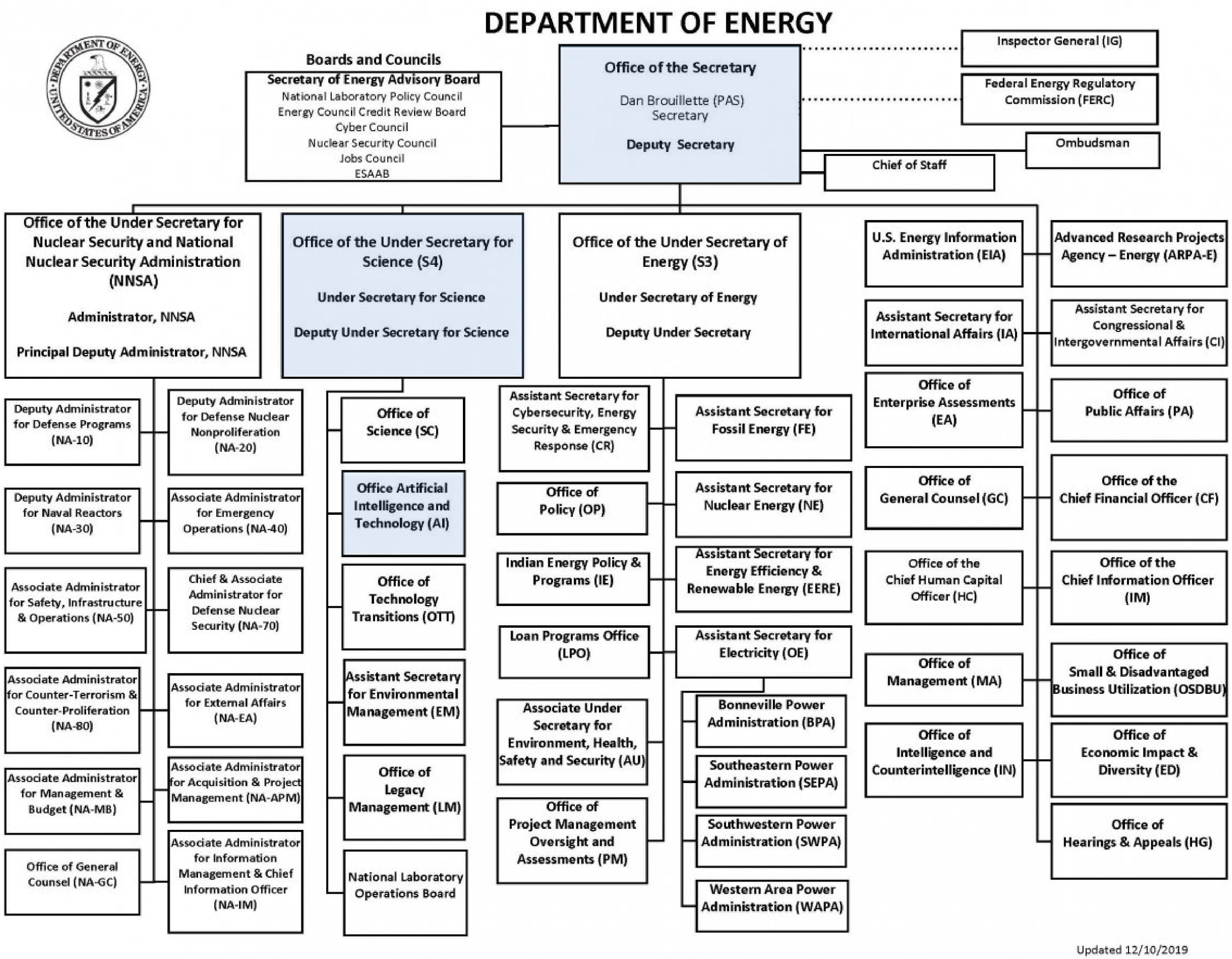
AITO: Artificial Intelligence and the DOE

Advanced Scientific Computing
Advisory Committee
January 13-14, 2020
Washington, D.C.

Fred Streitz
Artificial Intelligence and Technology Office
US Department of Energy

Artificial Intelligence and Technology Office created September 6, 2019

- Strategic Coordination Across DOE Enterprise
- Advance Frontier Technologies for the Department
- Accelerate Leadership Through Partnerships



U.S. DEPARTMENT OF ENERGY | Artificial Intelligence and Technology Office

How we think of *Artificial Intelligence*

AI is a disruptive, multidisciplinary field seeking to train computer systems to autonomously perform tasks that mimic the application of human intelligence.

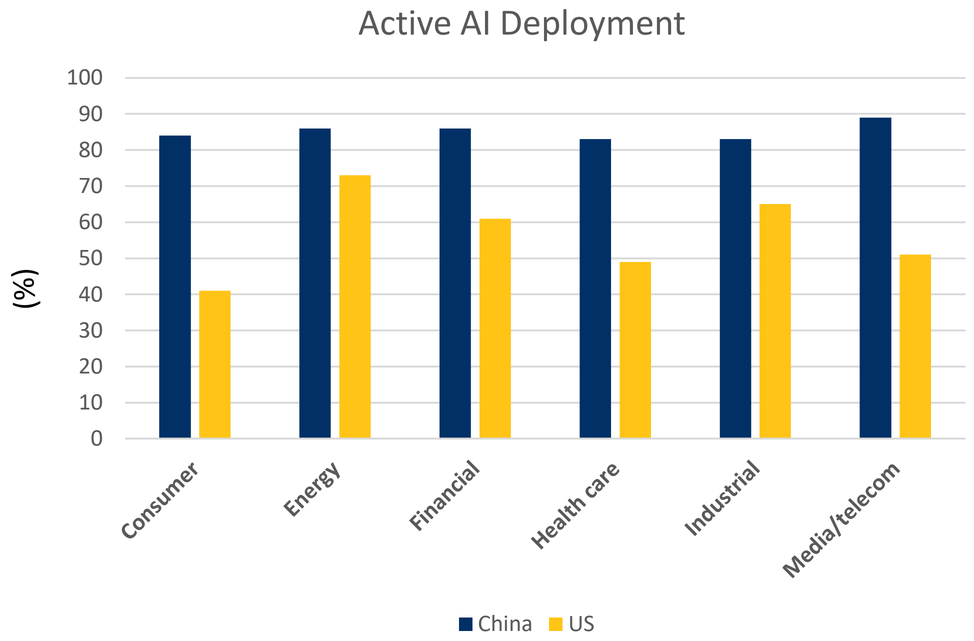
- AI lets computers continuously learn, make decisions and react like humans – but at scales beyond human capabilities.
- AI can learn from data that does not fit into spreadsheets: e.g., conversations, videos, attitudes, physical models or environmental systems.
- AI will be a driver for economic growth and productivity in the United States and the world. It will enhance our economic and national security and improve our quality of life.

AI matters for the nation

"The development of AI will shape the future of power." – NSCAI Interim Report

- AI will be the economic driver of the next decade – projected to add at least 14% ($4T) to the US economy (+26% in China)
- Dozens of countries developing strategies (but seen as a US/China race)
- China and Russia have stated strong intentions to develop global leadership
- China is outspending the US about 10:1 in AI research and development, and applying AI across their entire society

(e.g. City of Tianjin has a $16B AI municipal fund)



Forbes, December 2018

US failure to lead in AI will have grave consequences

U.S. DEPARTMENT OF ENERGY | Artificial Intelligence and Technology Office

HPC - Part of DOE’s technology edge

- **Computing has been a part of the DOE mission since the very beginning**
  - *Computing at scale from Univac in the early 50’s to Sierra and Summit*
- **Modeling and simulation has been our approach for problems that cannot be instrumented – a foundational concept of stockpile stewardship**
  - *Built around Uncertainty Quantification, Verification & Validation ; Many successes.*
- **The DOE has worked closely with vendors to push technology ahead of the curve**
  - *Massive parallelism (e.g., Blue Gene/L, Sequoia)*
  - *Graphics co-processors at scale (e.g., RoadRunner, Titan, Sierra, Summit, …)*
- **The focus has always been around specific needs**
  - *Hard problems we needed better ways to solve*
  - *Slow & fast dynamics, radiation & matter, transport in 3D, multiscale turbulence, etc.*
- **We were in the mode of trying to solve equations in finite time – AI was not on our radar**

Current opportunities in AI are driven by a confluence of factors

- **Prevalence of large data sets and streams**

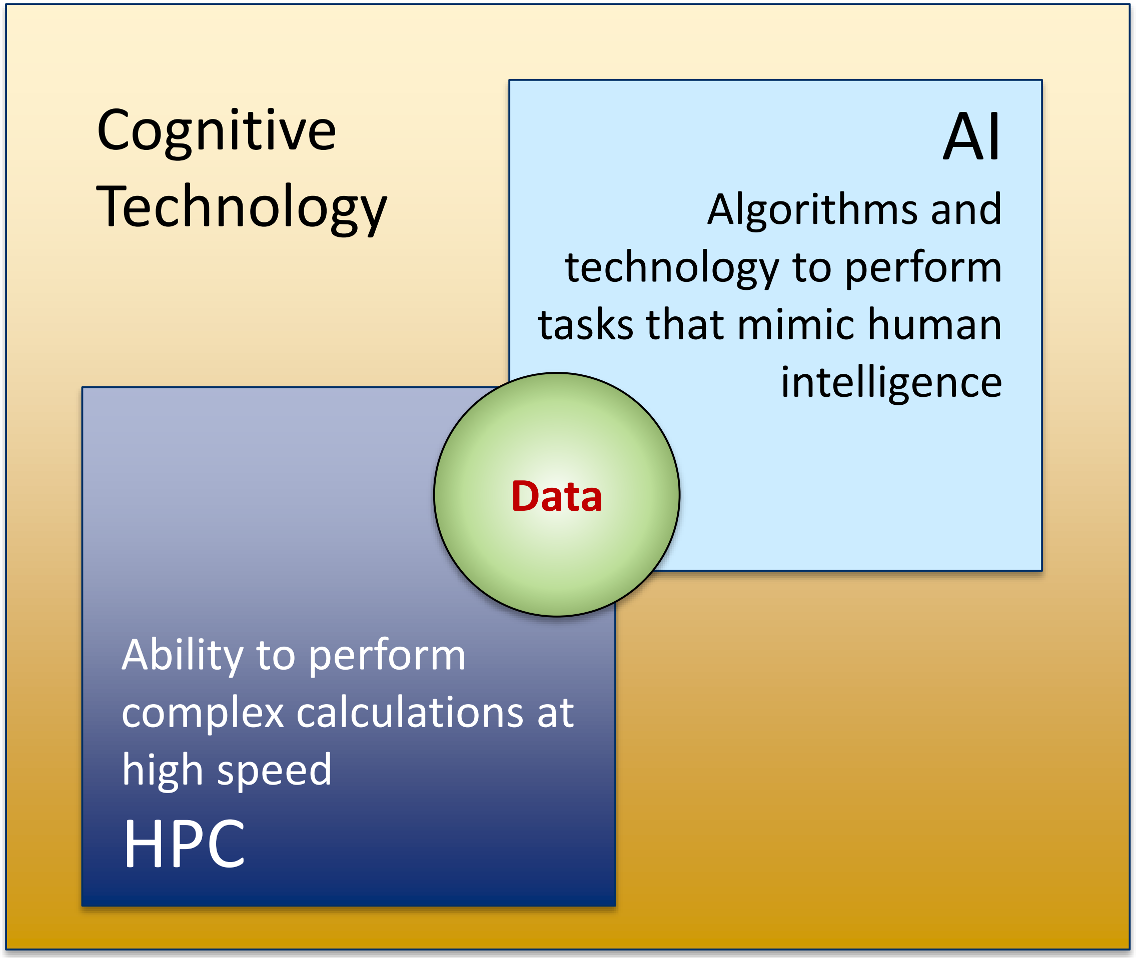
  Today we are creating, sensing, measuring, computing, imaging,… more data than we can humanly deal with. Decision support using all this data is needed in all mission, business and operational areas

- **Ubiquity in computing technologies**

  Both general purpose HPC and machines specifically designed to support large-scale data analytics and machine learning are becoming widely available

Both trends promise to continue - *Challenge is to push towards novel algorithms and more powerful compute on ever larger data sets, uncovering more beneficial use cases.*

AI is enabled by HPC, but has a different technology path & purpose



- Spans sensors, learning, deciding, autonomy, human interface, robotics, …
- Lives in rich data environments
- Can be applied at source of data creation
- Inference platforms need to be portable and power efficient
- Can surface questions and propose actions from data
- Together, HPC and AI form a powerful combination that will drive innovation for decades

DOE will focus on the application and advancement of AI very broadly

Applying AI/ML: Learning today across missions

- *Applications of techniques to existing data sets*
- *Pushing the scale of learning and graphs on large GPU-based systems*
- *Business and operations*

Advancing AI: Technologies tied to outcomes

- *Novel AI hardware architectures co-developed across DOE*
- *From edge to scale in AI technologies*
- *Particular classes of sensors, autonomy, large data acquisition, processing and learning, robotics, …*

Advancing AI: Hard problems that require our lead

- *UQ for AI*
- *Adversarial AI frontier*
- *AI inside and outside HPC – pushing cognitive simulation*
- *Hybrid simulated and measured data for learning*

Advancing AI: Data and its Environments

- *Broad diversity and scales of multimodal data*
- *Trusted data environments & data sciences at DOE scale*
- *Frontiers of data science*
- *Architectures built around data*

Some AI Drivers: Tidal Wave of Data & Information

Data is growing in all areas from business, management, operational and mission:

- Our research arms are focused on specific mission areas
- Much of the department is being inundated with data but has no research arm
- We are consequently accepting risk in ways we cannot quantify by selecting what and how we look at data
- We are in need of a more comprehensive approach to data in our business, operational and mission areas
- The pace of advancement in AI requires new approaches that deliver the speed and agility needed in this rapidly evolving domain. It cannot be realized as a side effort
- AI is a technology <u>today</u> and strongly rooted in the private sector – but they are not addressing key issues relevant to DOE

AITO's Plan of Action

1. Create DOE AI Strategy

- **DOE AI Strategic Plan** will define Departmental goals and determine a long-term **AI Roadmap** to ensure return on investments.

2. Institutionalize the AI Exchange

- Track progress towards DOE strategic goals & objectives, reduce overlap, and ensure mission alignment as Programs move forward.

3. Develop and Implement AI Leadership Training

- Enhance stakeholders required general knowledge of AI.

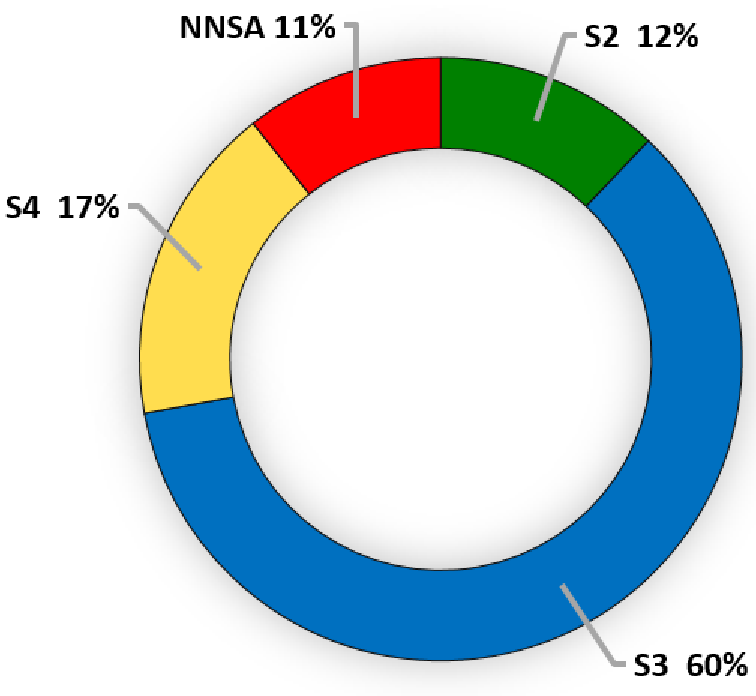
4. Conduct Workshops

- Identify future directions, develop relationships with potential partners

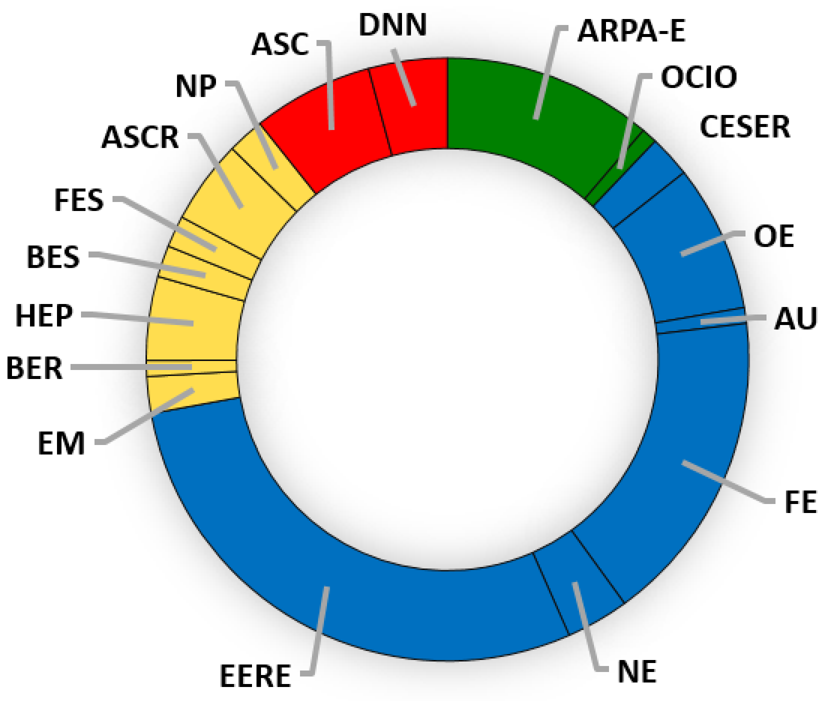
5. Prioritize & Develop Jointly Funded AI Partnerships

- Interagency (e.g., WAPA, EM...), International, Public/Private Activities

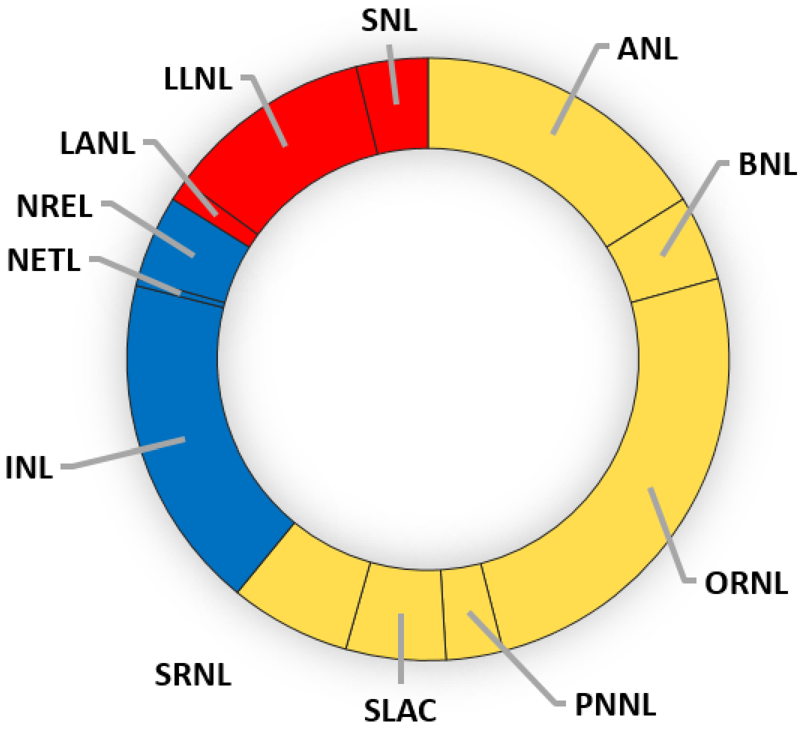
AI Exchange (AIX): capturing sweep of AI activities across DOE



AIX Content
By DOE Organization



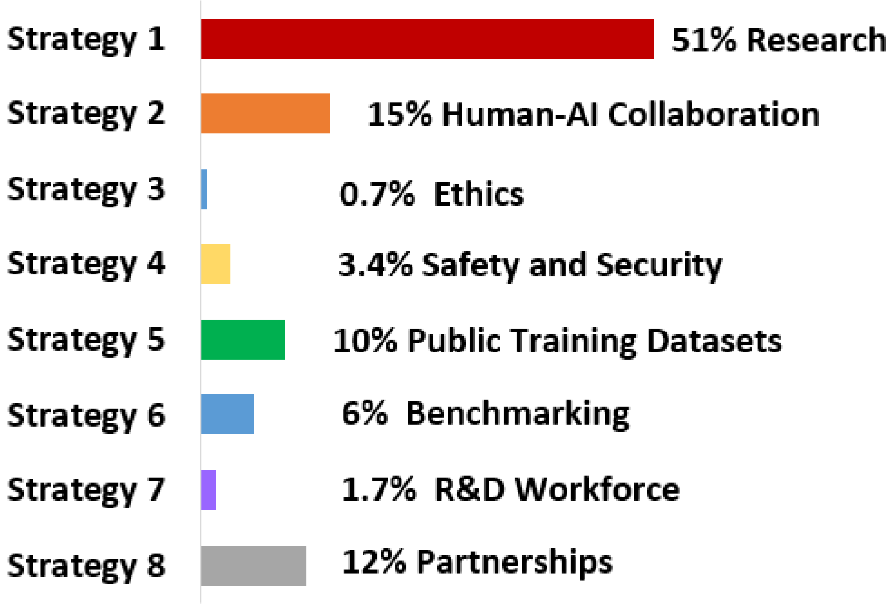
AIX Content
By DOE Programs



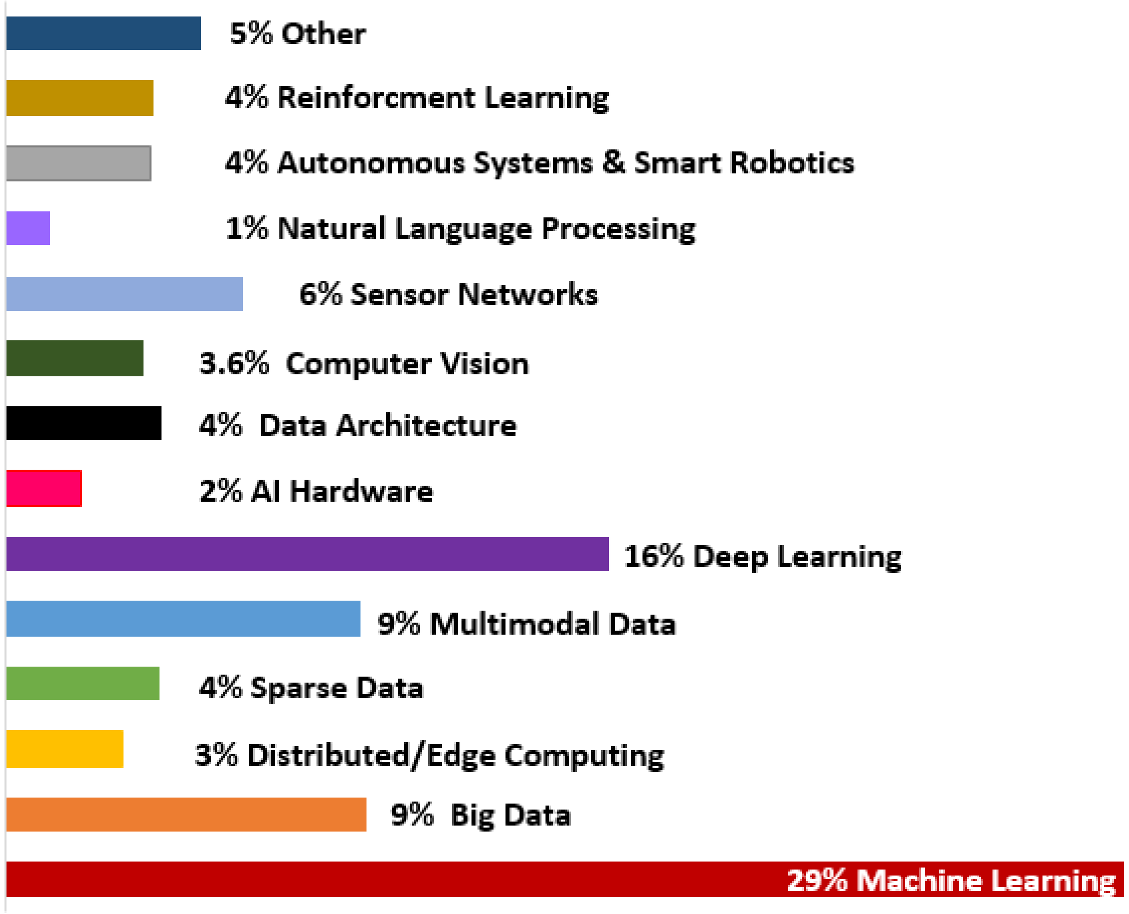
AIX Content
By National Laboratories

Initial parsing of AIX database

By OSTP AI Strategic Goals



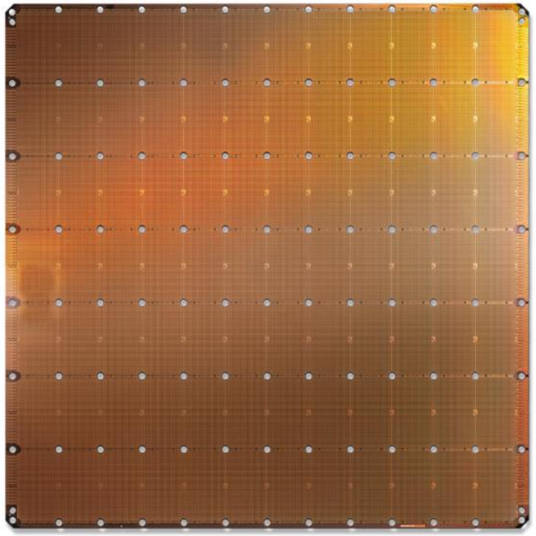
By Technology Type



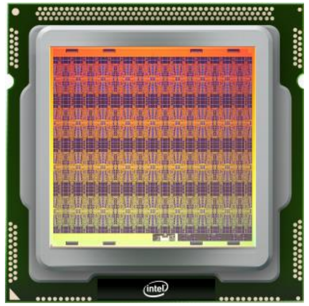
- Starting analysis of input to identify gaps and categorize overlaps
- Data will inform an optimized strategy to maximize return for agency

Bright future requires advances on multiple hardware fronts

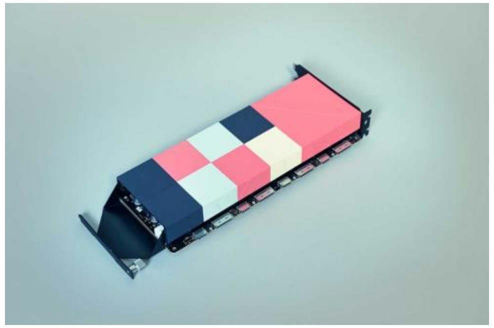
- AI accelerators/processors
  - *Cambrian Explosion needs to happen*
- AI frameworks, methods and algorithms
  - *Develop theory of data drive models*
- Brain-inspired synapse and neuron-based architectures
  - *Move from neuroscience experiments to production*
- Quantum and analog computing and infrastructure
  - *Move from physics experiments to production*
- Progress in traditional CMOS technology/infrastructure
  - *Memory, Fabric, Processors, Storage*
- System architectures to exploit all of the above



Cerebras



Intel Loihi

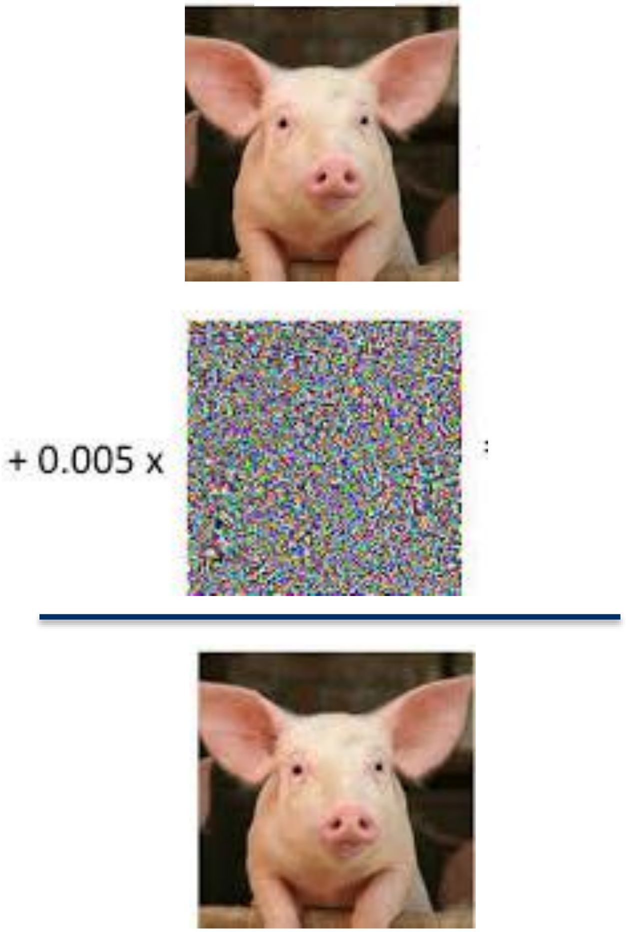


Graphcore

Ubiquitous AI poses serious concerns as well

- AI today is powerful but fragile
- Like the internet – before we worried about cybersecurity
- Today it can be easily fooled
  - *Single pixel defeats of methods – identical to humans, distinct to AI*
  - *Miss-identification/Impersonation/Dodging*
  - *Ignoring/masking visual objects/events (e.g. stop signs,…)*
  - *Data poisoning/reverse engineering models…*
- Many new weaknesses being surfaced
- For decision support, we need measures of certainty in predictions
- Tools – new chips and methods – are being developed today and we have the opportunity to build in what we really need

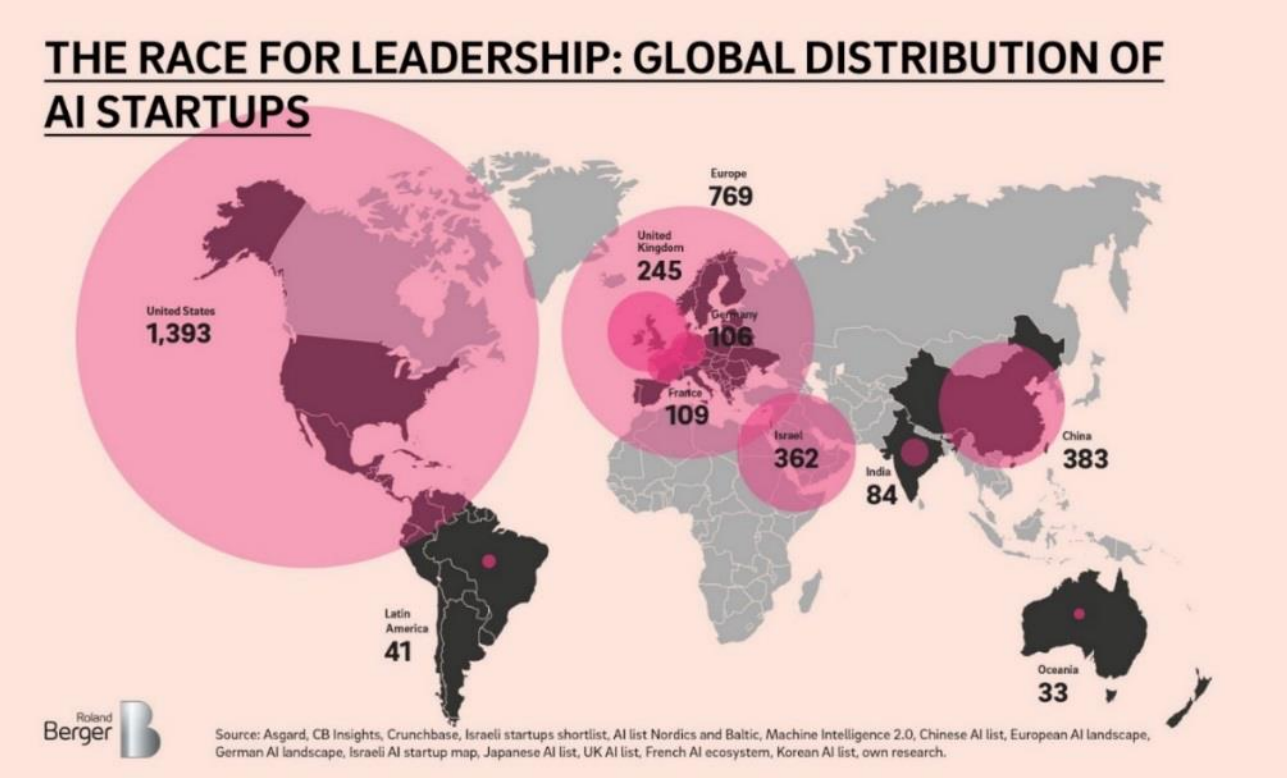
"pig"



"airliner"

US leadership in AI will require strategic partnerships

- Partnership with academia, industry, and governments to develop tools, hardware, data, etc,
- Challenges:
  - Frontier technology is currently driven by the private sector
  - Leadership is forming in competitive and adversarial countries
  - Growing technology gap in AI expertise in government
  - Competition for talent is global



The DOE will be engaging broadly to help the U.S. retain leadership in AI technology

Summary

- The AI landscape is enormous – DOE is in a unique position to make an impact
- We are seeing just the tip of the iceberg in hardware and algorithms innovations- these innovations needs to be encouraged
- HPC and AI are distinct technologies that each need to advance (for mutual benefit)
- The powerful use of the two together will lead to remarkable innovations
- Broad and deep partnerships will be necessary to accelerate advancement
- AITO is organizing across DOE (defining gaps, identifying needs and fostering partnerships) to establish and maintain U.S. leadership in AI

AITO

U.S. DEPARTMENT OF ENERGY

Artificial Intelligence and Technology Office

Joint Design of Advanced Computing Solutions for Cancer

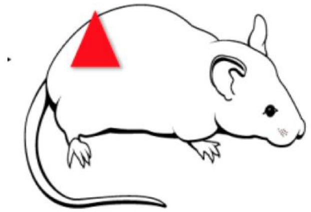
A collaboration between DOE and NCI

NCI
DOE

CANCER MOONSHOT

Pilot 1

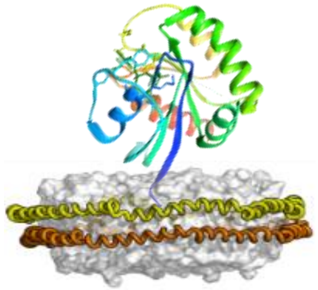
Pre-clinical Model Development



Yvonne Evrard (FNLCR)
Rick Stevens (ANL)

Pilot 2

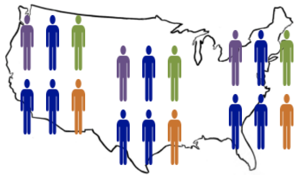
RAS Biology in Membranes



Dwight Nissley (FNLCR)
Fred Streitz (LLNL)

Pilot 3

Precision Oncology Surveillance



Lynn Penberthy (NCI)
Gina Tourassi (ORNL)

Investigate RAS biology on lipid membranes

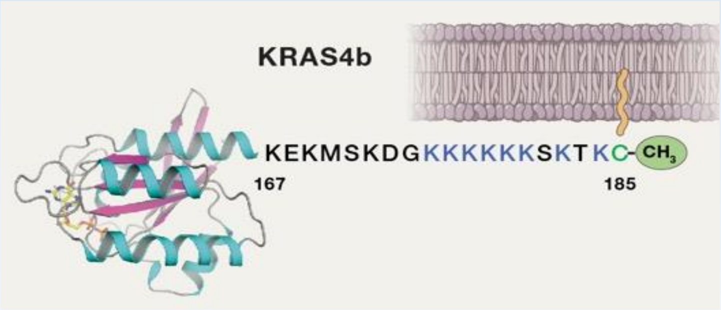
93% of all pancreatic

42% of all colorectal

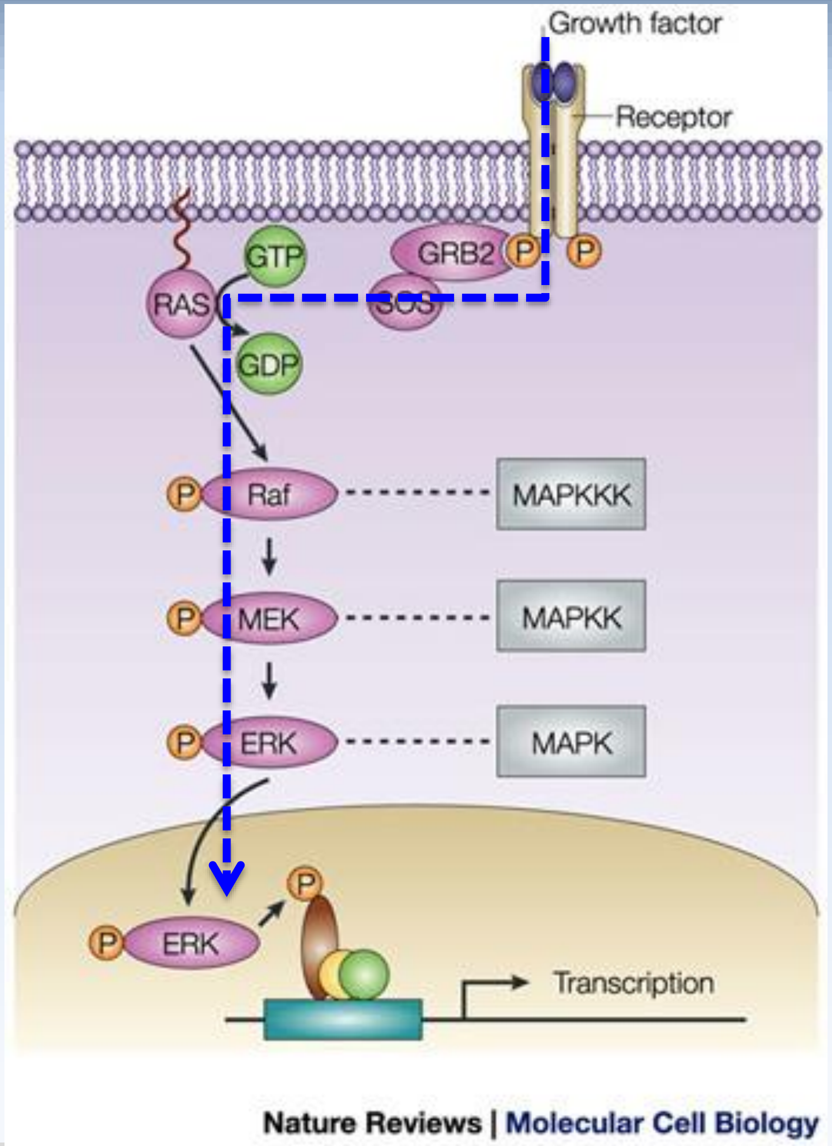
33% of all lung cancers

1 million deaths/year world-wide

No effective inhibitors



Simanshu,Cell 170, 2017



Pathway transmits signals to the nucleus

RAS is a switch- oncogenic RAS is "always on"

RAS localizes to the plasma membrane

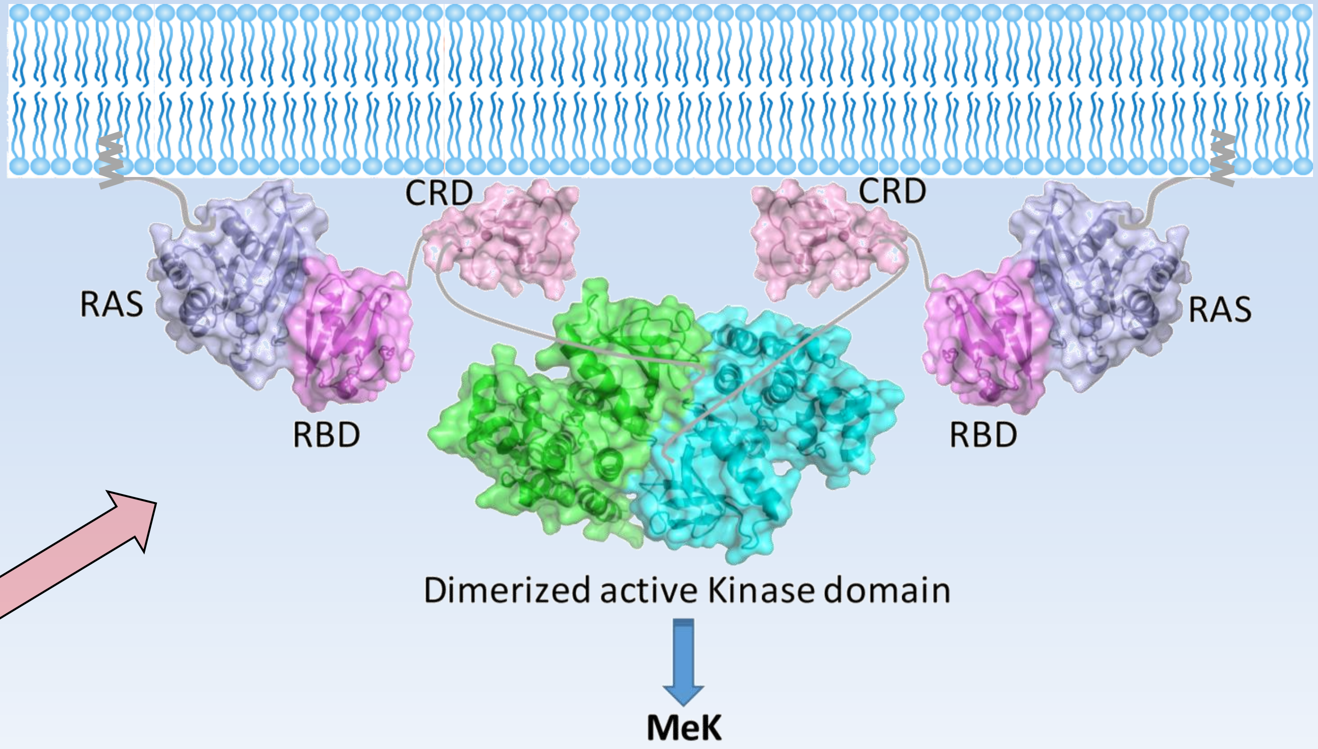
RAS binds effectors (RAF) to activate growth

Essential strategy: utilize appropriate scale methodology for each component

A problem of length and time scale:

Membrane evolves on ms time frame across µm…

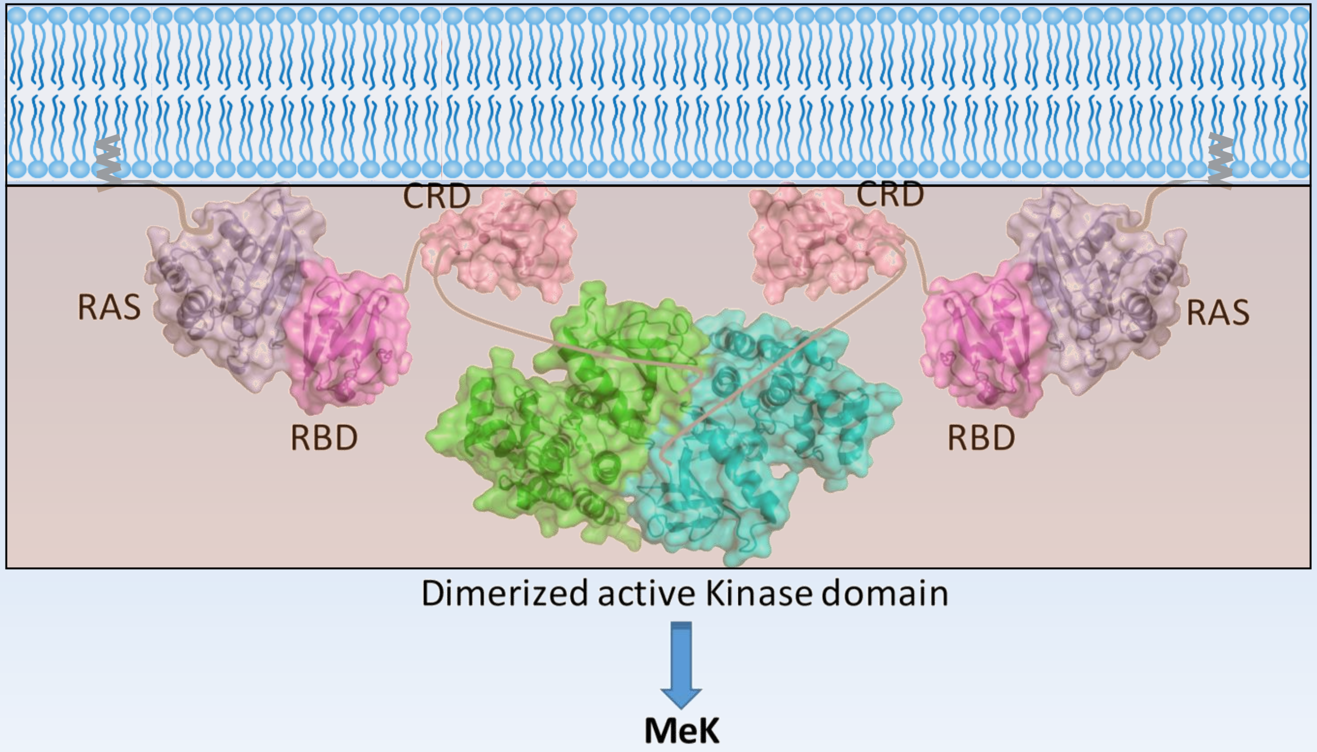
…while protein interactions involve µs across nm



Essential strategy: utilize appropriate scale methodology for each component

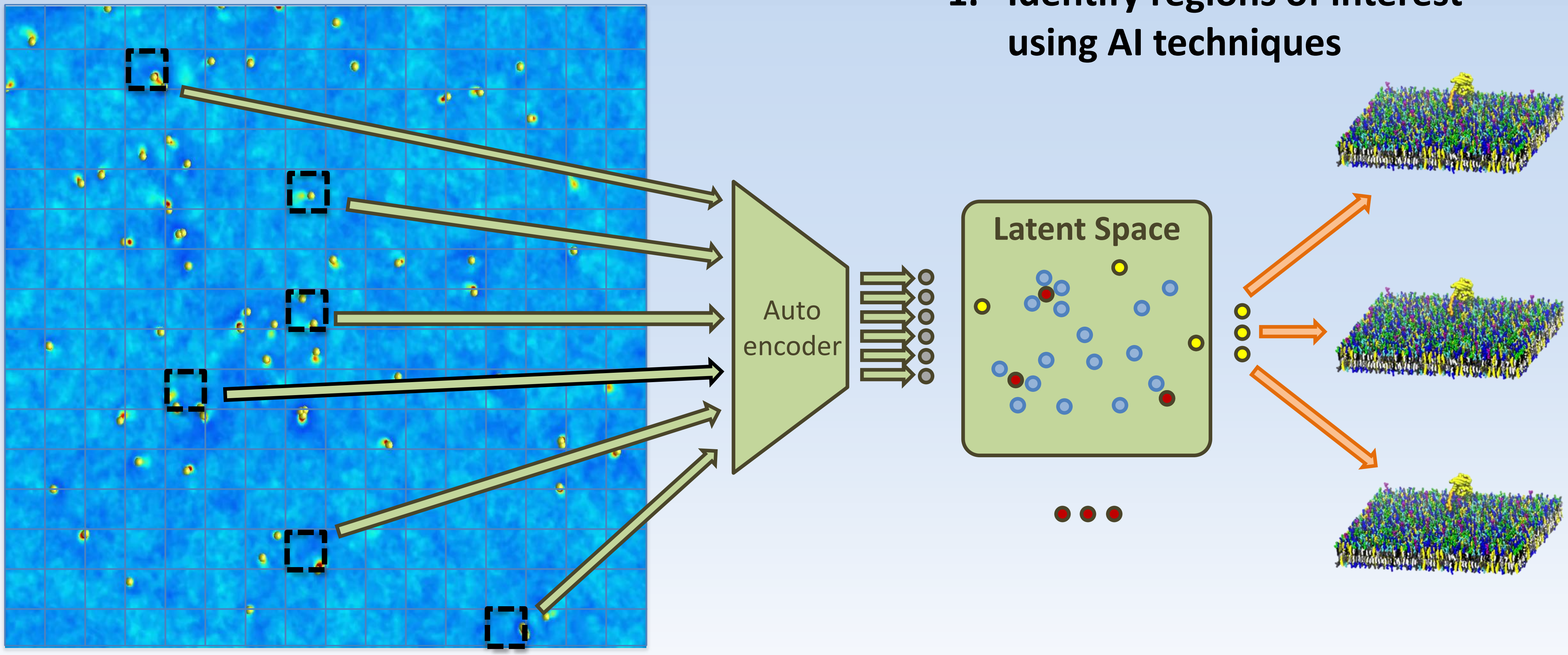
Model membrane dynamics with RAS at micron scale

Model protein behavior in membrane at molecular scale

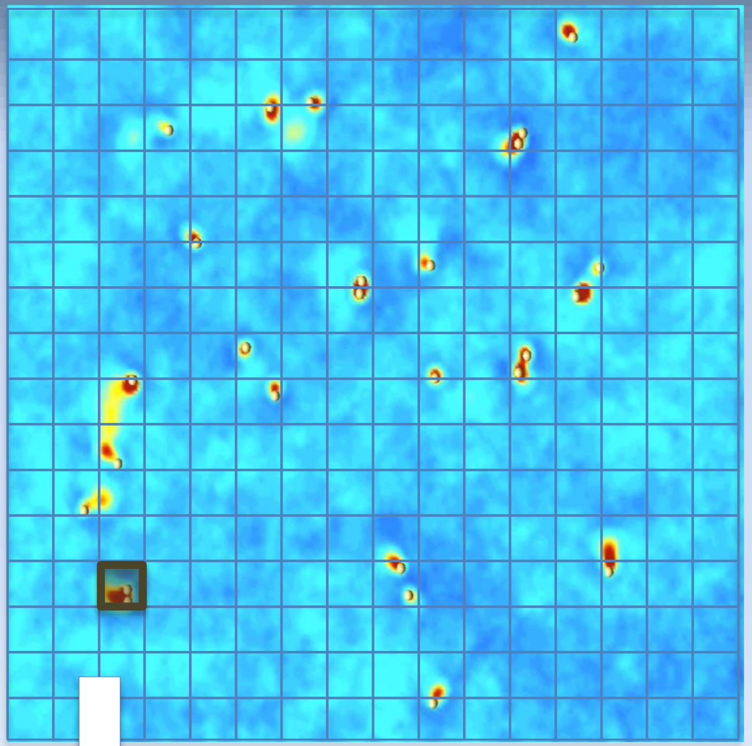


From macroscopic to molecular modeling

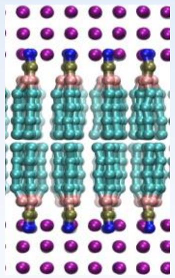
1. **Identify regions of interest using AI techniques**



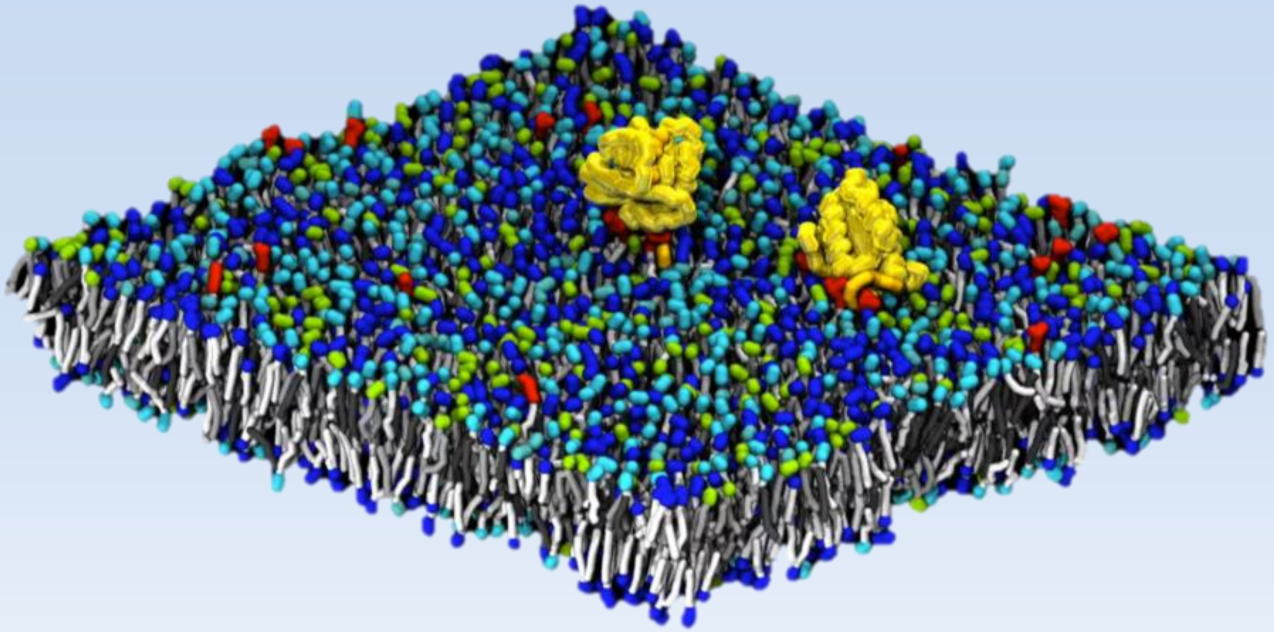
From macroscopic to molecular modeling



1. **Identify region**

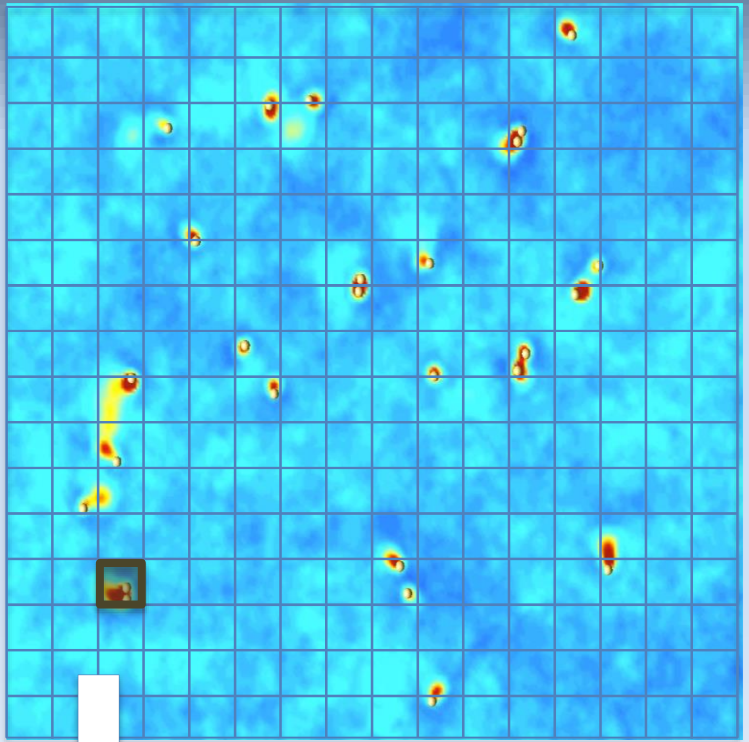


2. **Generate particle positions**

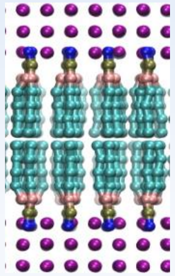


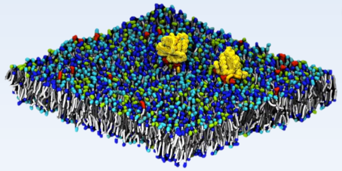
- Lipid positions are generated consistent with composition using *insane* membrane building tool
- Ras proteins, if present, are inserted with appropriate conformation
- System is initialized using GROMACS (CPU-only)

From macroscopic to molecular modeling



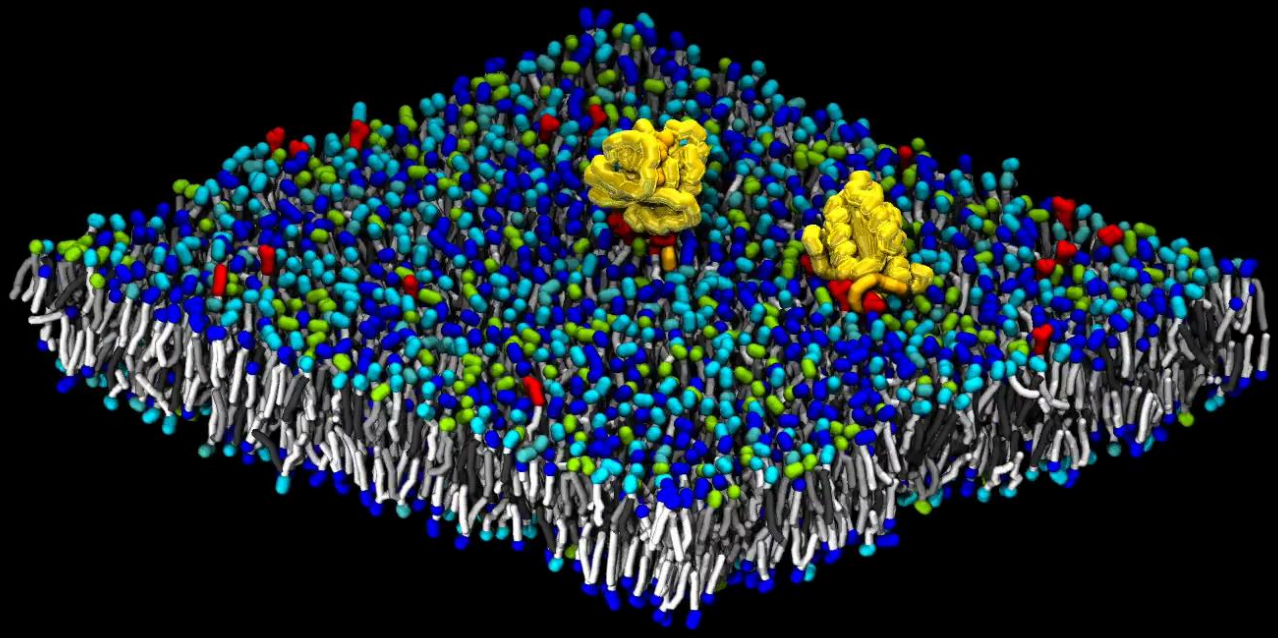
1. Identify region



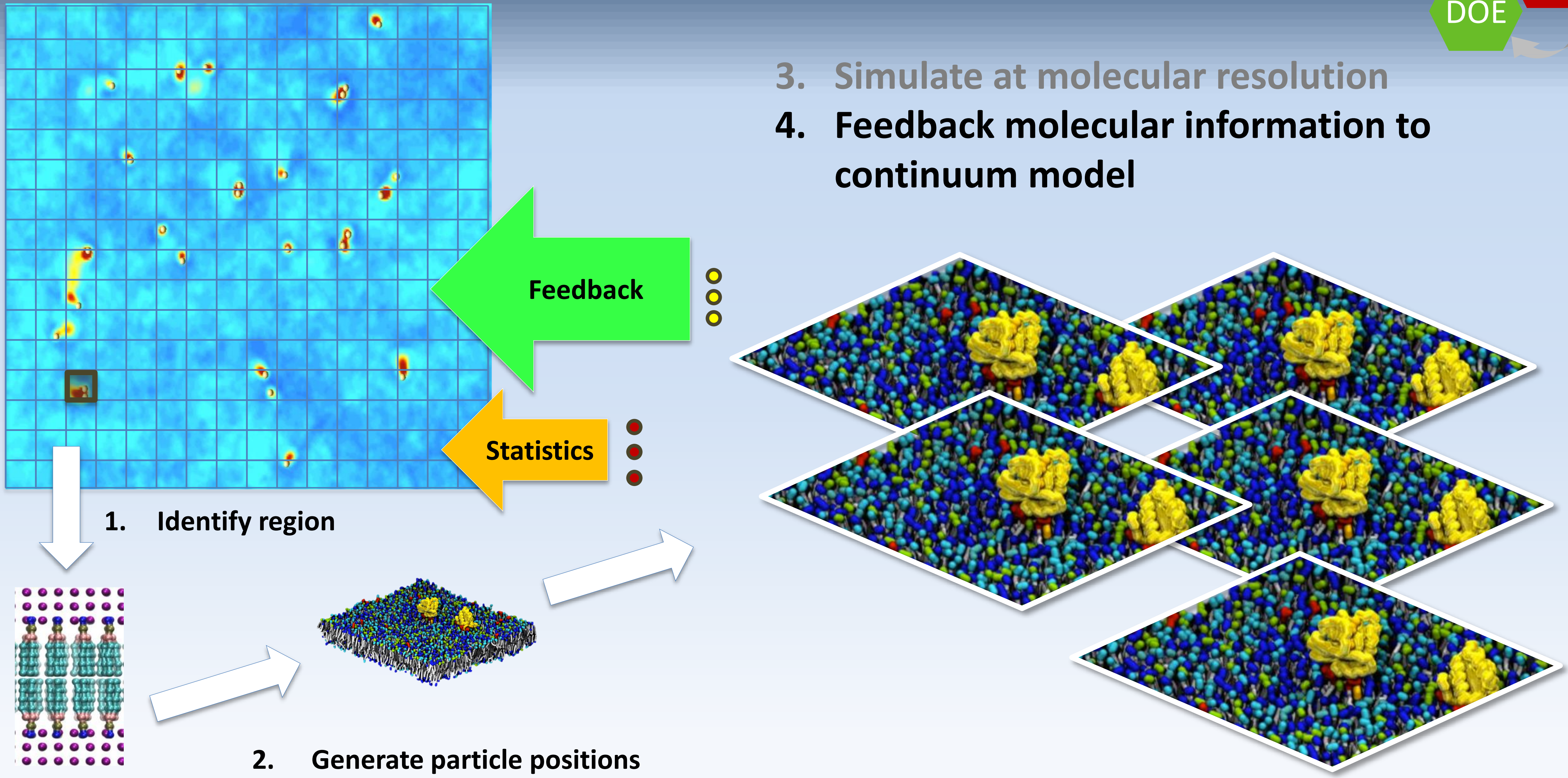


2. Generate particle positions

3. Simulate at molecular resolution using ddcMD

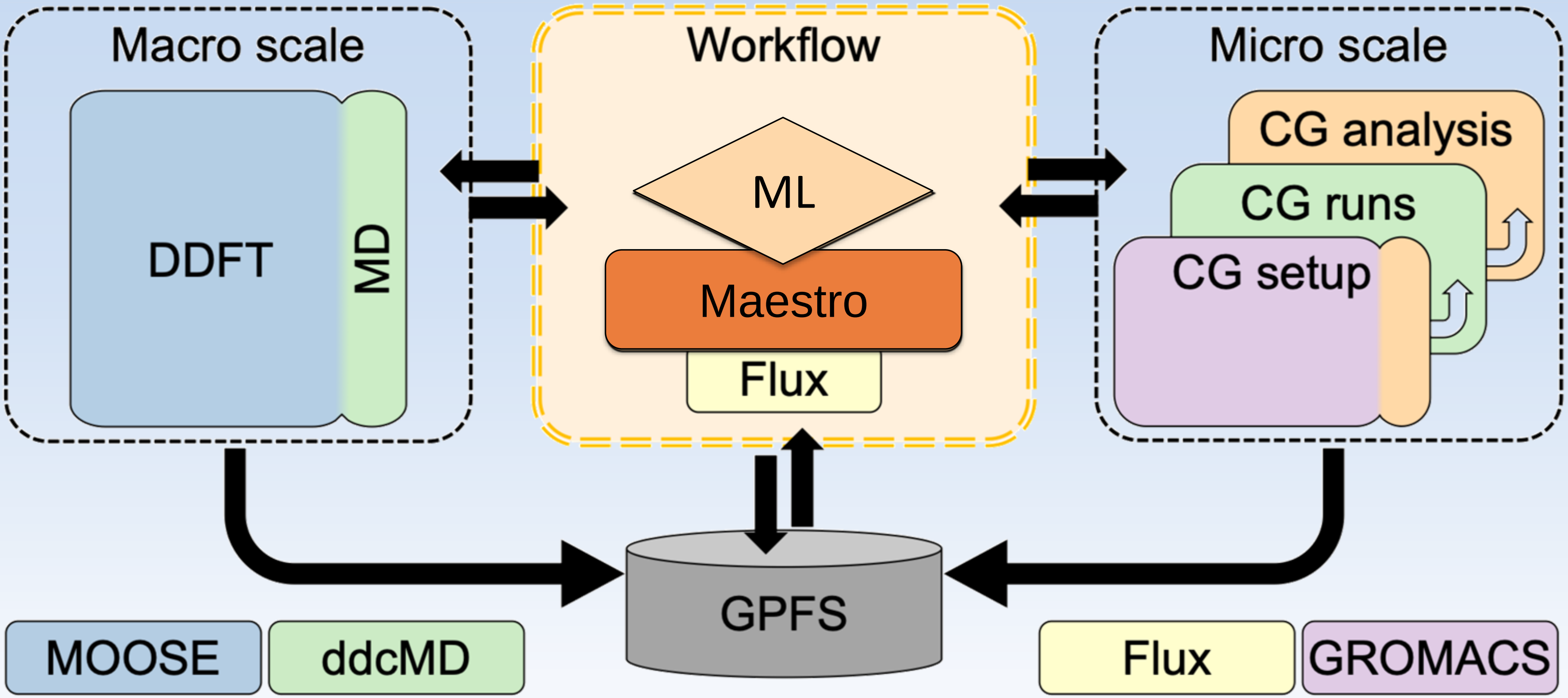


From macroscopic to molecular modeling



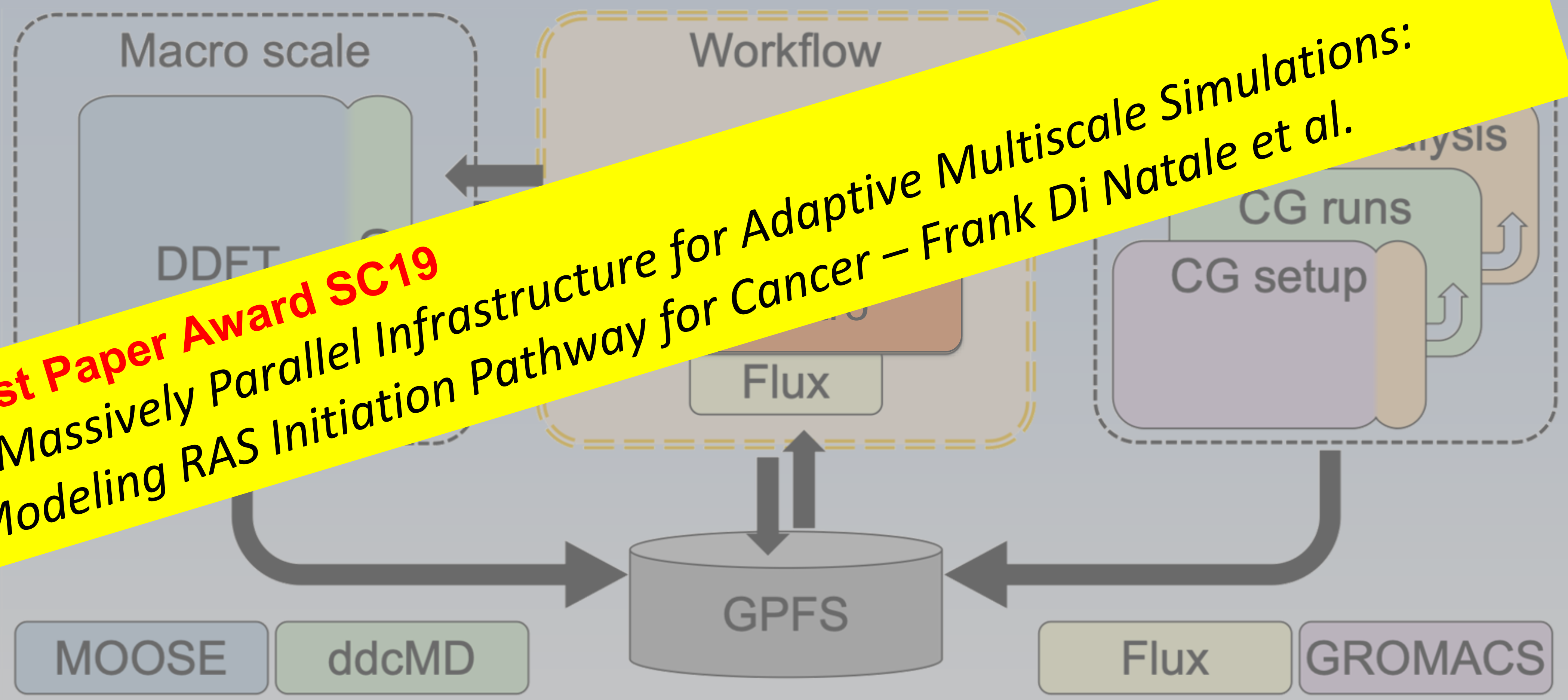
1. Identify region
2. Generate particle positions
3. Simulate at molecular resolution
4. Feedback molecular information to continuum model

Connecting different scales using multiple codes across a heterogenous architecture required a complex workflow



Connecting different scales using multiple codes across a heterogenous architecture required a complex workflow

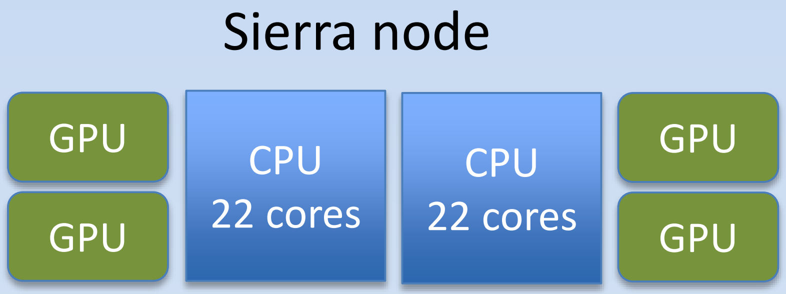
NCI
DOE

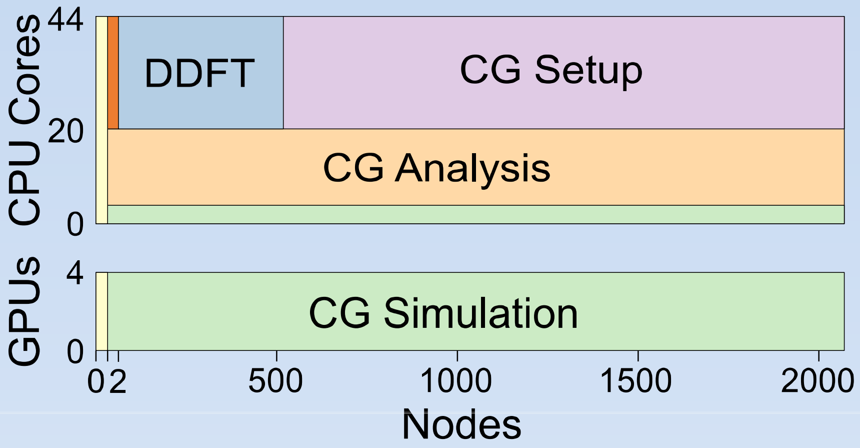


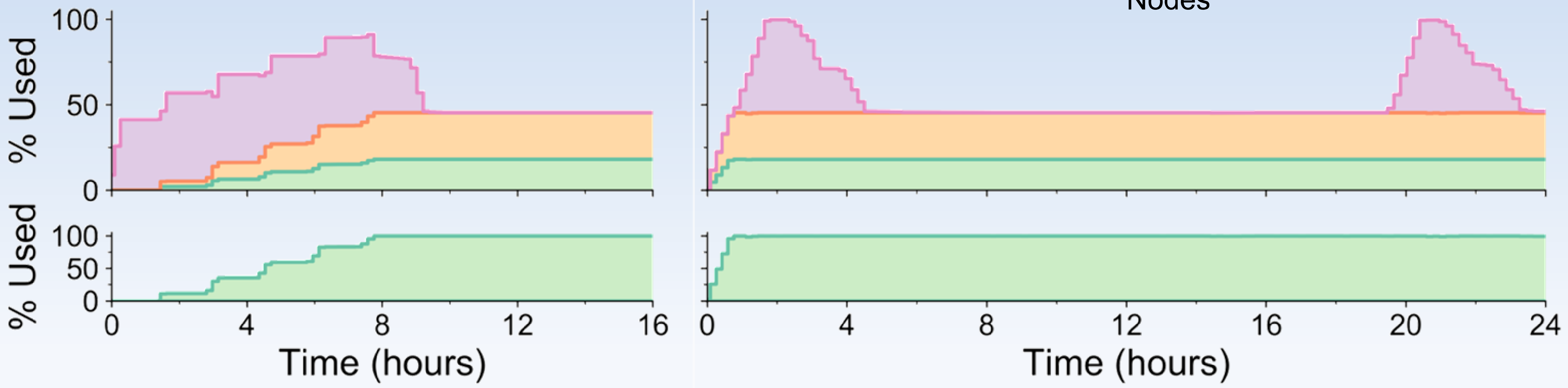
Best Paper Award SC19

A Massively Parallel Infrastructure for Adaptive Multiscale Simulations: Modeling RAS Initiation Pathway for Cancer – Frank Di Natale et al.

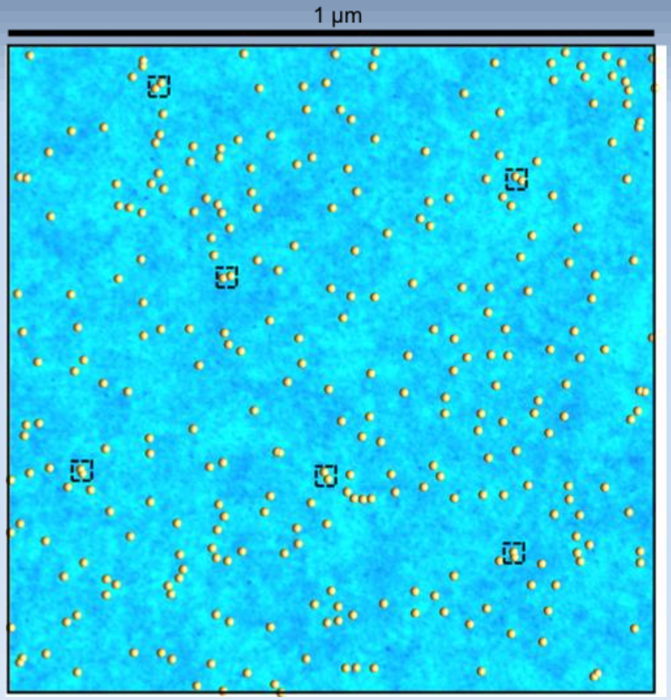
Maestro and Flux enable effective use of Sierra architecture throughout campaign workflow





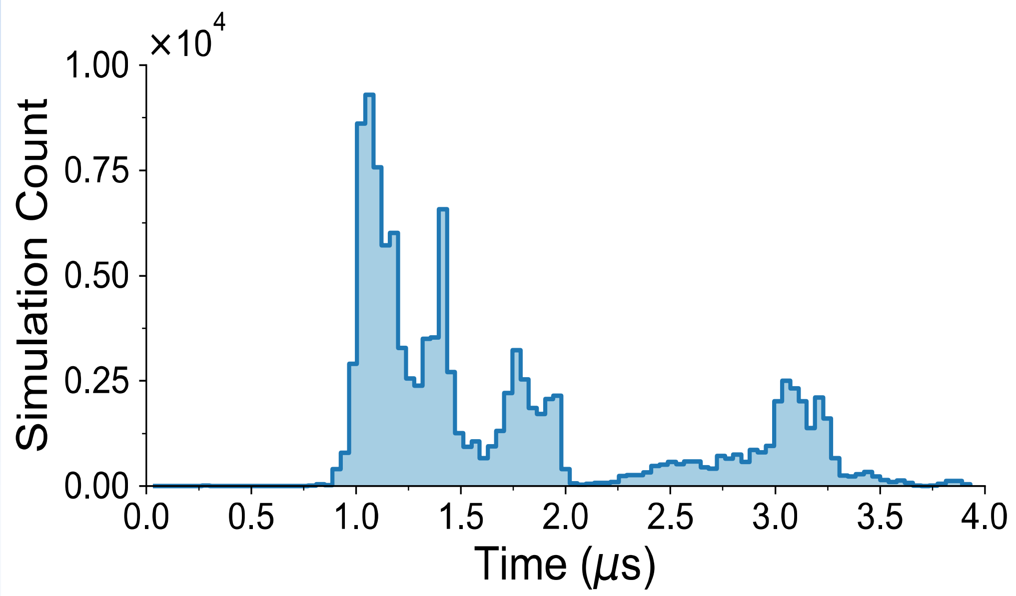


Campaign 1 – RAS membrane dynamics



DDFT simulation:
- 1 µm^2
- 300 KRAS
- 152 µs
- 2,200,000 patches

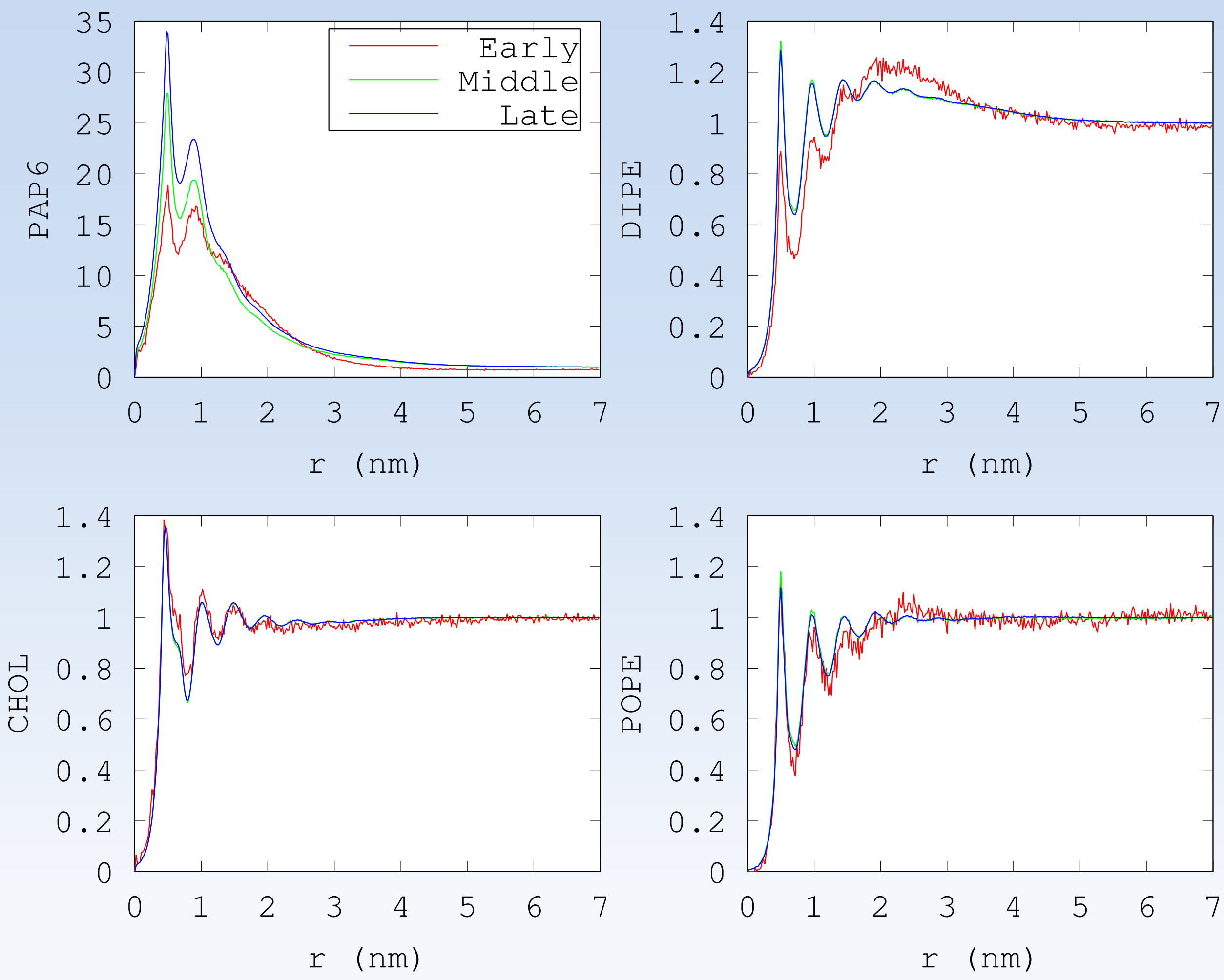
CG simulations:
- 30x30 nm^2
- 140,000 particles
- 116,008 simulations
- 1-4 µs each
- 200 ms aggregated
- *9.998 Trillion MD steps*



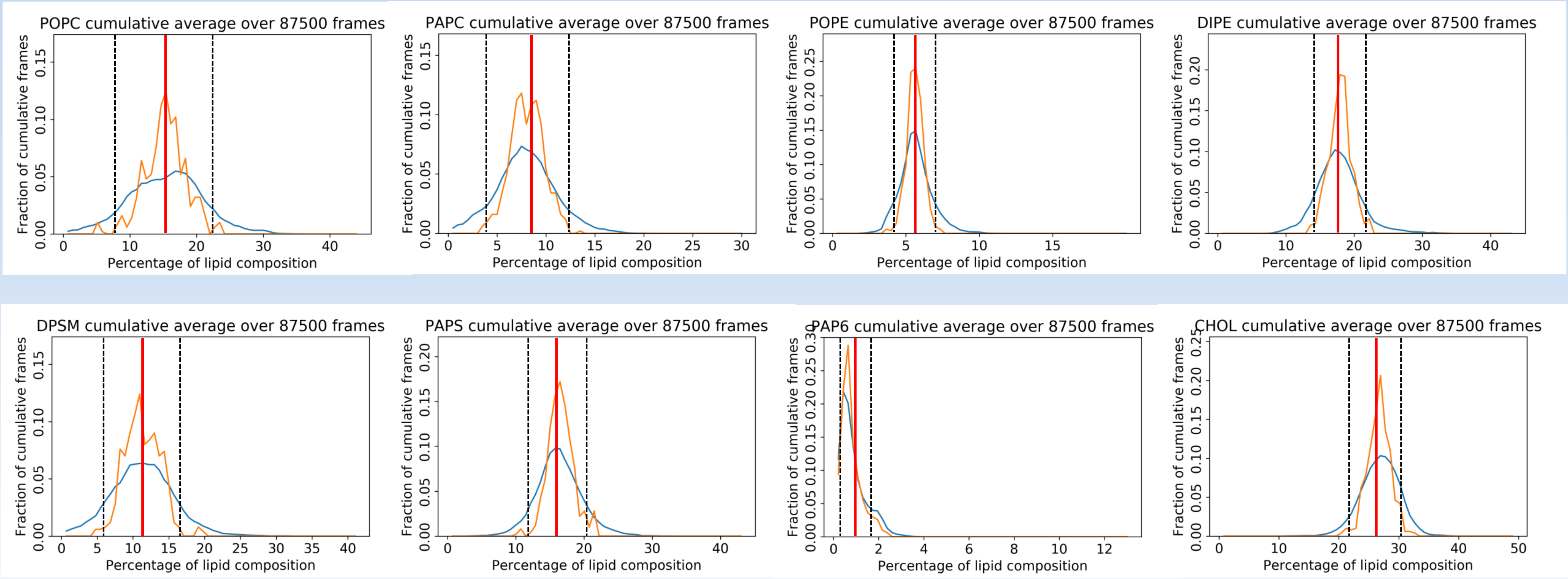
Feedback improves macro model parameters in running simulation

- Simulations initially parameterized with RDF's calculated from single RAS
- Full campaign reveals significant change in PAP6 RDF as RAS aggregate
- Feedback mechanism adapted simulation parameters accordingly
- Minor adjustments seen in other lipid RDF's

RAS-lipid RDFs at three times during campaign



Massive simulation campaign explores complex membrane environment



Global average

10th and 90th percentiles

Pilot 2 Team

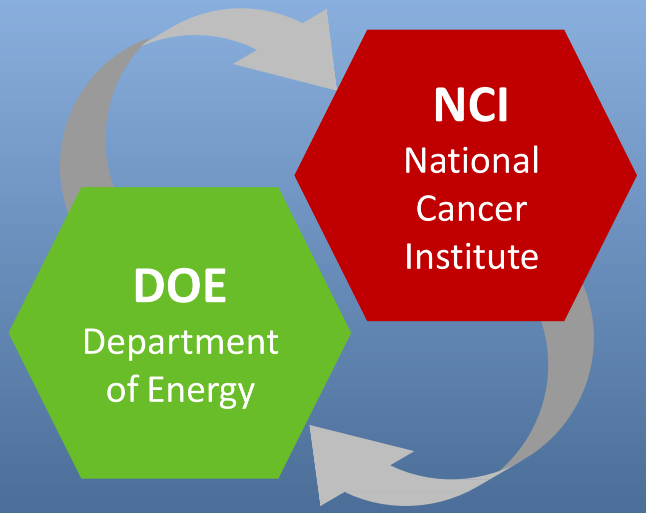
Argonne National Laboratory: Arvind Ramanathan

Lawrence Livermore National Laboratory: Ryan Berg, Harsh Bhatia, Timo Bremer, Tim Carpenter, Gautham Dharuman, Francesco Di Natale, Jim Glosli, Helgi Ingolfsson, Felice Lightstone, Tomas Oppelstrup, Fred Streitz, Brian Van Essen, Xiaohua Zhang

Los Alamos National Laboratory: Boian Alexandrov, Angel Garcia, Nick Hengartner, Jeevapani Hettige, Christoph Jungans, Cesar Lopez, Chris Neale, Sandrasegaram Gnanakaran

Frederick National Laboratory for Cancer Research: Debanjan Goswami, Gulcin Gulten, Dwight Nissley, Rebika Shrestha, Andrew Stephen, Tommy Turbyville, Que Van

Oak Ridge National Laboratory: Debsindhu Bhowmik, Chris Stanley



Lawrence Livermore
National Laboratory