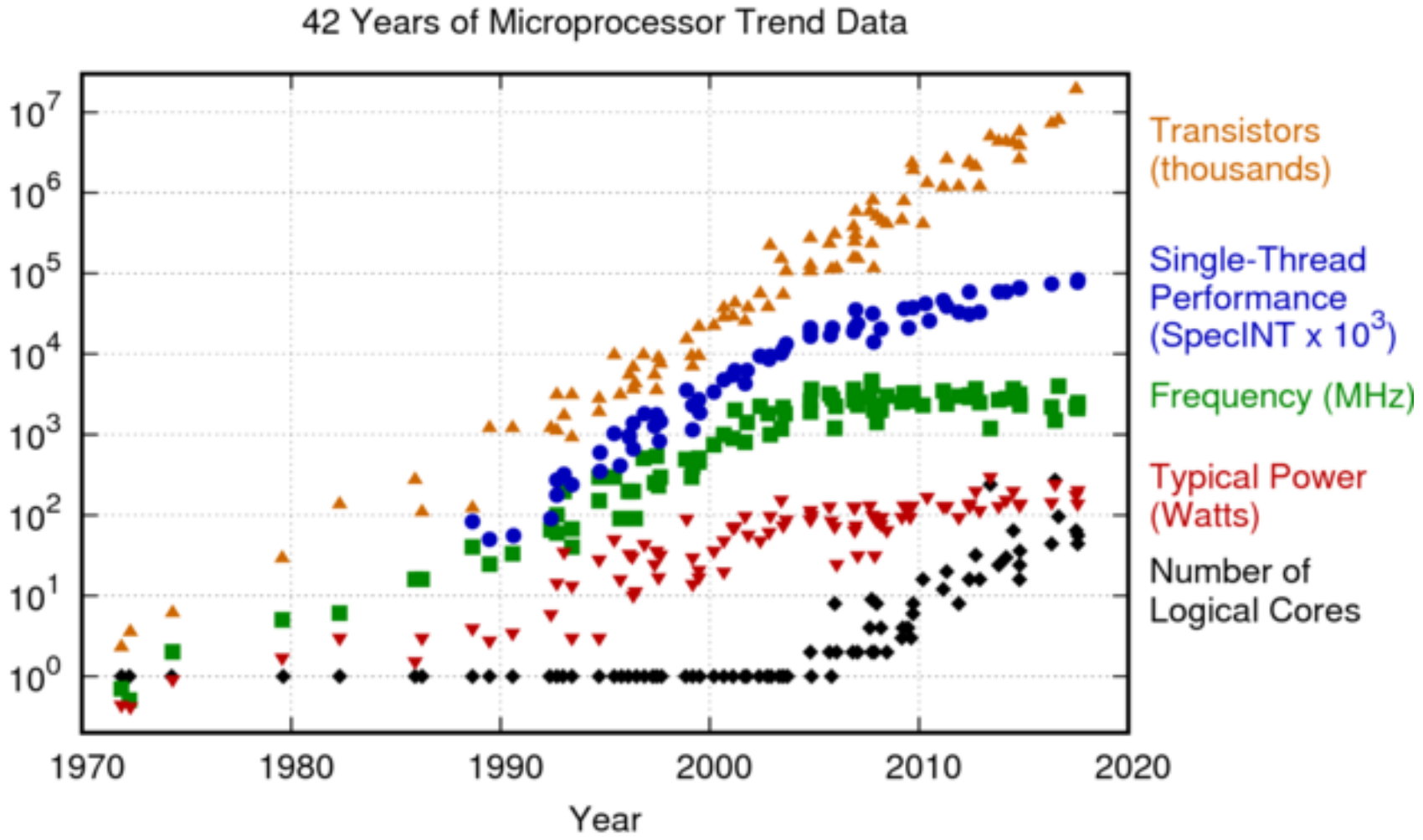
EXASCALE COMPUTING PROJECT

Why Now? End of Moore's Law

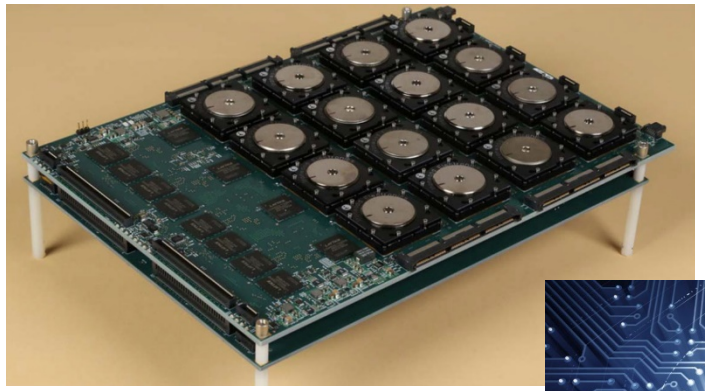
'Same as Usual' has done well, until now.

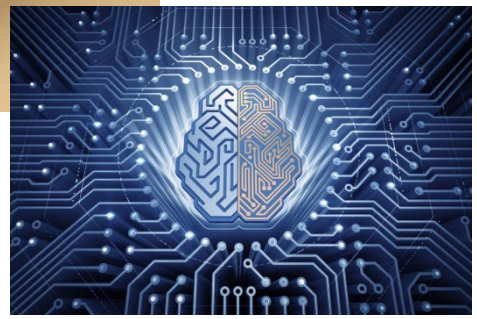


Von Neumann architecture

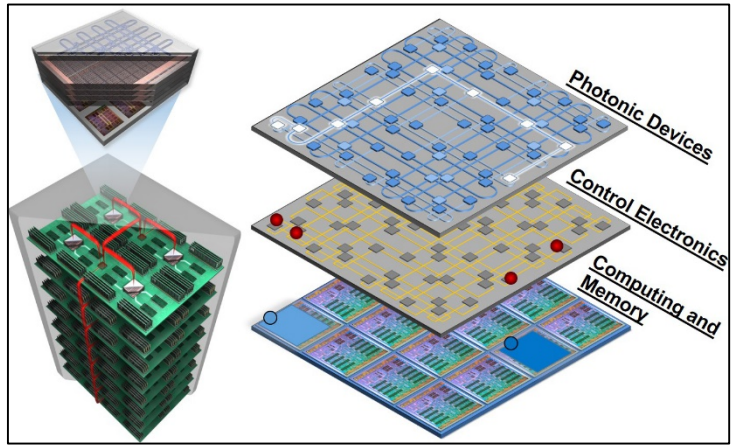


Why Now? Renaissance in Computing Technology and Applications

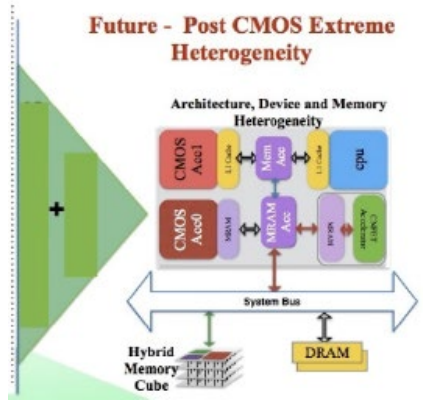




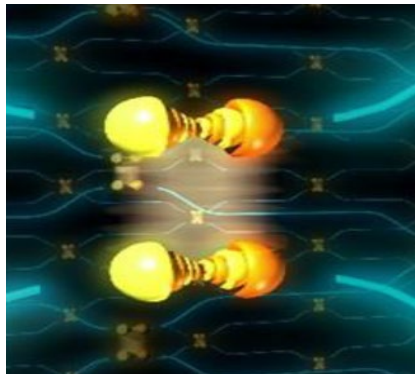
Neuromorphic Computing
"Brain Inspired" Chips
(IBM TrueNorth Chip)



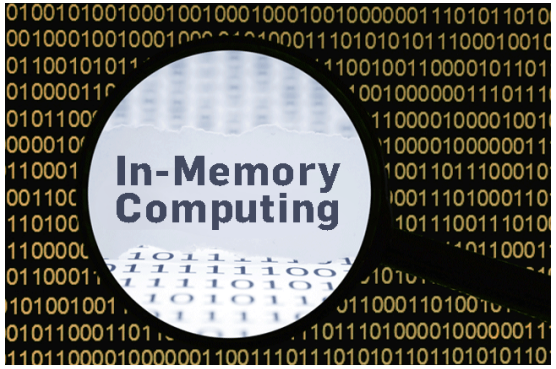
Photonic Communication in Computers
(*Karen Bergman, Columbia Univ)*



Progressively Heterogeneous Systems



Quantum Computers



Field Programmable Gate Arrays

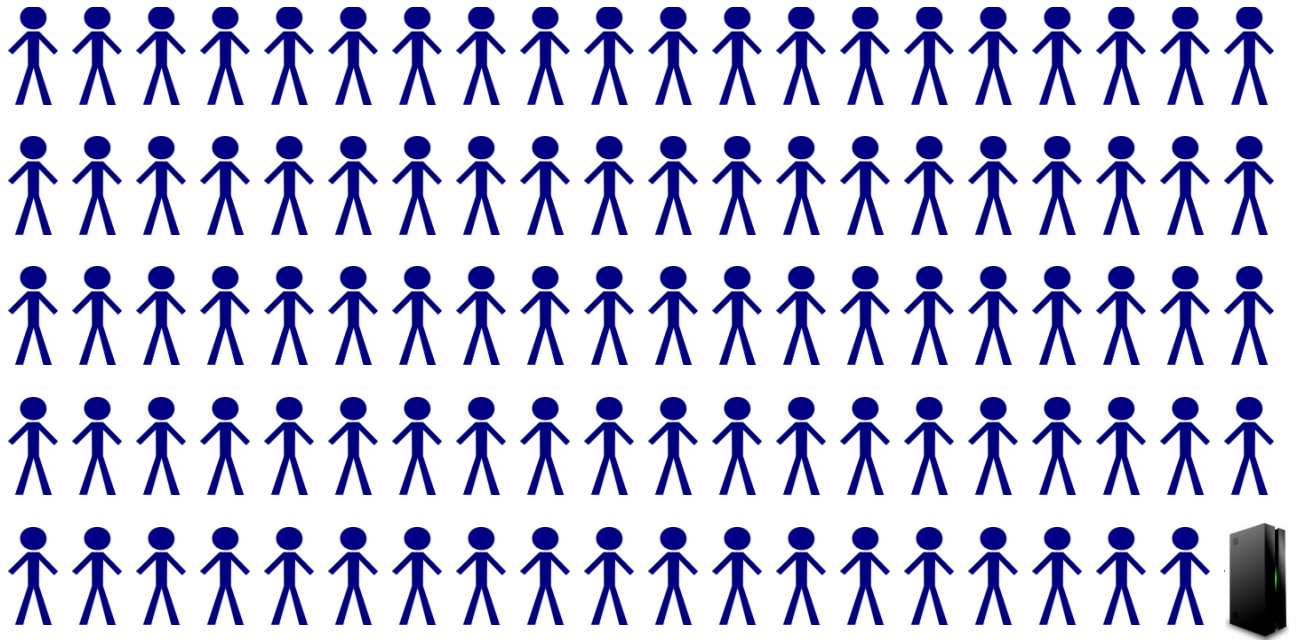
How do we drive computing solutions in the new technology era?

Near term: Artificial Intelligence/Machine Learning, heterogeneous architectures and memory, data movement and processing…

We need partnerships in science and technology fields outside DOE traditional spaces to inspire fresh thinking!

Why the Digital Twin for Predictive Oncology?

- Traditional approaches involve many individuals for general predictions
- Results take time to achieve
- Imprecise conditions
- Explorations limited by available physical models, samples, data



"The Digital Twin Approach"

- Digital twin approaches involve many models for precise predictions
- Explore many possible treatments
- More rapid explorations
- Set specific conditions
- Integrate understanding

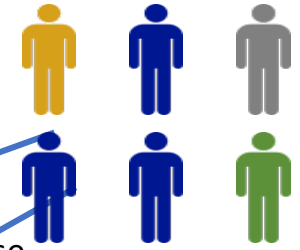
Critical insights for the individual cancer patient

Cancer Patient Digital Twin:

New insights and approaches from molecular to patient scale!

New observables
New data
New approaches

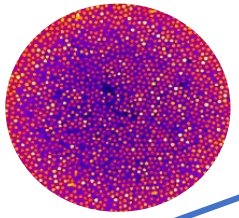
Multiple perspectives
Integrated objectives



Prescriptive cancer models

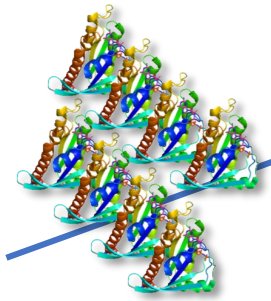
Patient specific disease ensembles

Multiple distinct cancer models



Single cancer ensemble

Multi-scale
Multi-cell models

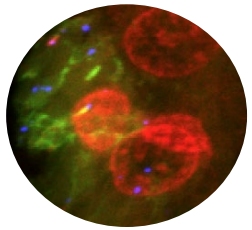


Multi-mechanism
Multi-ensemble

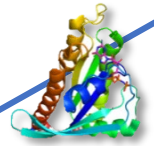
Multi-scale
Single cell models

Single mechanism
Single ensemble

Multi-mechanism
Single ensemble



Single mechanism
Single trial

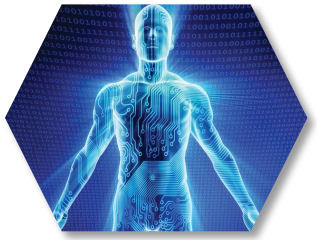


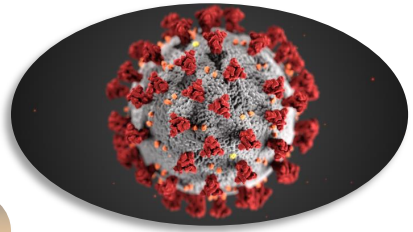
Reaching the Patient for Impact

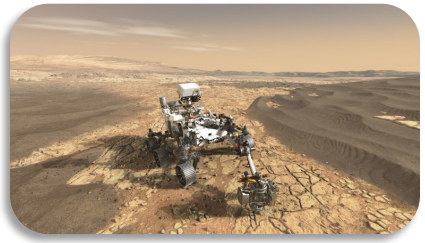
Multiple starting points
Team collaborations

Challenges – Inspiring new ideas and new horizons!

- Sequence the human genome
- Apollo mission to the moon
- Manhattan project







- COVID-19
- Mission to Mars
- *Cancer Patient Digital Twin*

Data Challenges Can Feel Like a Black Hole –
Don't Get Sucked In!

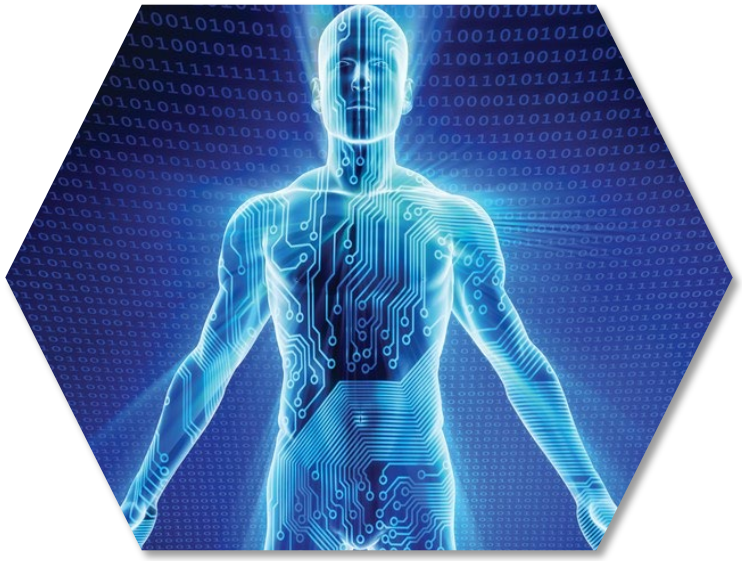
Consider data needed to fulfill the concept but do not limit thinking by access/feasibility concerns such as:

- Bias/noise
- Sample size
- Sharing/acquisition policies
- De-identification of PHI/PII
- Standardization and labeling
- Ease of integration/linkages/comparisons

Think beyond standard data generation/annotation projects!

The Charge: Chart the course for a CPDT!

- Think beyond the immediate….
- What would be possible if…?
- What will be possible when…?
- What will be different when current efforts finish?
- Where will technology be in ten years?
- As current barriers are surpassed, what follows?



- How to get started?
- What are the steps to move ahead?

Guiding long-term objective: Deliver a digital twin that provides critical, timely and responsive insight for the *individual* cancer patient!