



DIGITAL TWINS FOR SCIENCE

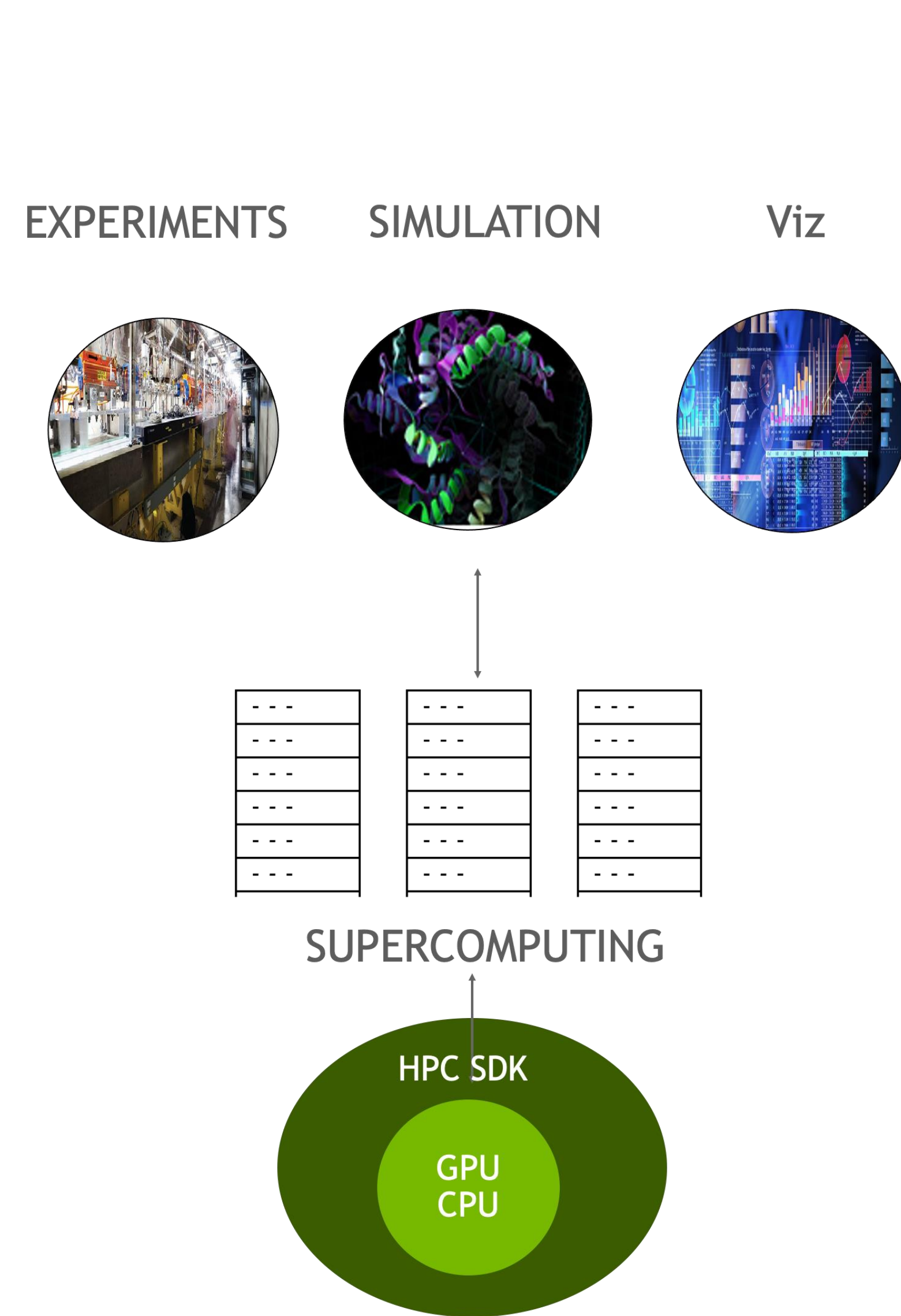
SCIENTIFIC COMPUTING IS EVOLVING

Thought Leaders Mapping Out the Opportunity and Constraints Given Current Market Reality



A DEEPER LOOK AT THE NEW HPC

ADAPTING TO THE CONSTRAINTS, MARKET REALITIES AND EXPANDING OPPORTUNITY



[ACM HPC Forecast
Reed and Dongarra](#)

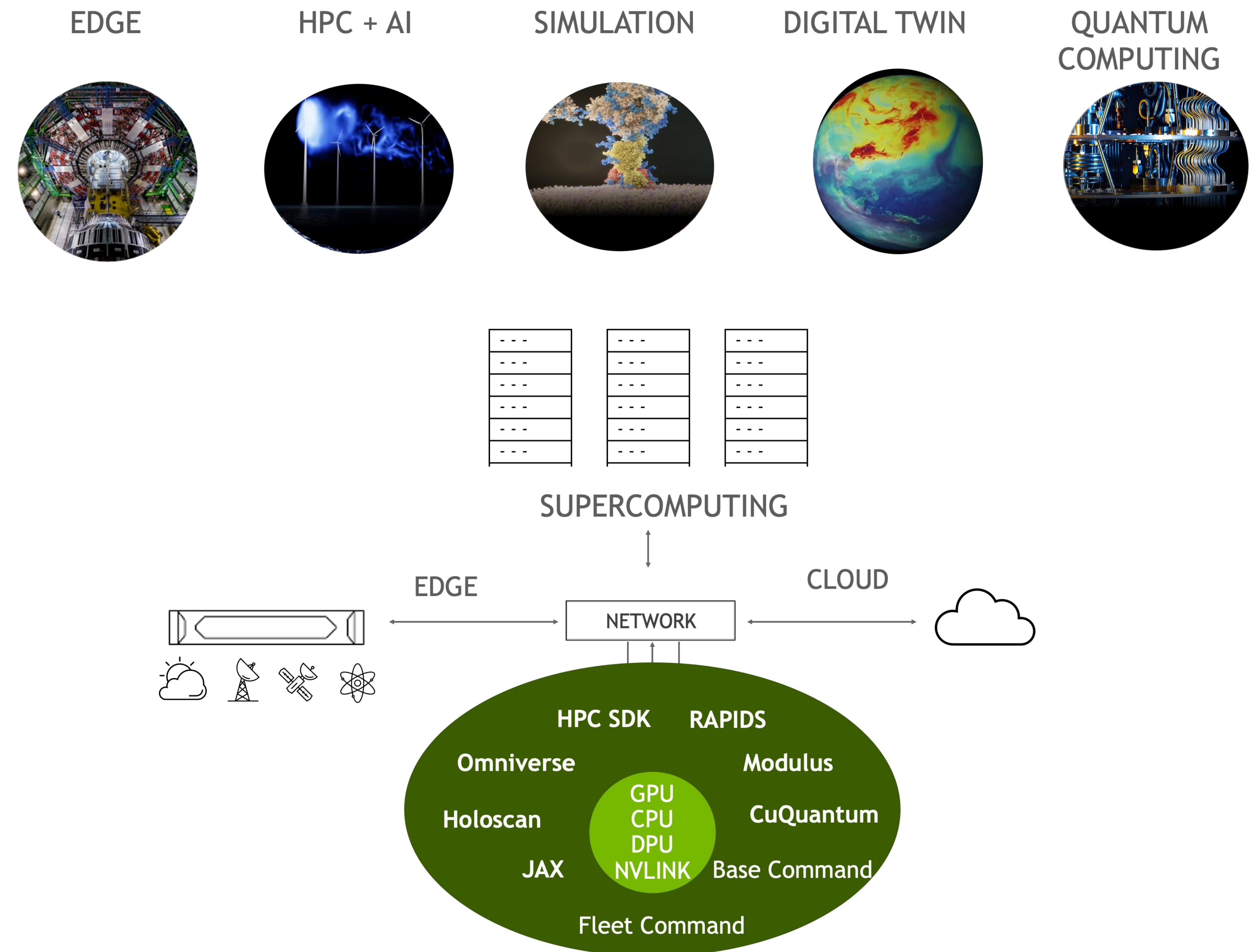
Semiconductor Constraints limit the potential increase in scale for legacy algorithms

New Algorithms offer potential for dramatic increase in model scale and reduction in latency

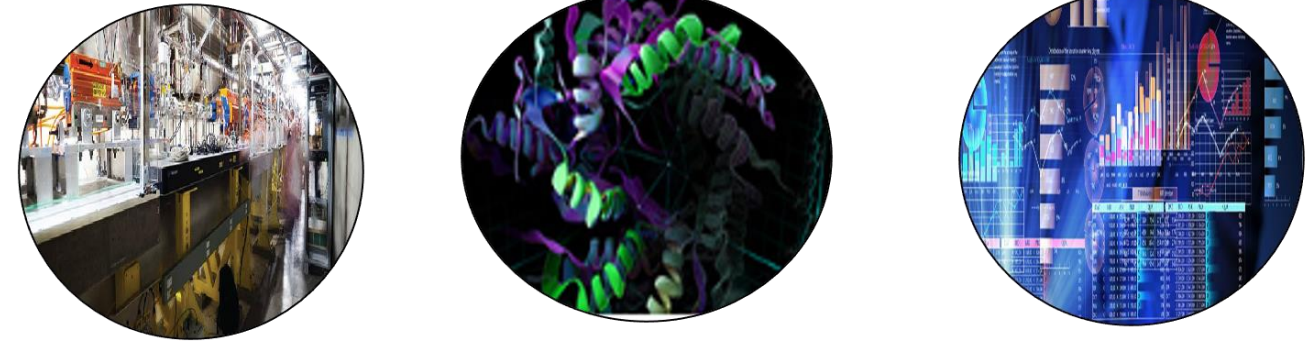
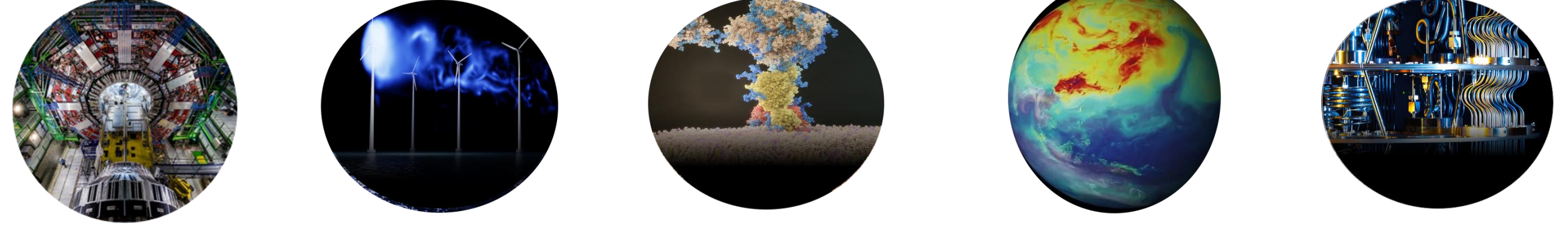
Cloud Economics have changed the supply chain ecosystem

[Charting a New Path Post Exascale Computing NNSA](#)

OVERARCHING FINDING: The combination of increasing demands for computing with the technology and market challenges in HPC requires an intentional and thorough reevaluation of algorithms, software development, system design, computing platform acquisition, and workforce development.



TRANSITION TO POST EXASCALE ERA

FEATURE	<p>EXPERIMENTS SIMULATION Viz</p>  <p>TERA THROUGH EXASCALE</p>	<p>EDGE HPC + AI SIMULATION DIGITAL TWIN QUANTUM COMPUTING</p>  <p>POST EXASCALE</p>
USAGE	BATCH & MOSTLY LOCAL TO A SITE	INTERACTIVE & DISTRIBUTED WITH MULTIPLE SITES
WORKLOAD	SINGLE SIMULATION/ENSEMBLE	WORKFLOW COMPRISED OF SIMULATION ENSEMBLES, AI TRAINING AND INFERENCE, LIVE DATA ANALYTICS
EXPERIMENTS	OFFLINE DATA ANALYSIS FOR EXPERIMENTS	MIX OF REAL-TIME ANALYSIS, STEERING AND OFFLINE
DIGITAL TWINS	IN-SITU VISUALIZATION OFFLINE	INTERACTIVE COMBINATION OF SIMULATION AND OBSERVATIONAL DATA
QUANTUM COMPUTING	SIMULATION	SIMULATION PREPARING FOR A HYBRID MODEL
PROGRAMMING MODELS	FORTRAN, C++, MPI, OPENMP	STANDARD PARALLELISM SUPPORT IN FORTRAN, C++, MPI, OPENMP, OPENACC, PYTHON, JULIA, PYTORCH, TENSORFLOW
SYSTEM CONFIGURATION	MONOLITHIC	MODULAR
CLOUD	GRID	BURST CAPABILITIES, FASTER REFRESH CYCLE, ACCESS TO LATEST TECHNOLOGY AT SCALE

DIGITAL TWIN FOR SCIENCE?

A Relatively New Modeling Concept That is Just Emerging for Science

Digital Twin Definition from Wikipedia

Digital model of an intended or actual real-world physical product, system, or process (a physical twin) that serves as the effectively **indistinguishable digital counterpart of it for practical purposes.**

The concept and model of the digital twin was first publicly introduced in 2002 by Michael Grieves, at a Society of Manufacturing Engineers conference as the conceptual model underlying Product Life Cycle Management

National Academy of Science Report: Foundational Research Gaps and Future Directions:

A digital twin is a set of virtual information constructs that mimics the structure, context, and behavior of a natural, engineered, or social system (or system-of-systems), is dynamically updated with data from its physical twin, has a predictive capability, and informs decisions that realize value.

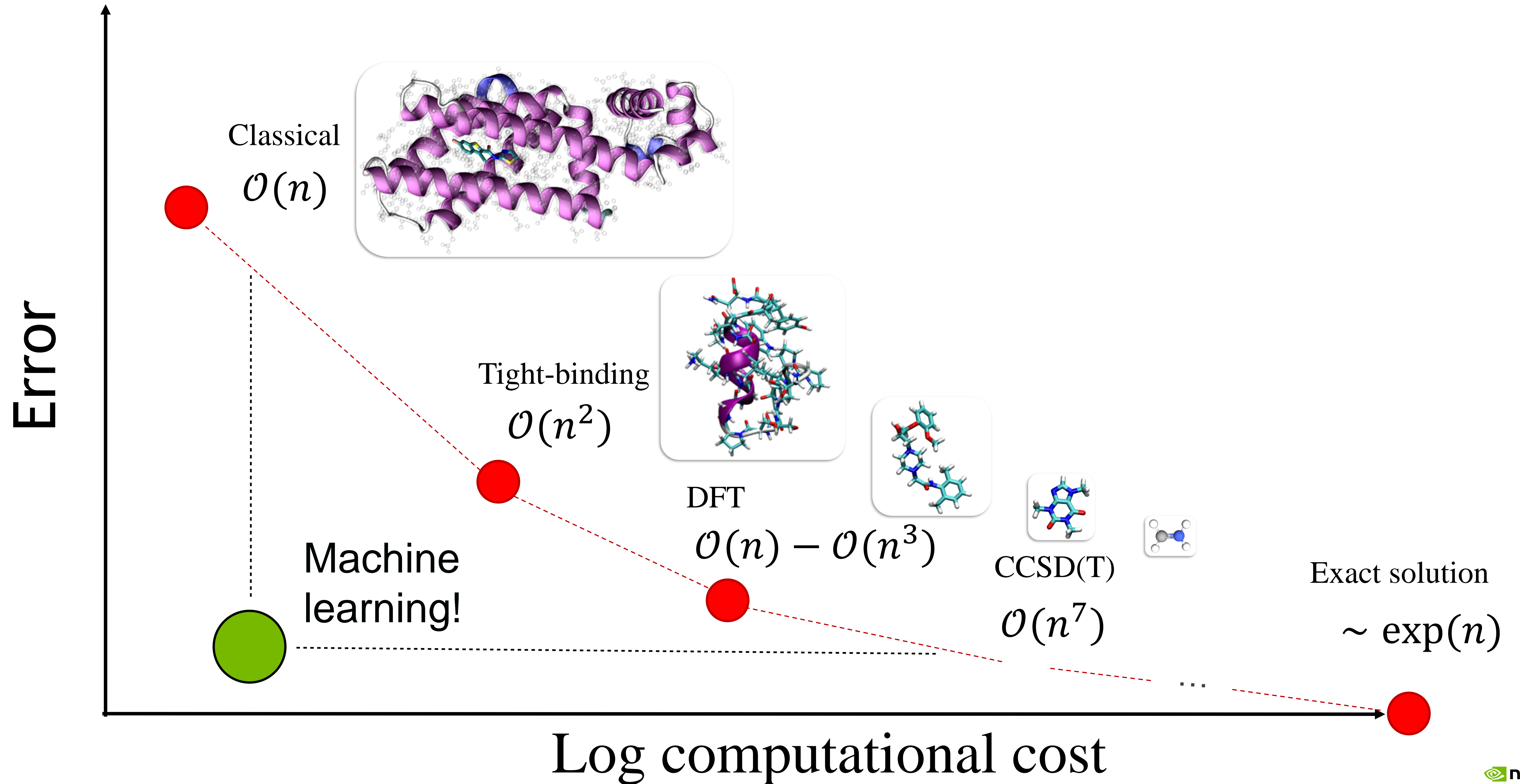
The bidirectional interaction between the virtual and the physical is central to the digital twin.

A Digital Twin for Science

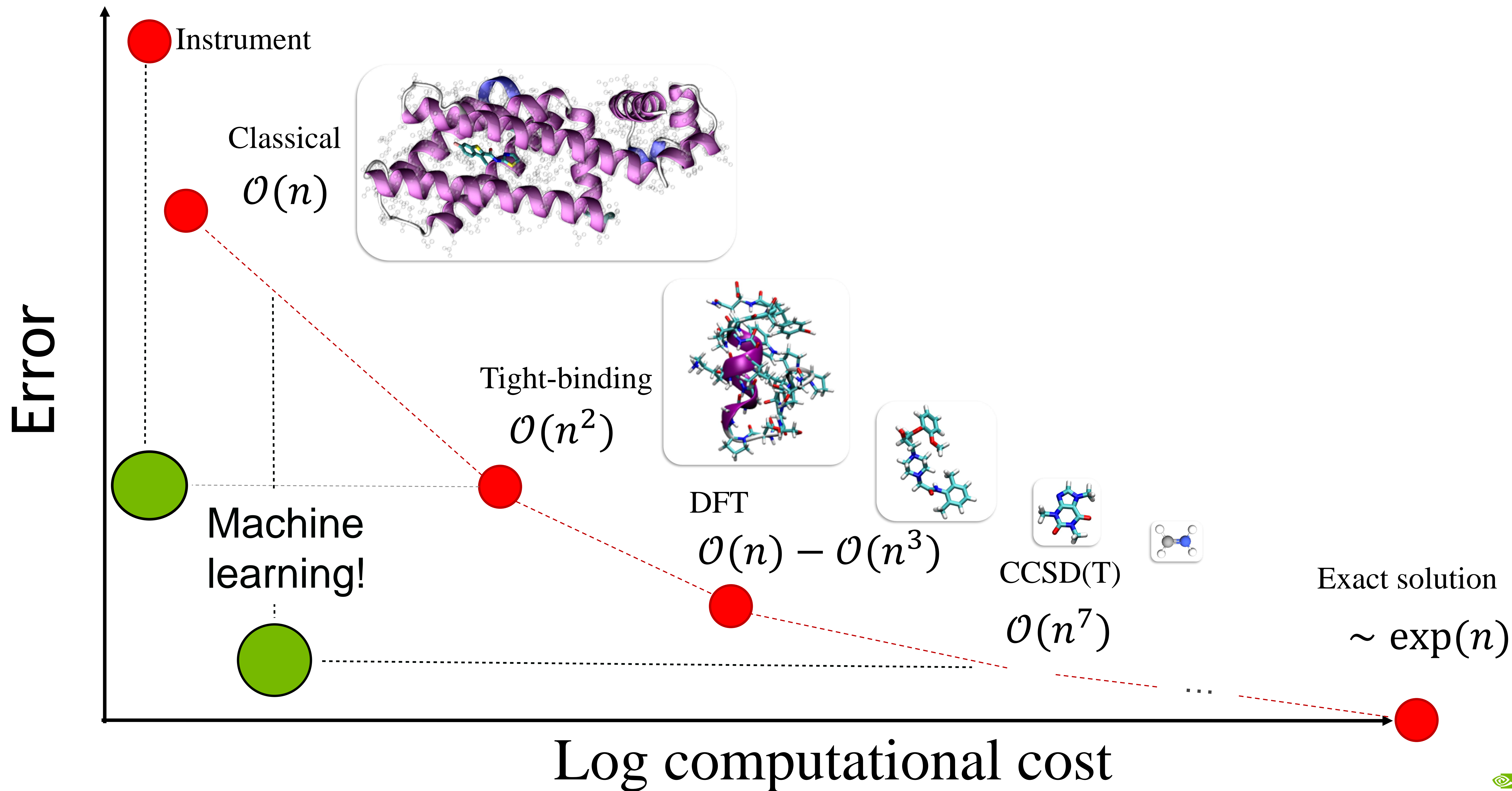
A digital model with sufficient scale and accuracy to be indistinguishable from the real-world physical object, system or process that is dynamically updated with data from its physical twin, where the model has sufficient fidelity to meet the requisite accuracy with time to solution to inform decisions that improve the operation of the asset or process being modeled

ML ENABLES REQUISITE ACCURACY/TIME TO SOLUTION

Allows the Model to be Indistinguishable and Interactive

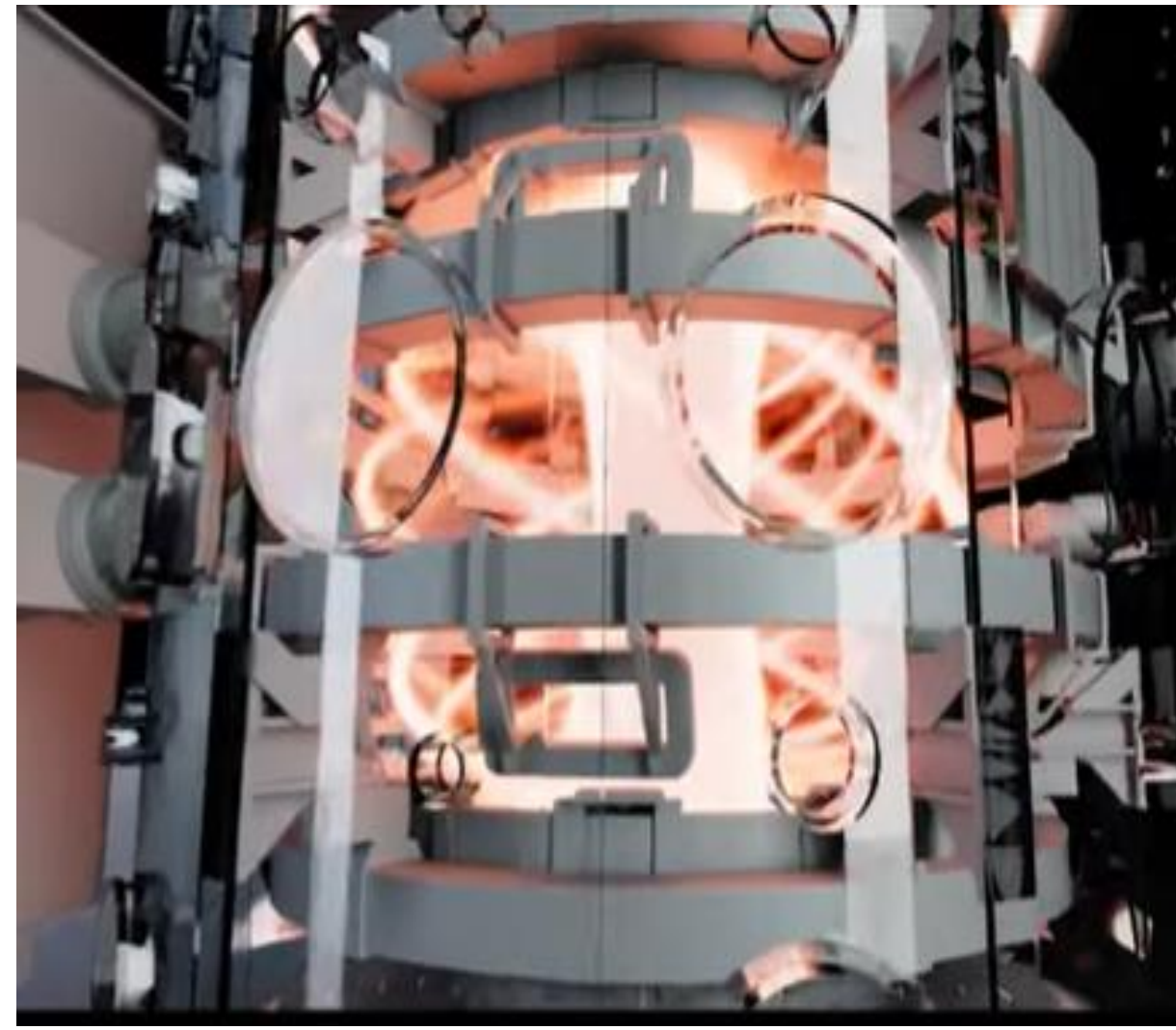


ML POTENTIAL TO BRIDGE MODELS WITH CONTROL



EXAMPLES OF DIGITAL TWINS FOR SCIENCE

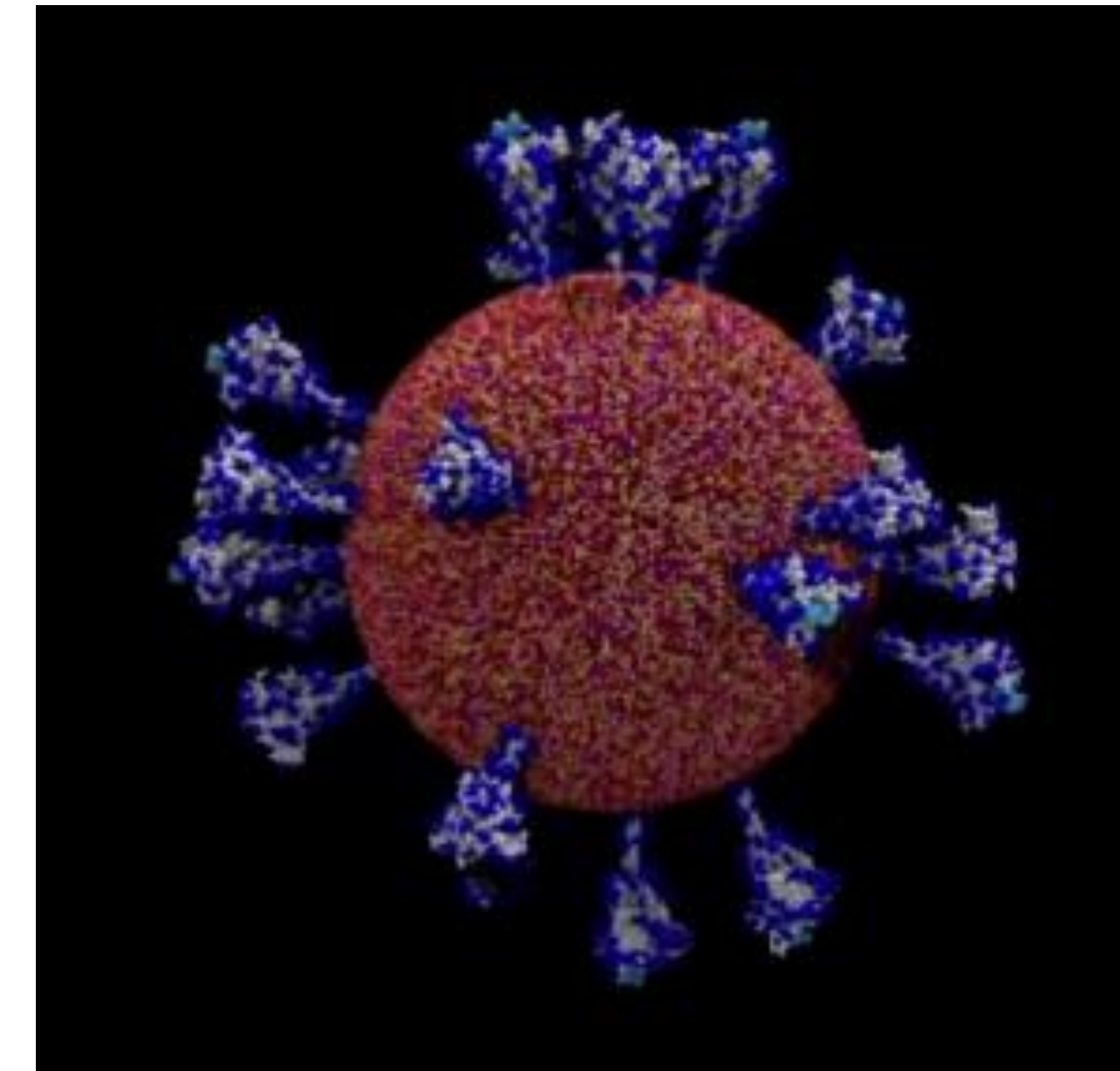
Collaborate with the Global Research Community to Pursue Science Discovery that Benefits Mankind



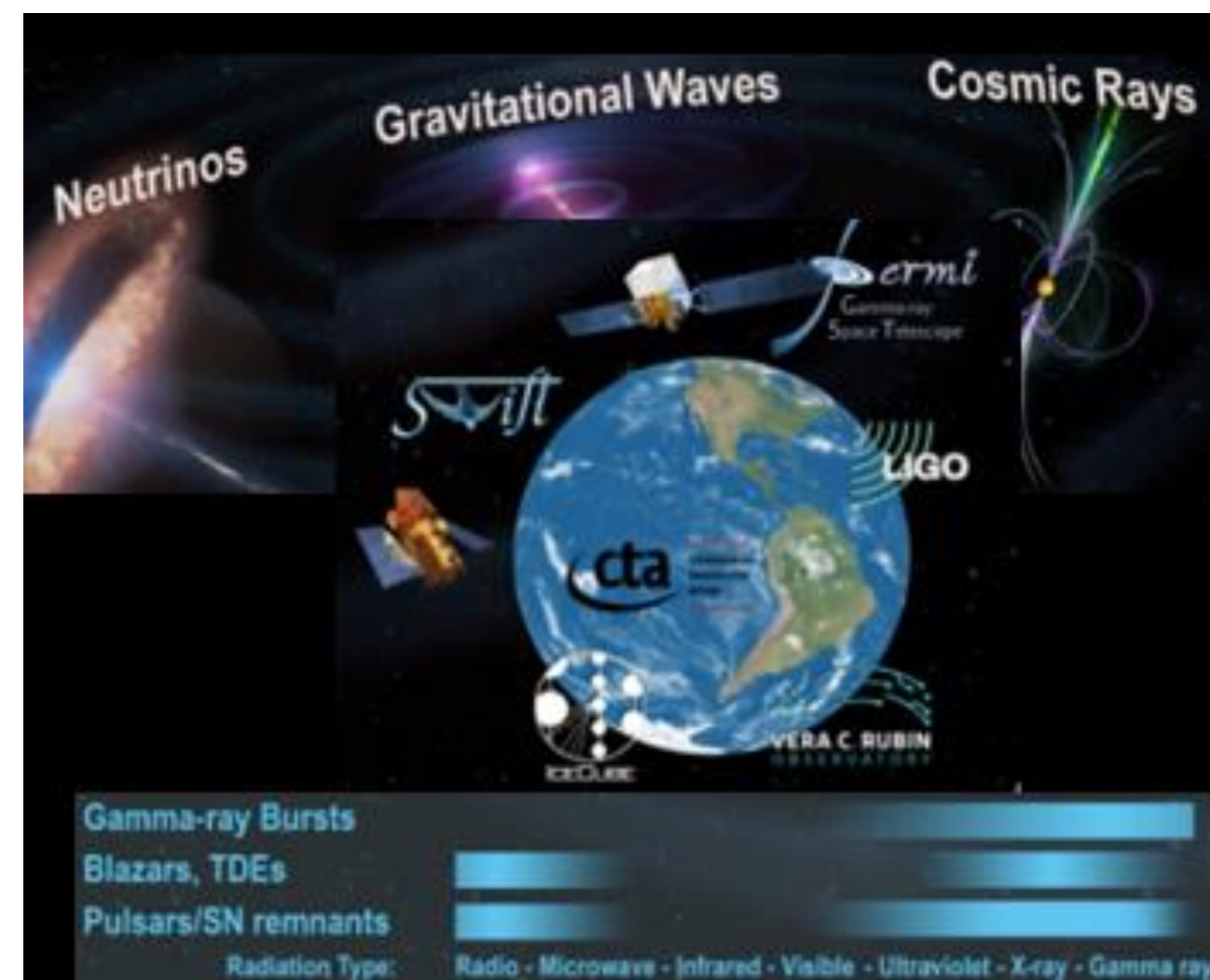
Towards Real time Fusion Reactor Design
Generative AI to Predict Disruption



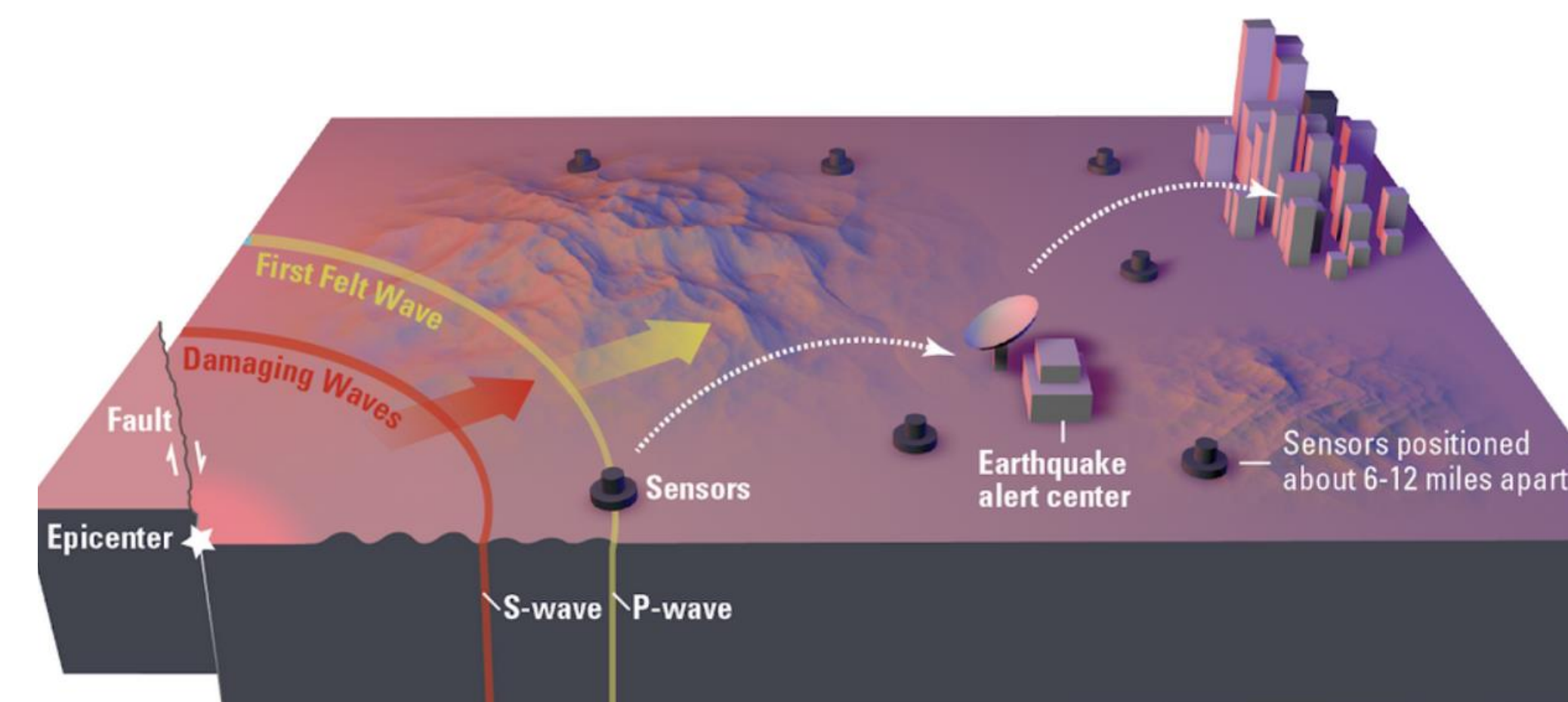
Large Hardon Collider



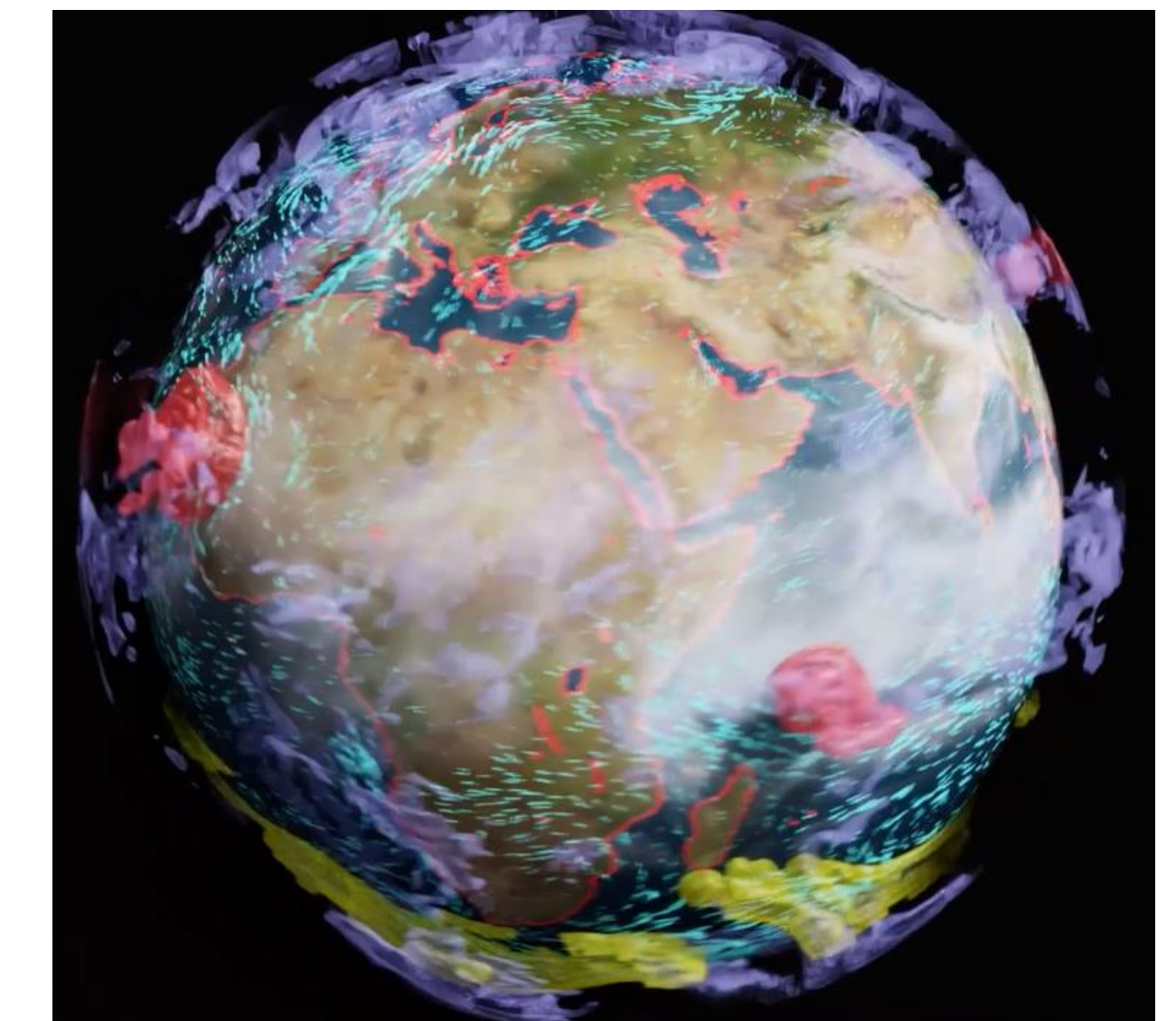
Genome Scale LLMs for Covid



Real Time Multi-Messenger Astrophysics



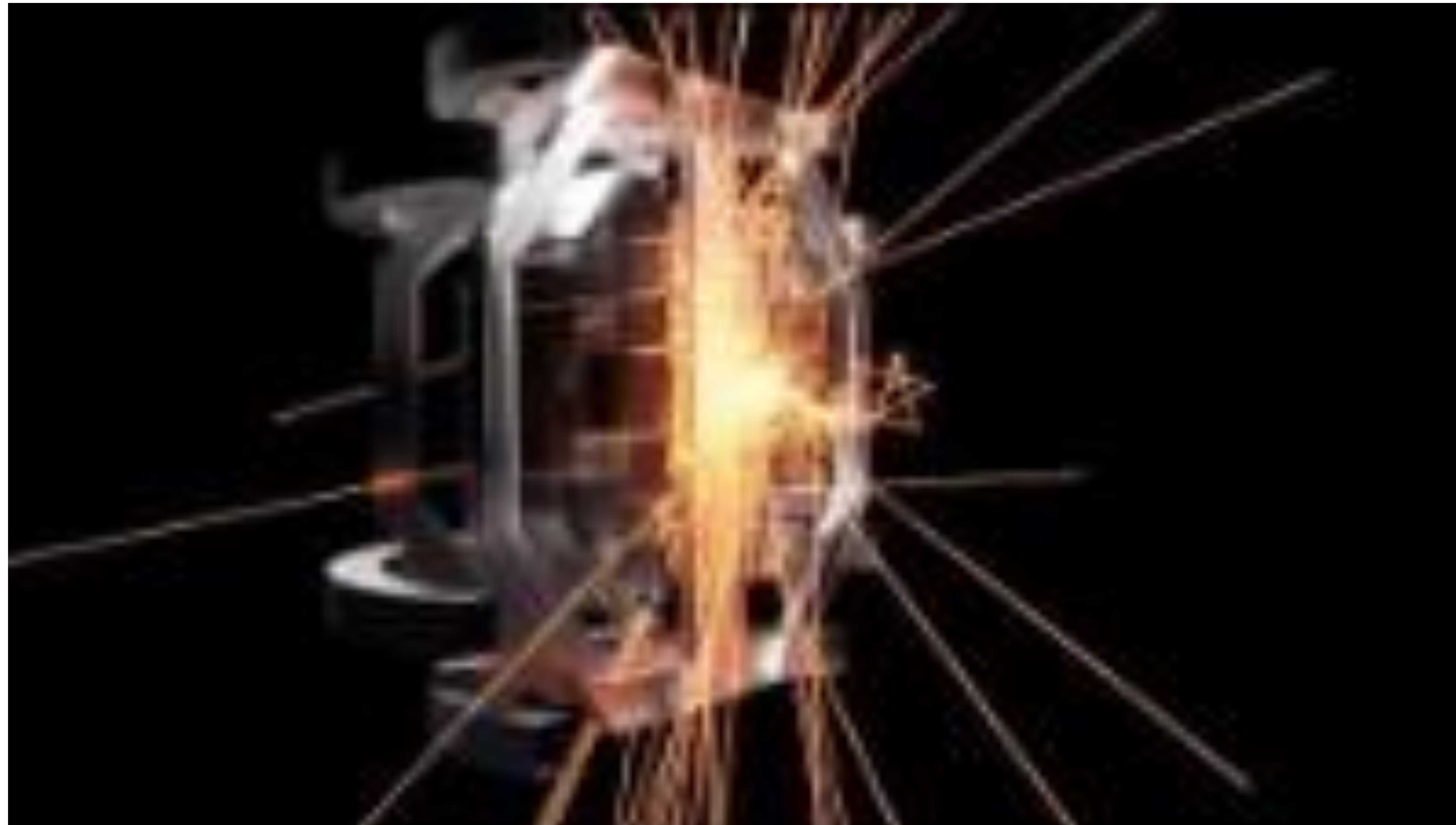
Earthquake Model with Machine Learning



Destination Earth

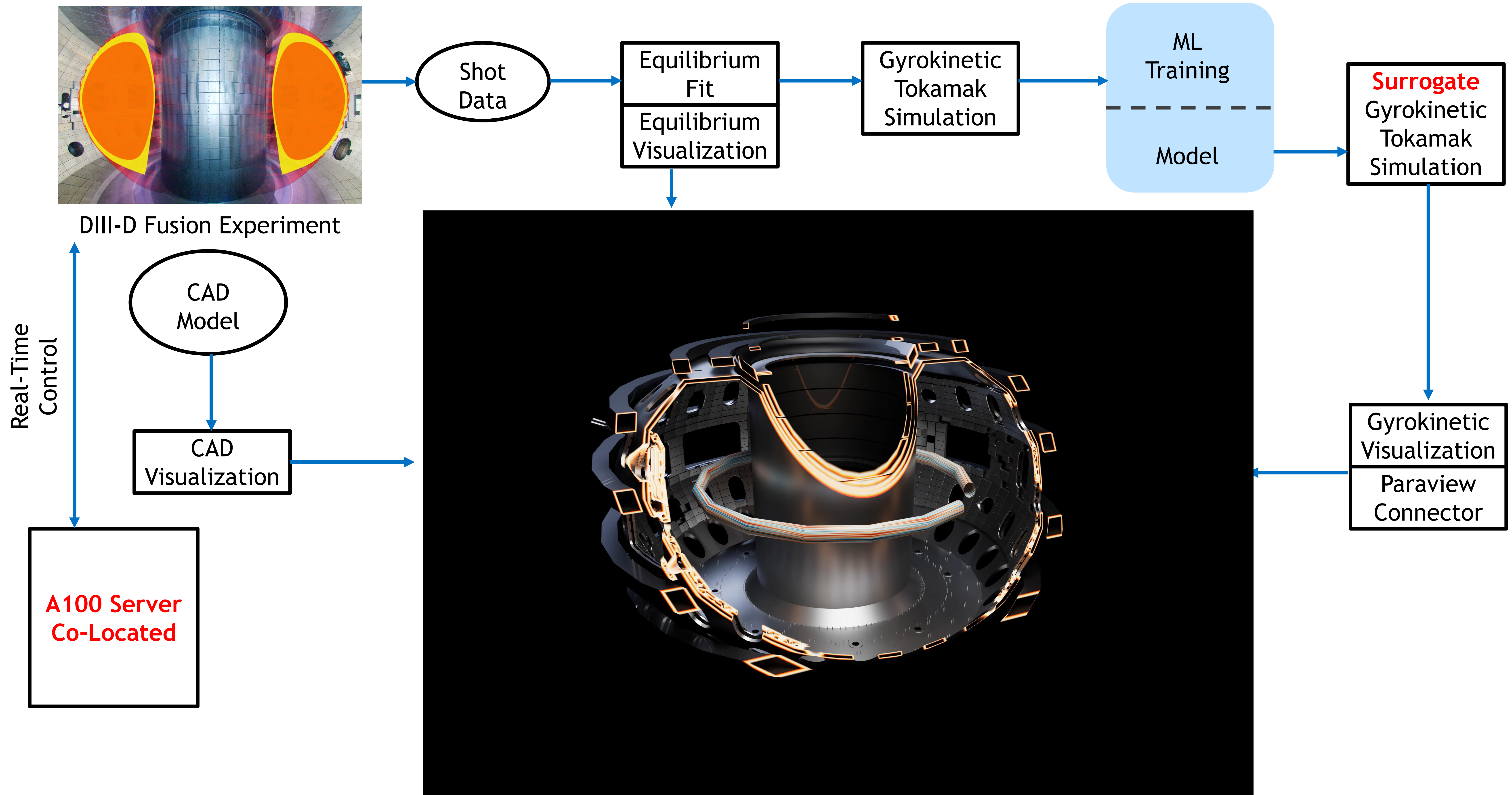
FUSION DIGITAL TWIN WITH CONVENTIONAL SIMULATION

<https://www.youtube.com/watch?v=tJgR1TSBD0k>



FUSION DIGITAL TWIN WORKFLOW WITH SGTC SURROGATE

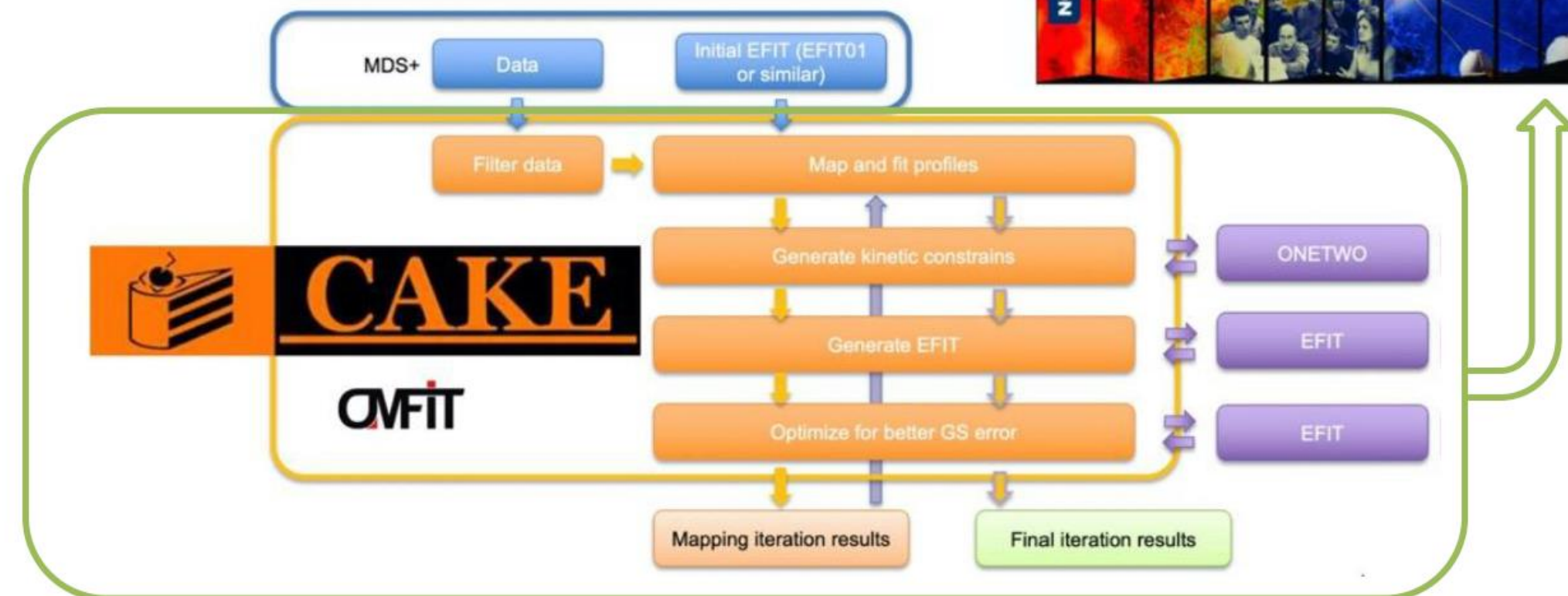
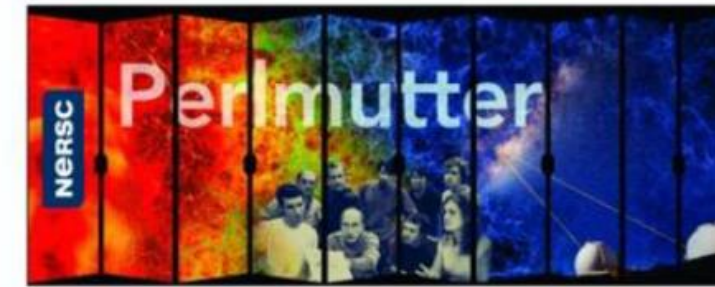
File-Based Prototype



SUPERFACILITY WITH NERSC AND GA DIII-D FUSION REACTOR

Couple EFIT Simulation with Experiment operation

- Dedicated low-latency queue on NERSC
- Test bed for data processing in support of ITER operation

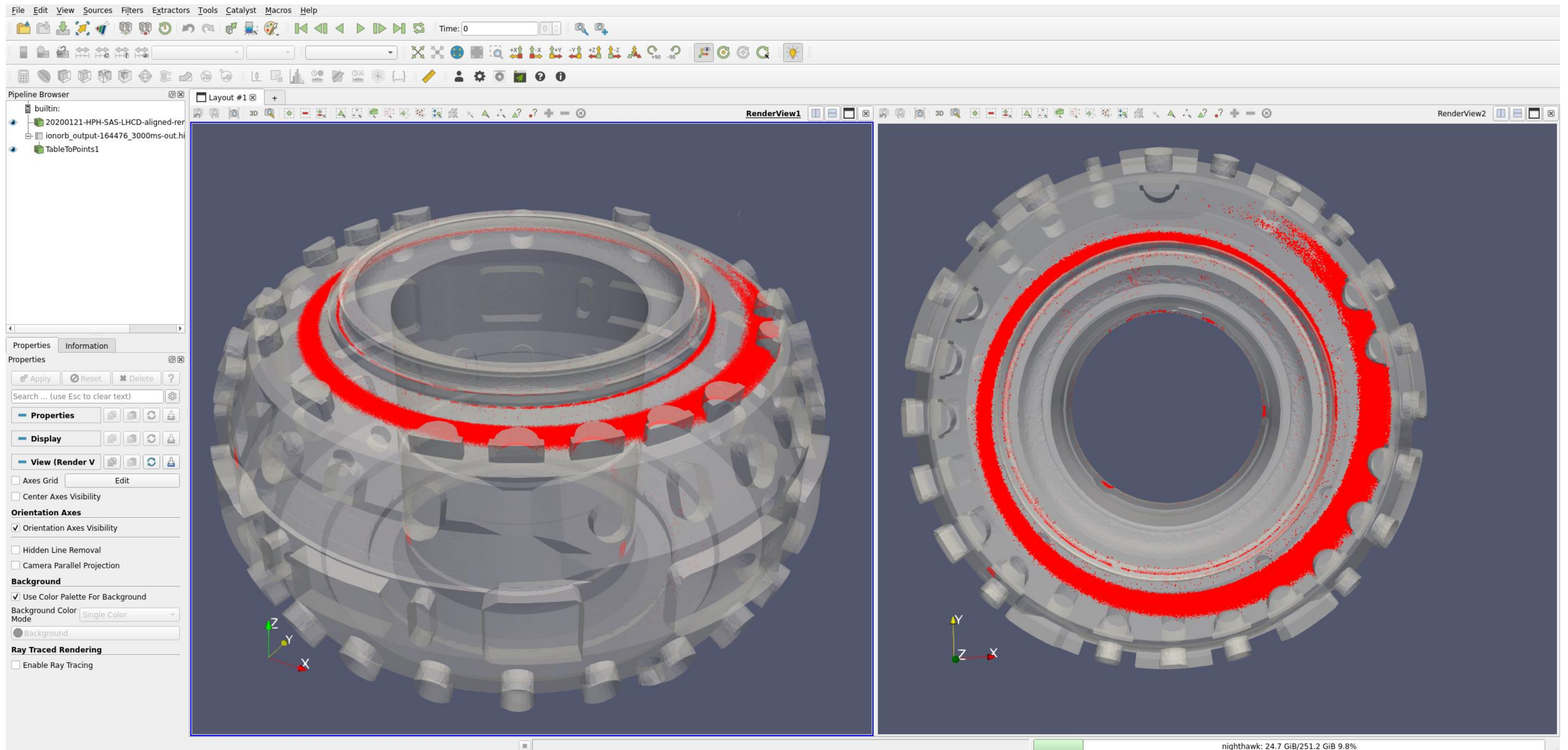


Stefano Smith & Raffi Nazikian/DOE CAKE Superfacility Review / 8 April 2023 - 12

Current Conventional Model with the Twin

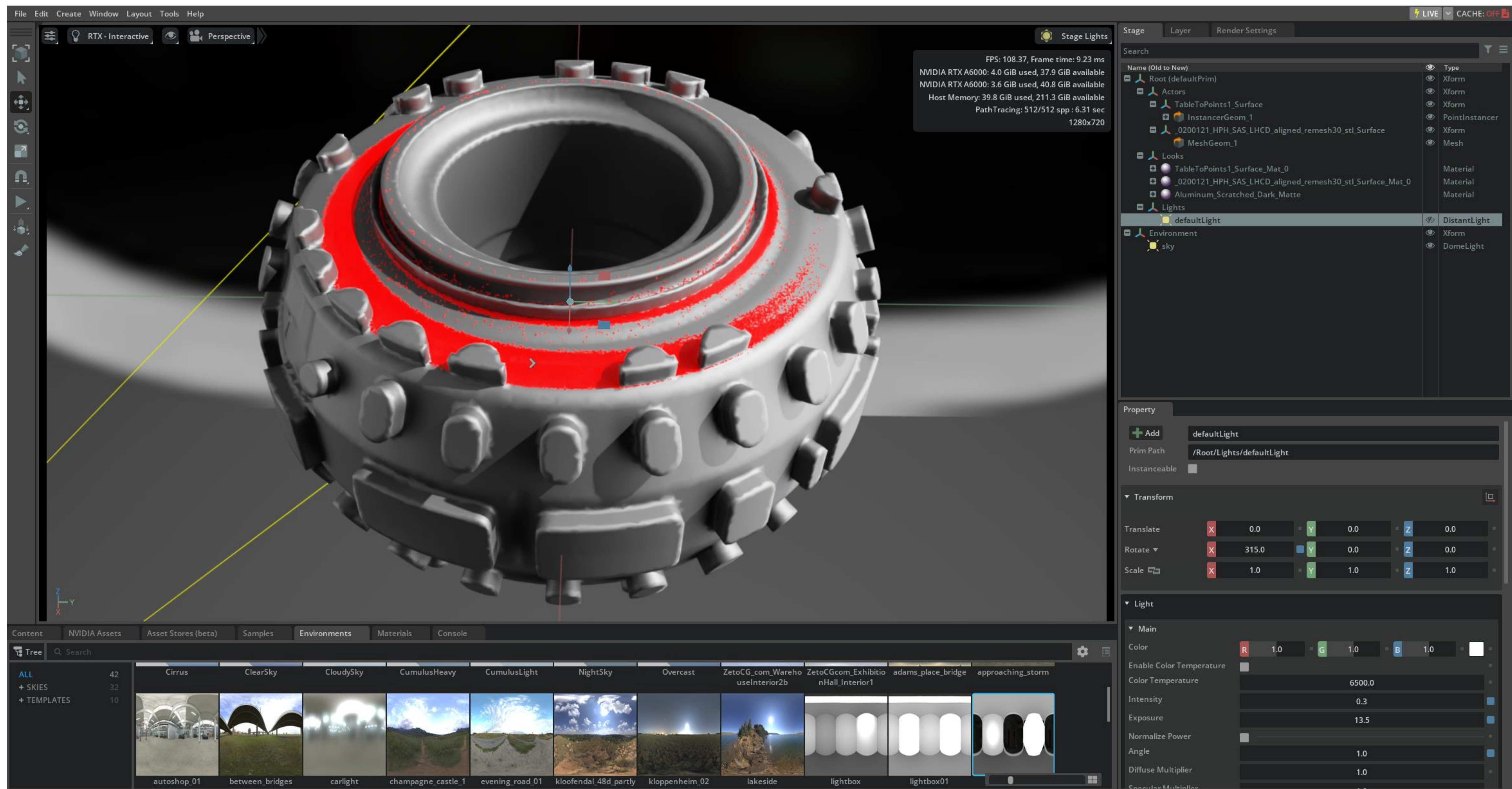


GA Ion Orbiter DIII-D Vacuum Vessel Impacts Postprocessing/Reprojection pipeline -> ParaView



GA Ion Orbiter DIII-D Vacuum Vessel Impacts

Postprocessing pipeline -> ParaView -> Omniverse

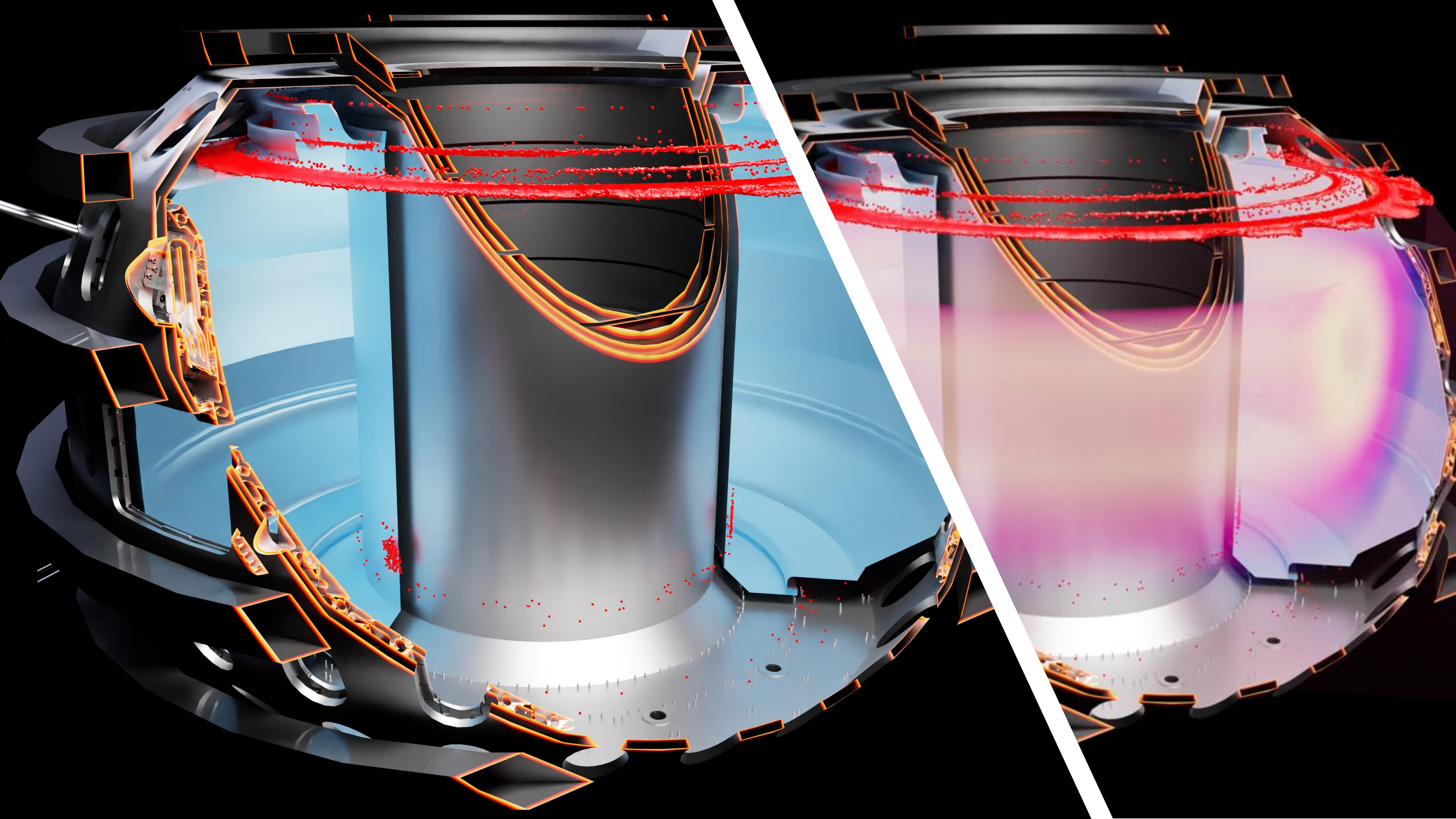




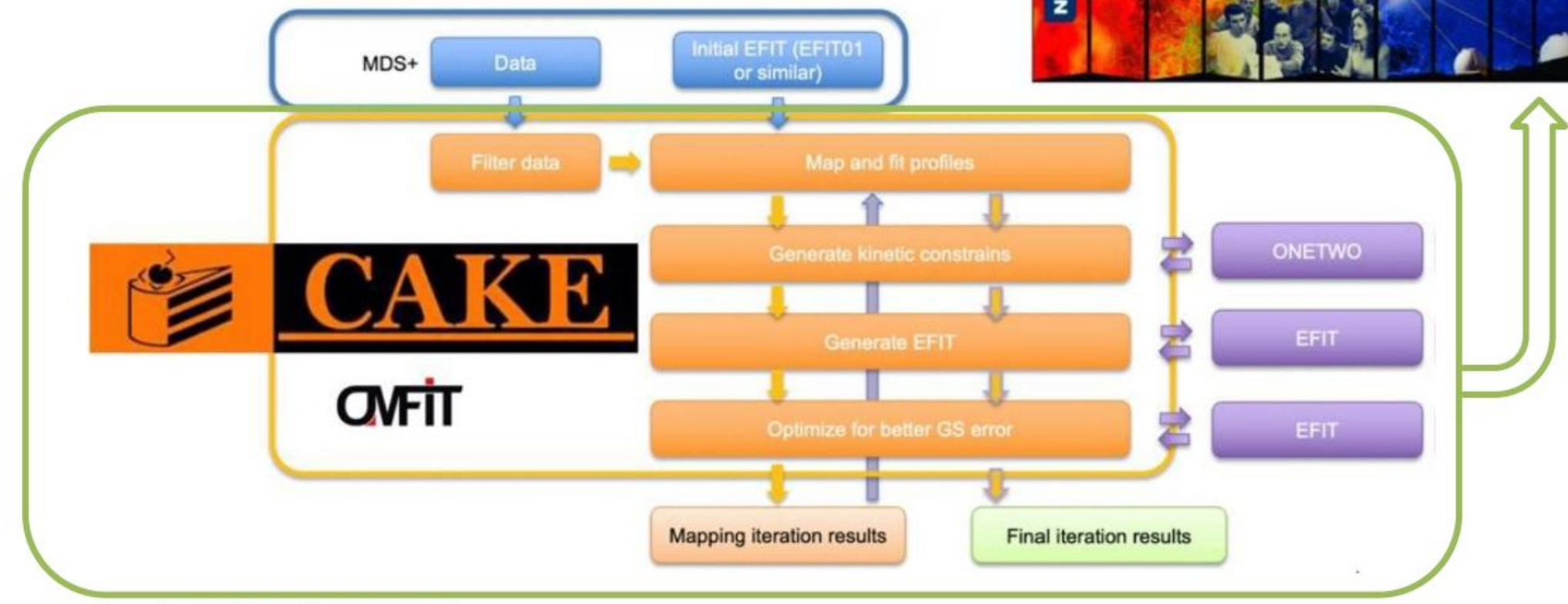
GA Ion Orbiter DIII-D Vacuum Vessel Impacts



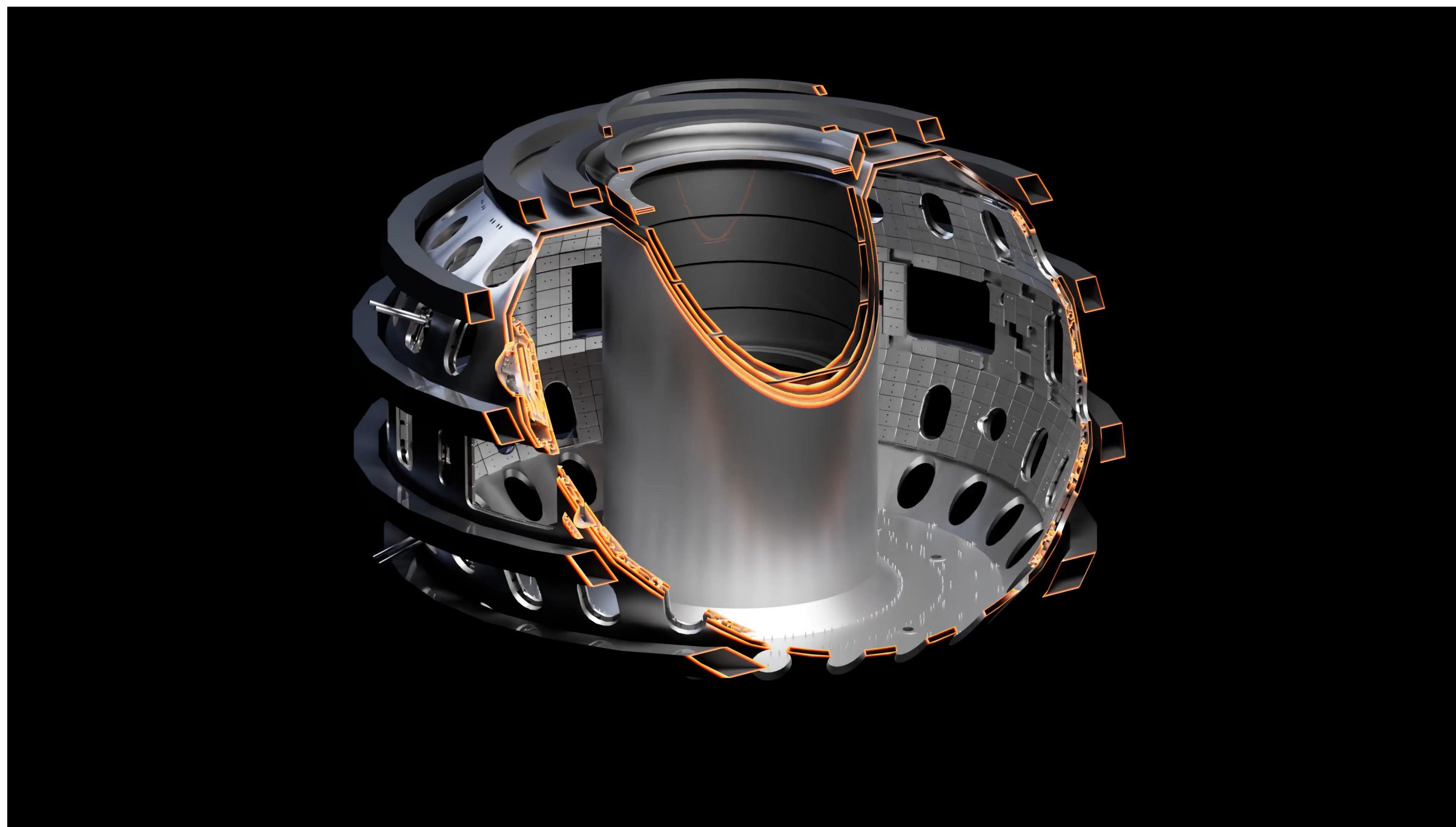
GA Ion Orbiter DIII-D Vacuum Vessel Impacts



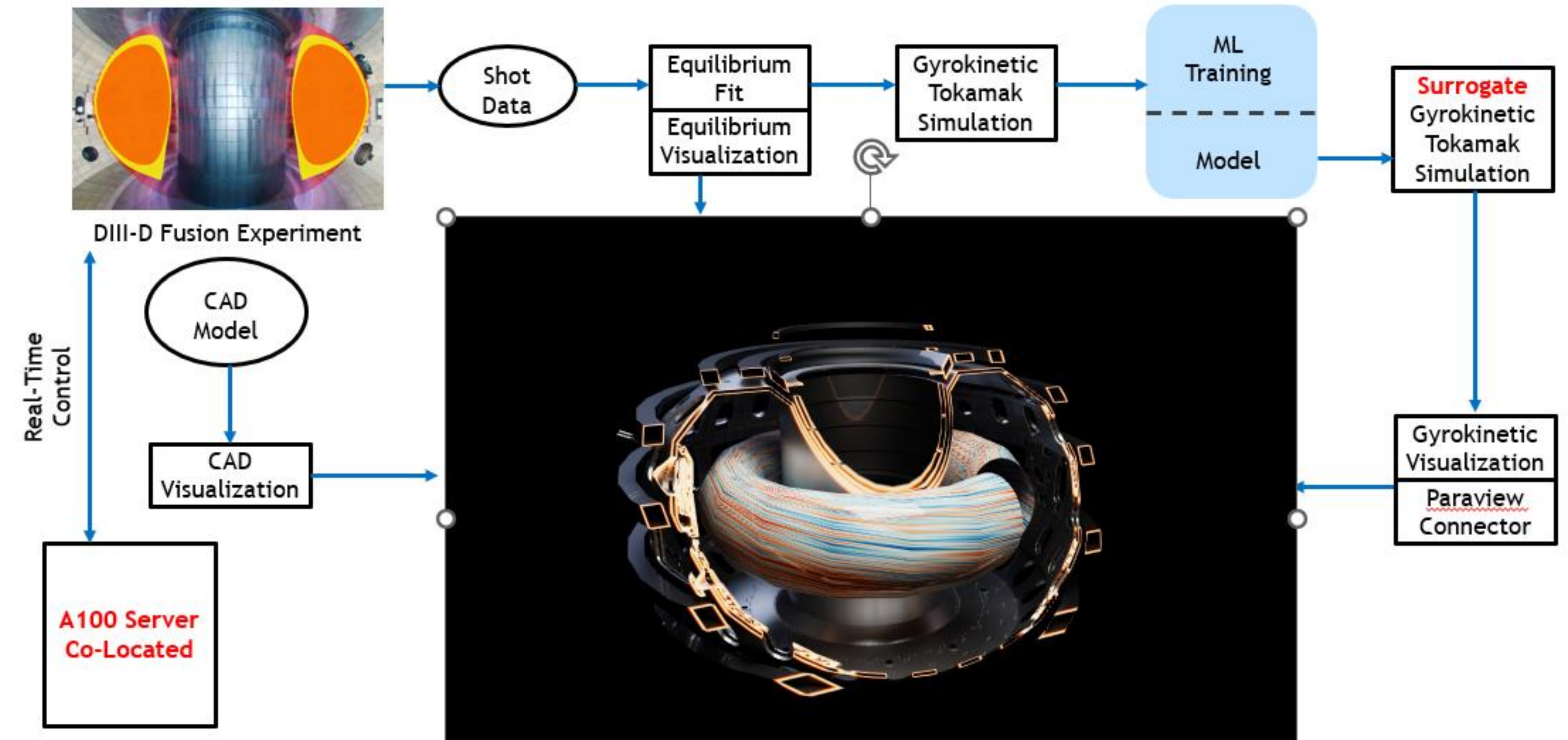
- Dedicated low-latency queue on NERSC
- Test bed for data processing in support of ITER operation



Sterling Smith & Raffi Nazikian/DOE CAKE Superfacility Review / 6 April 2023 - 12



GA PLAN: DIGITAL TWIN ENABLED CONTROL ROOM

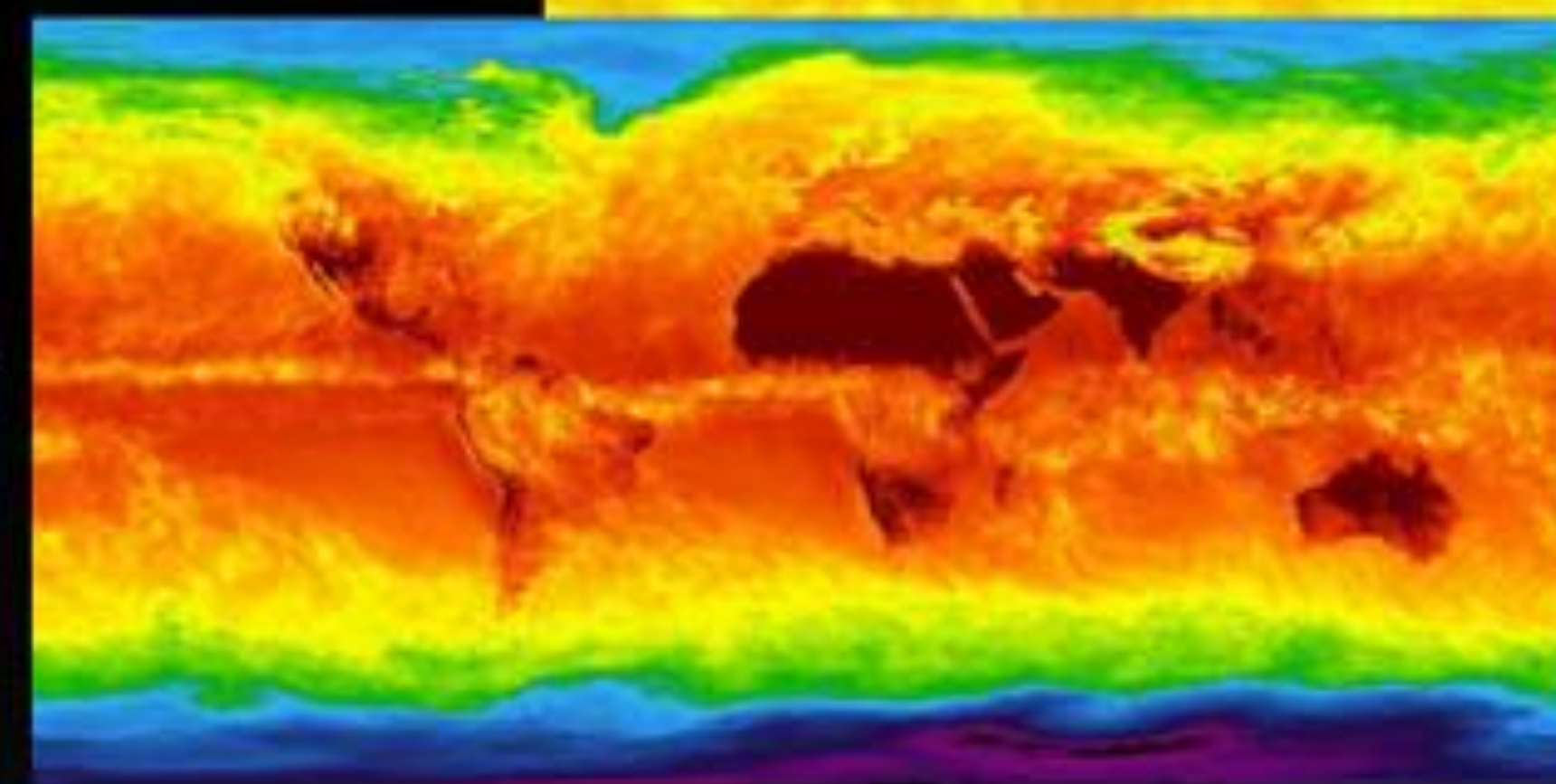
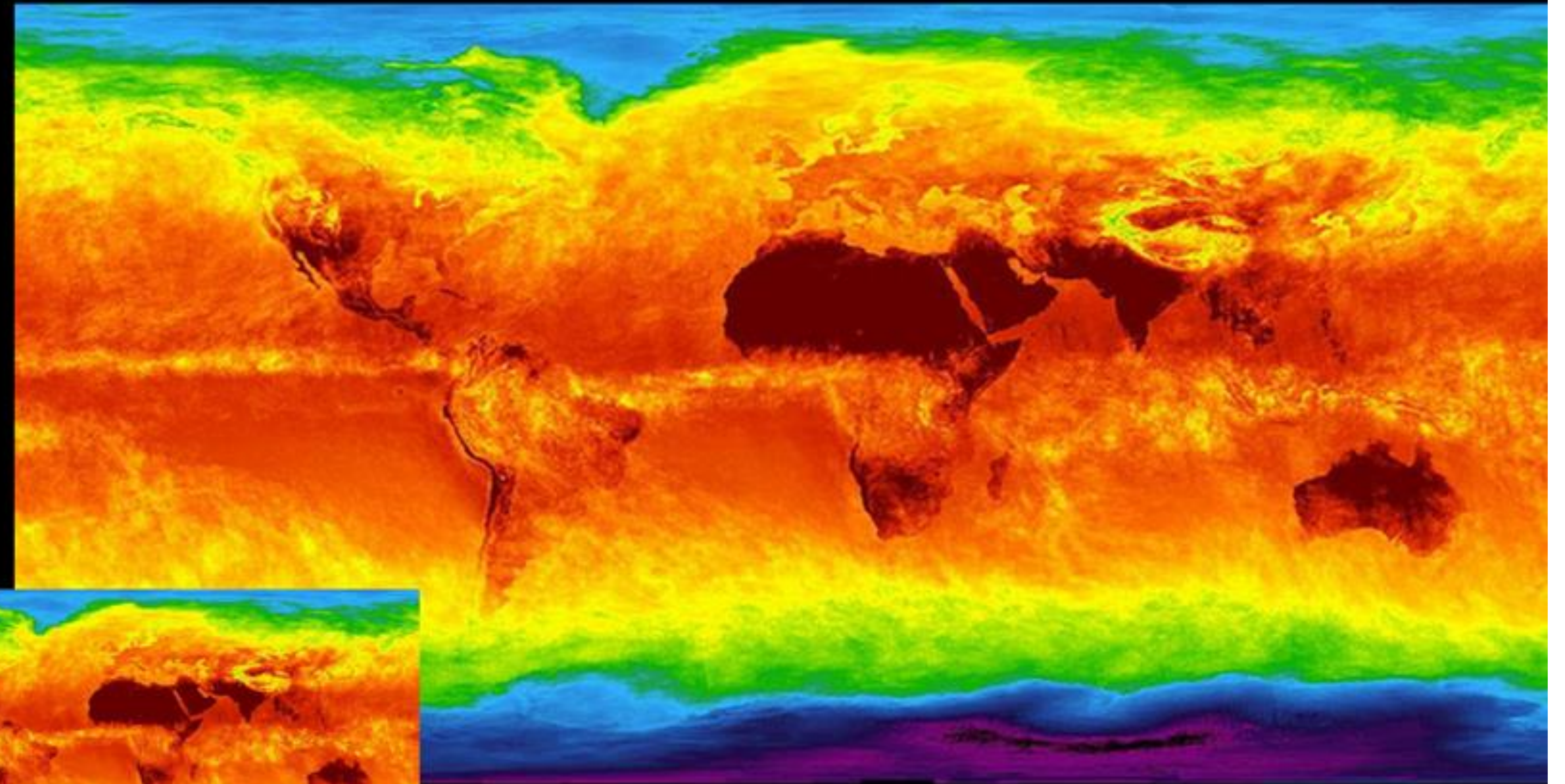


ALGORITHMS EVOLVING AT UNPRECEDENTED PACE

FourCastNet High Resolution for Data-Driven Weather Models

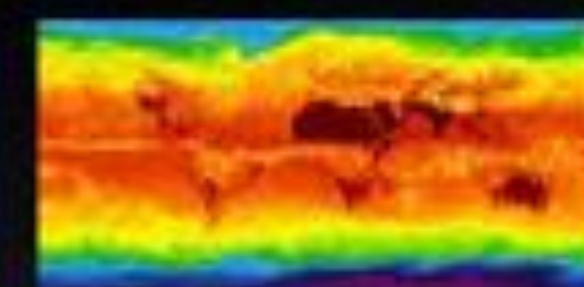
Comparison of resolutions for data-driven weather models since 2018 (Dueben & Bauer)

SOTA evolving rapidly
Recent Pre-print Kang Chen et al (2023) extend forecast to 10 days with 0.25° resolution using “cross modal Transformer”

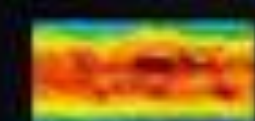


FourCastNet, Pathak et al. (2022), 0.25°, ~1,000,000 Pixels, ViT+AFNO

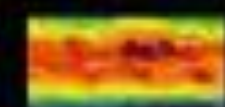
GNN, Keisler et al. (2022), 1°, 64,000 Pixels, Graph Neural Networks



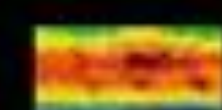
DLWP, Weyn et al. (2020). 2°, 16K pixels, Deep CNN on Cubesphere/(2021) ResNet



Weyn et al. (2019), 2.5° N.H only, 72x36, 2.6k pixels, ConvLSTM



WeatherBench, Rasp et al. (2020). 5.625°, 64x32, 2K pixels, CNN



Deuben & Bauer (2018), 6° , 60x30, 1.8K pixels, MLP

DIGITAL TWIN WITH EARTH2

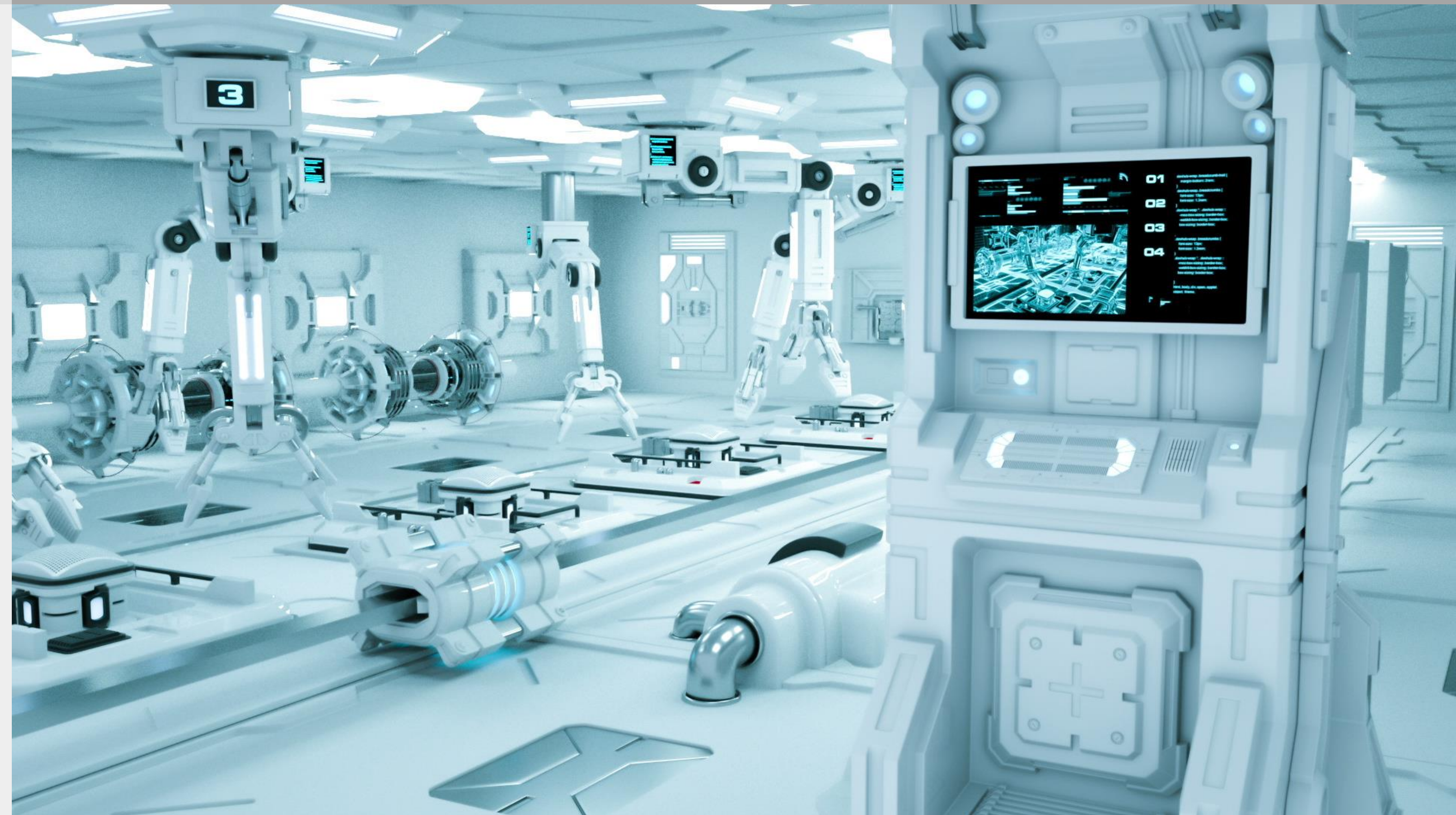
https://www.youtube.com/watch?v=8cQoYcbUG_M



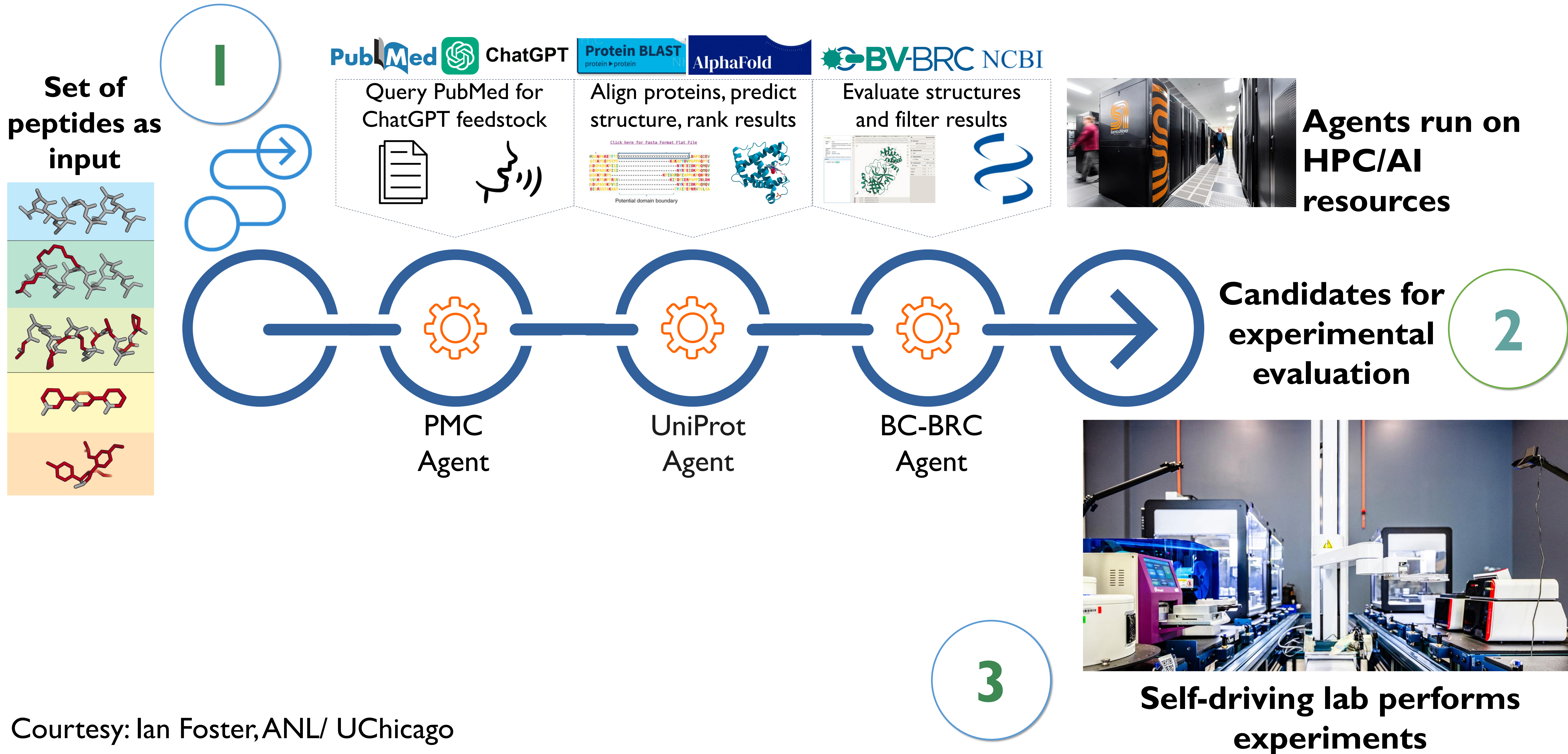
Our Vision: Smart “Factory” for Probing & Designing Complex Biological Systems

ARTIFICIAL INTELLIGENCE GUIDED, ROBOTICALLY EXECUTED EXPERIMENTS

- Accelerate the discovery process
- Elevate human creativity to higher level goals
- Democratize biological systems design approaches
- Unbiased data collection and evaluation



Link AI driven workflow with self-driving laboratory



Embodied Agent for Automated Lab Code Generation

```
##### Performing Task 1... #####
Reasoning: Based on the information provided, it seems like the next logical step
would be to prepare the master mix for the PCR reaction. This involves combining
various reagents in specific volumes to create the master mix solution.
```

Task: Prepare the master mix for the PCR reaction.

Generation Candidate Code



Memory of Tasks



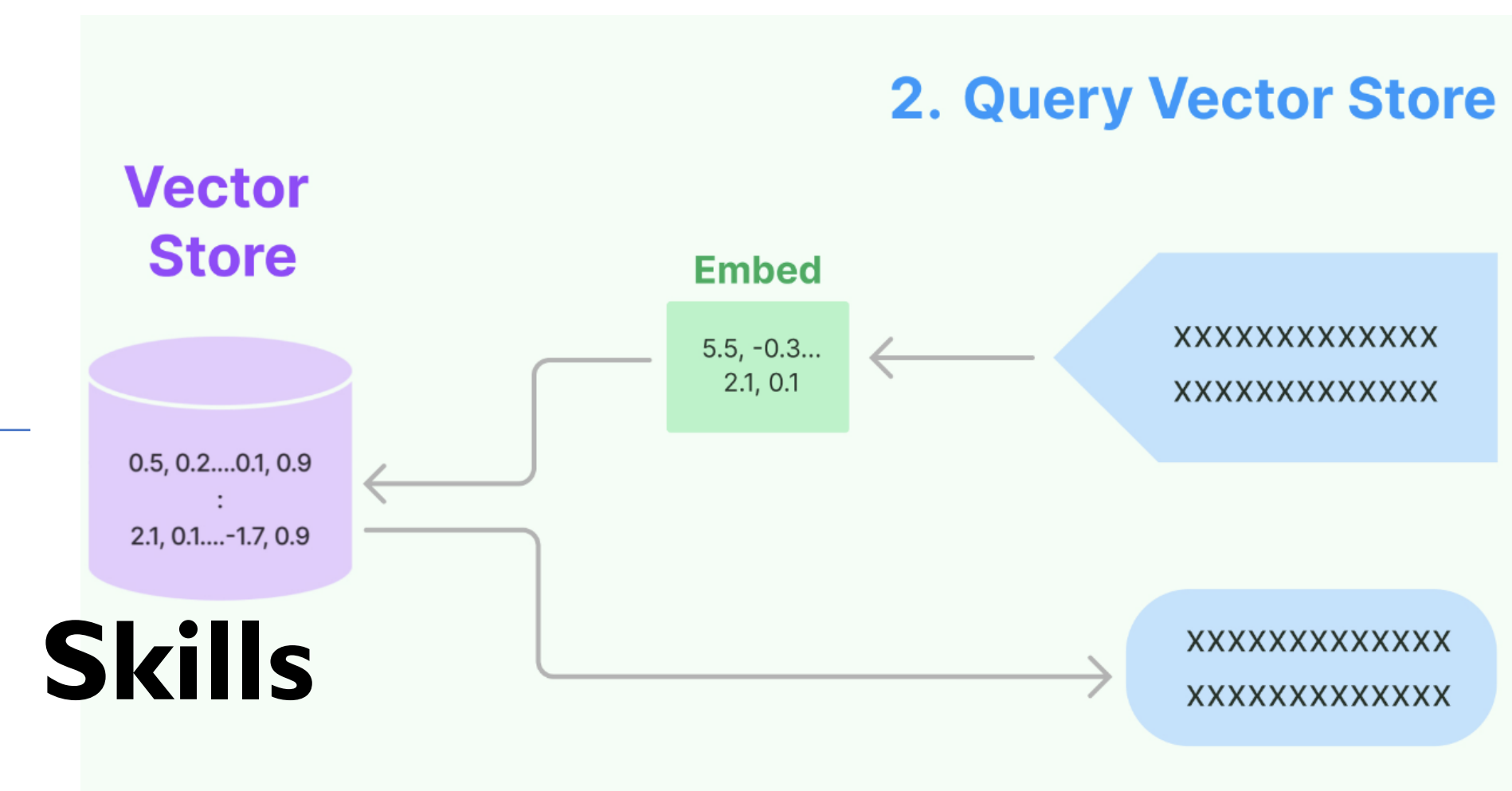
```
Useful Programs:

def PCR_Master_Mix(labware_info, protocolContext):
    */
    Input: labware_info --> json-str
    Pass in a variable labware_info that contains labware information and
    quantities used

    Output: function call that creates master mix DNA and assigns to appropriate locations
    */

Human:
labware_info = {"number_of_samples":96,
"right_pipette":"flex_8channel_1000",
"left_pipette":"flex_8channel_1000","mastermix_volume":18,"DNA_volume":2}
```

Retrieve



2. Query Vector Store

Skills

Task Decomp.

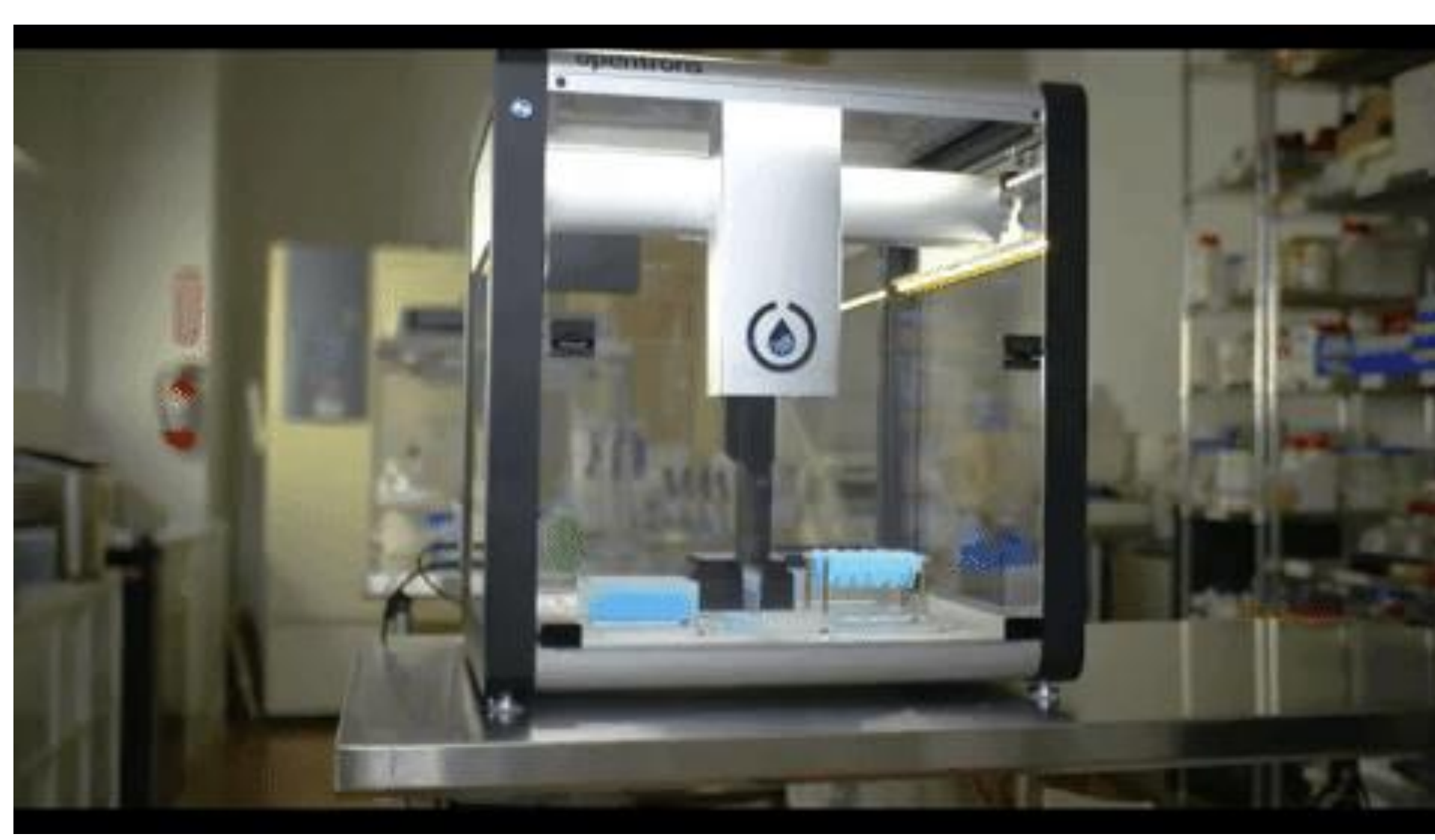
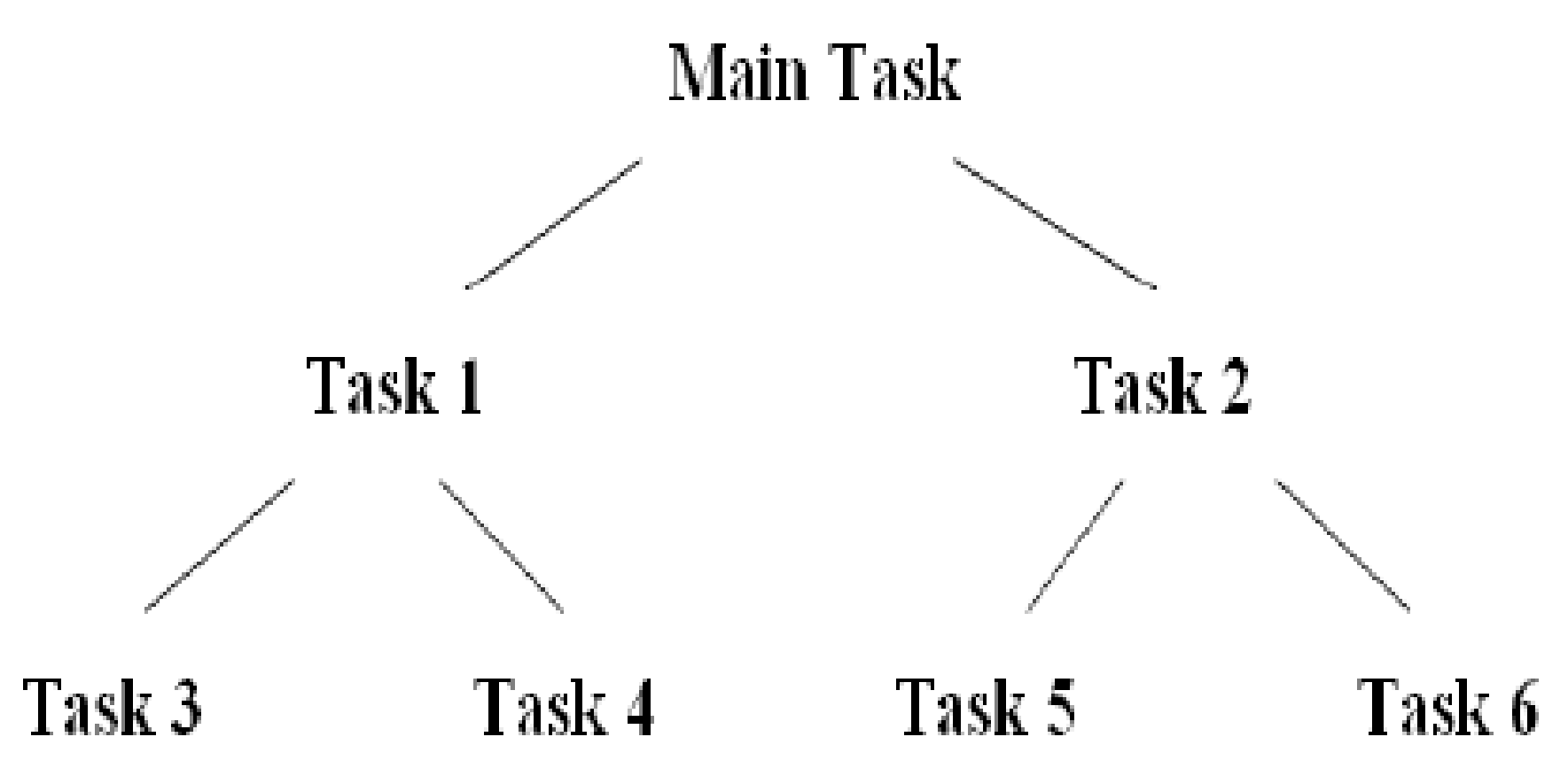
Goal Tracking

Code Action

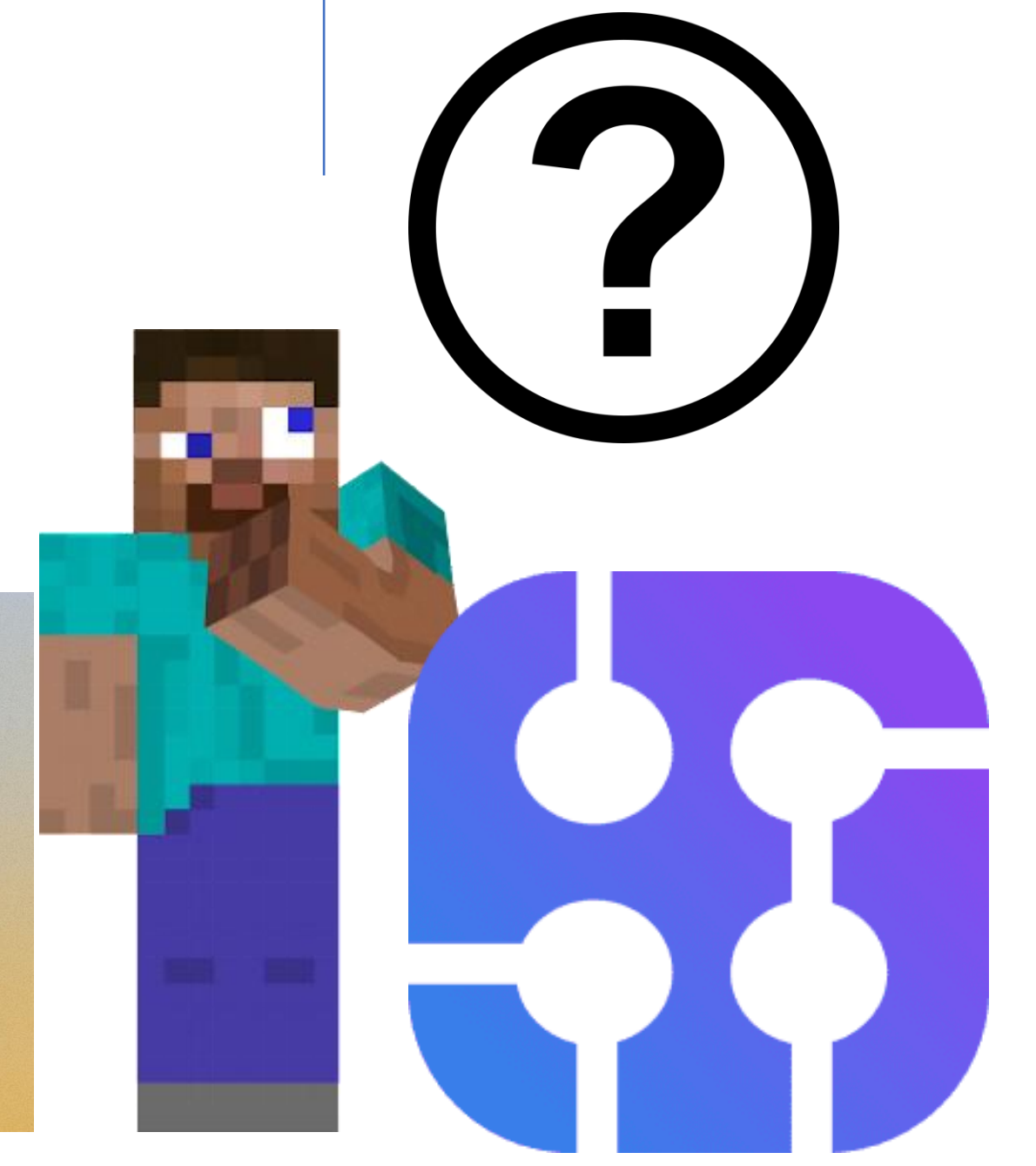
Execution Error

Refine Code

Add Code Skill

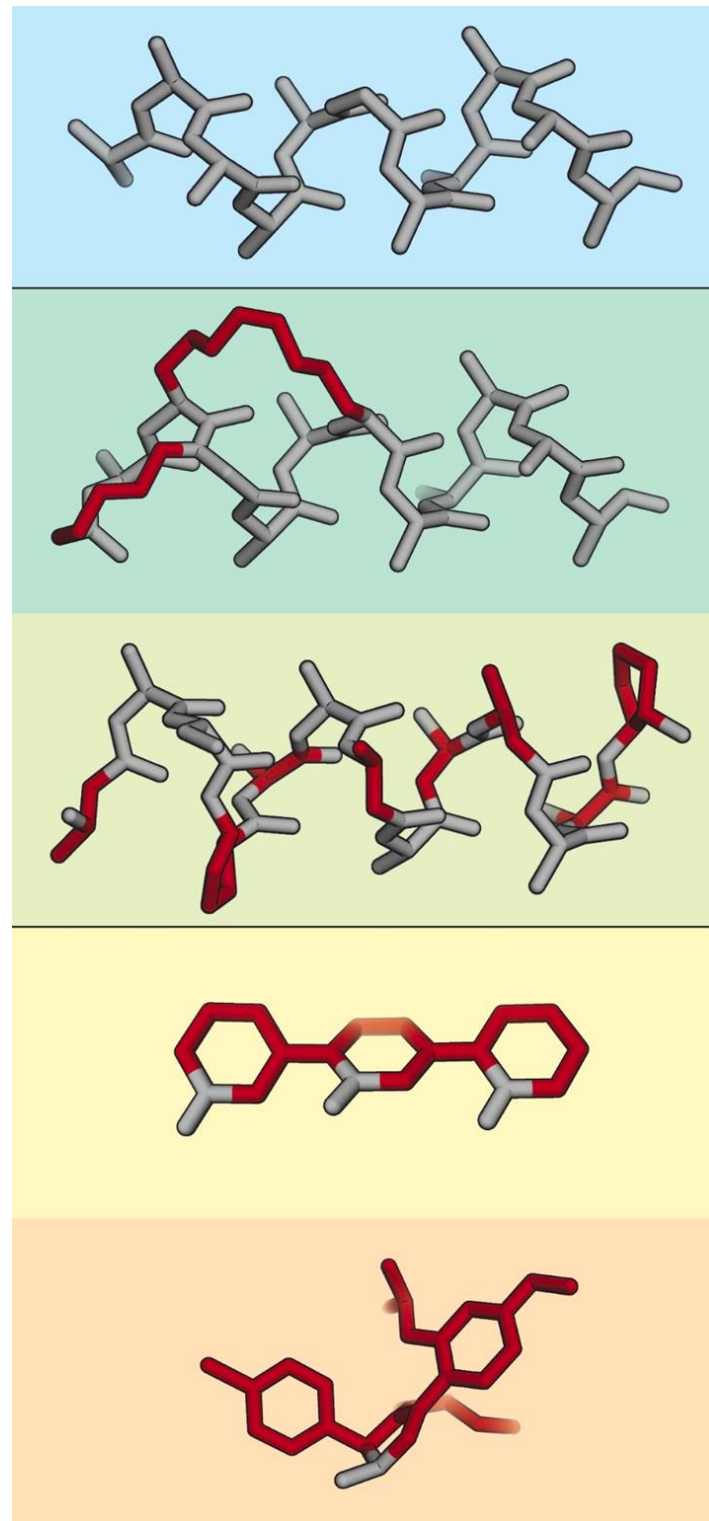


Verify Code



Feedback to define additional experiments

Set of peptides as input

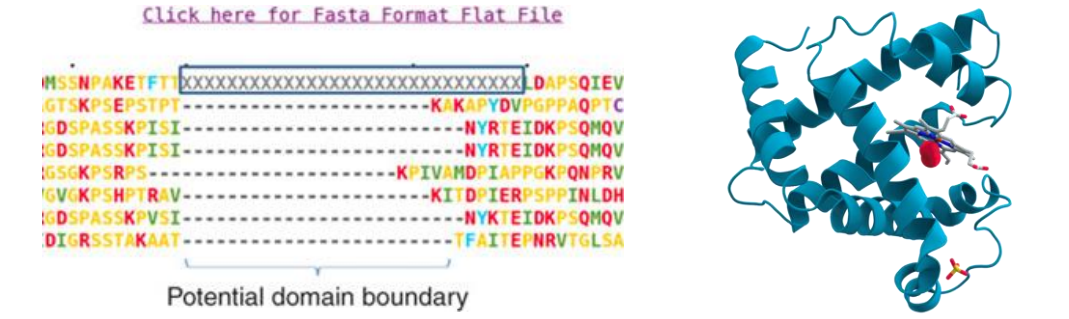


PubMed ChatGPT Protein BLAST AlphaFold BV-BRC NCBI

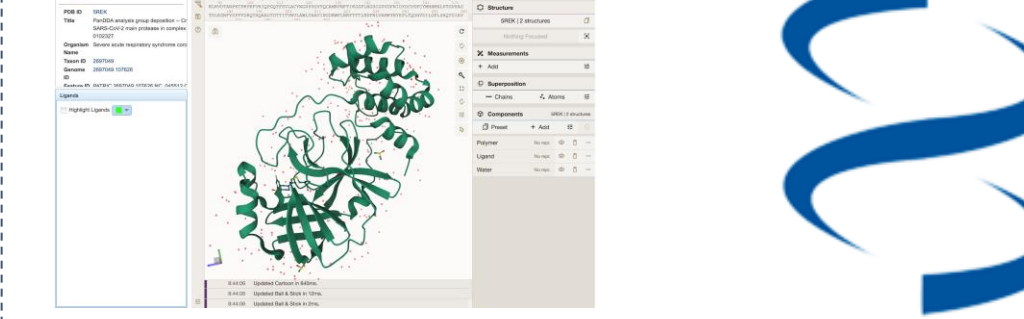
Query PubMed for ChatGPT feedstock



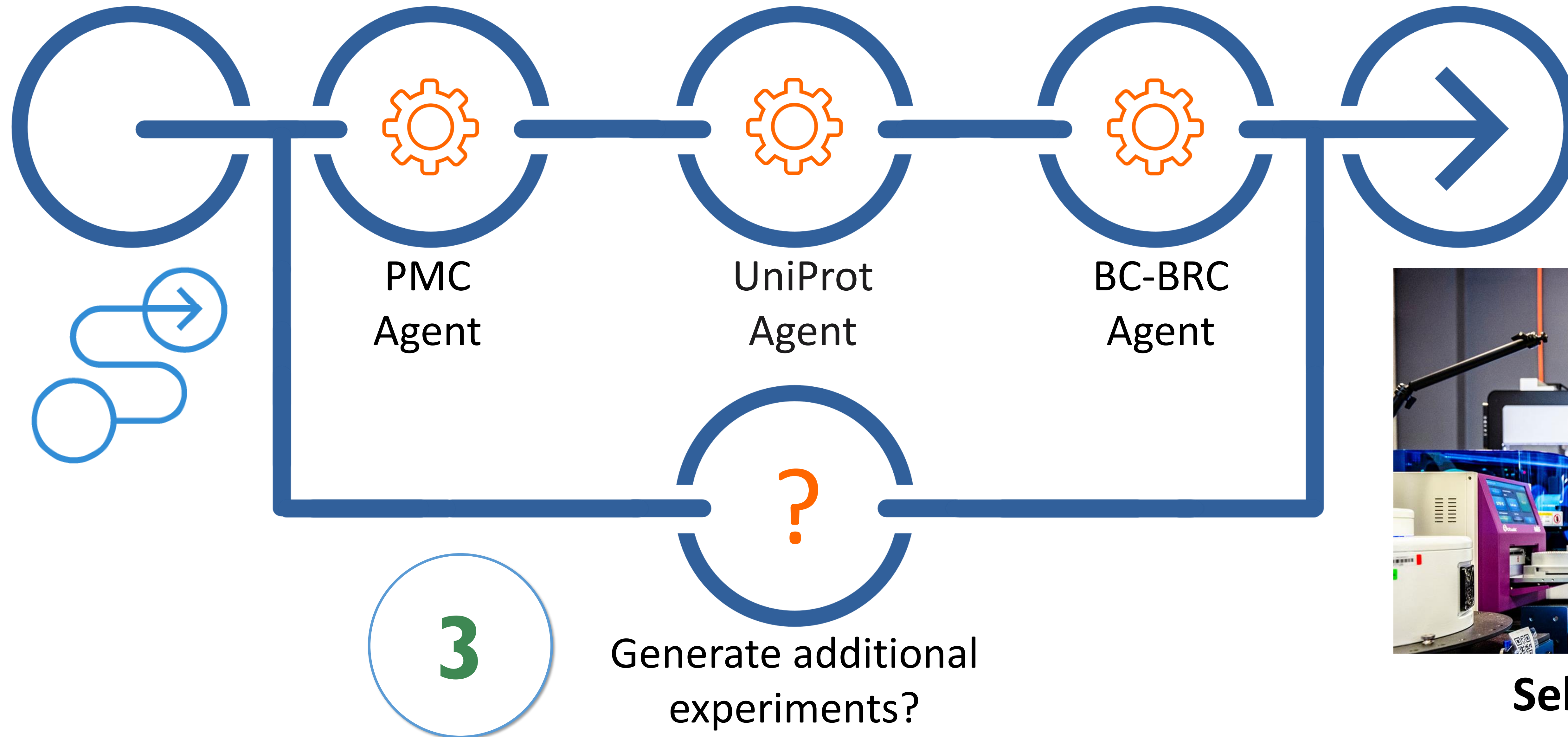
Align proteins, predict structure, rank results



Evaluate structures and filter results

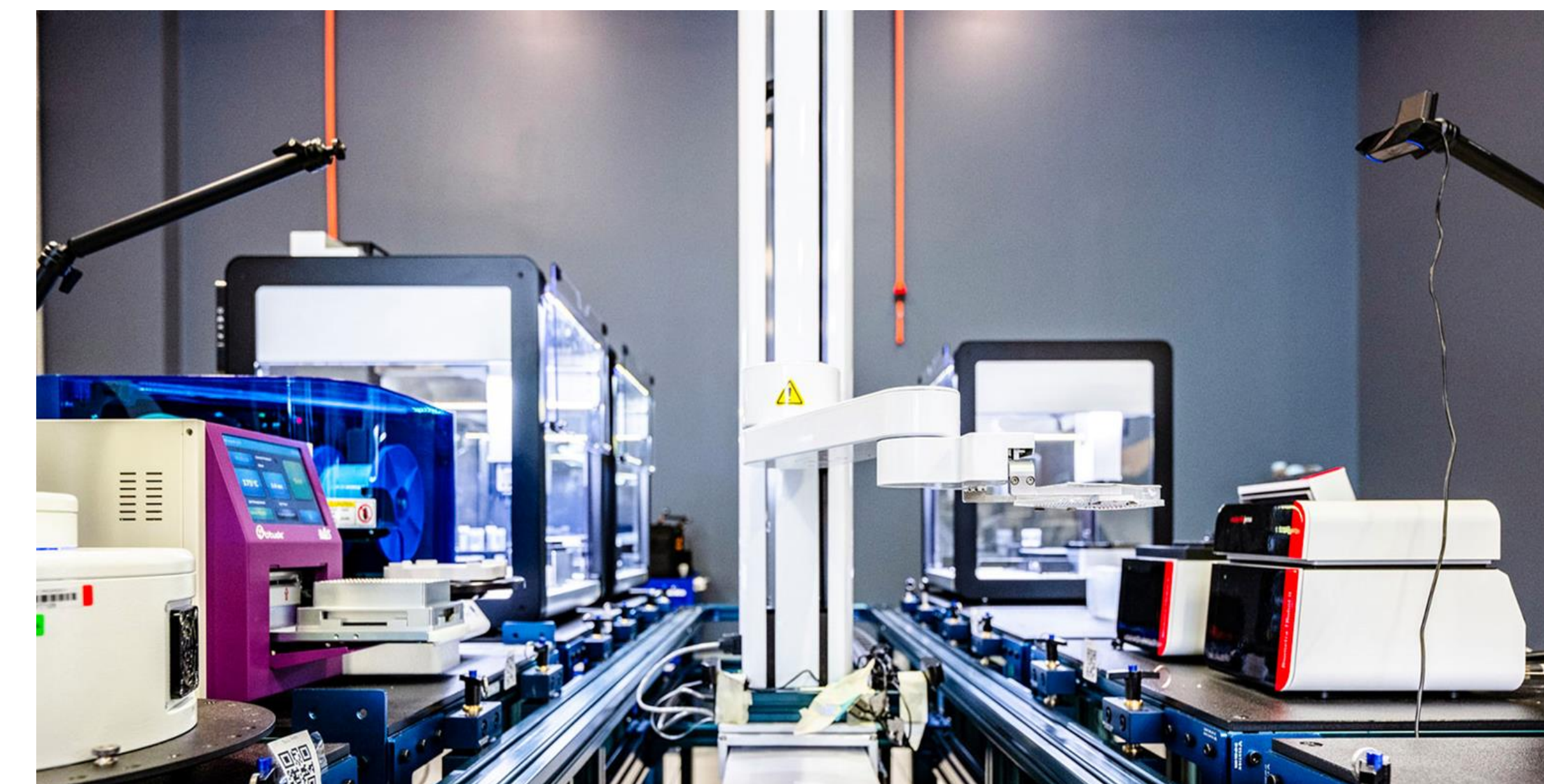


Agents run on HPC/AI resources

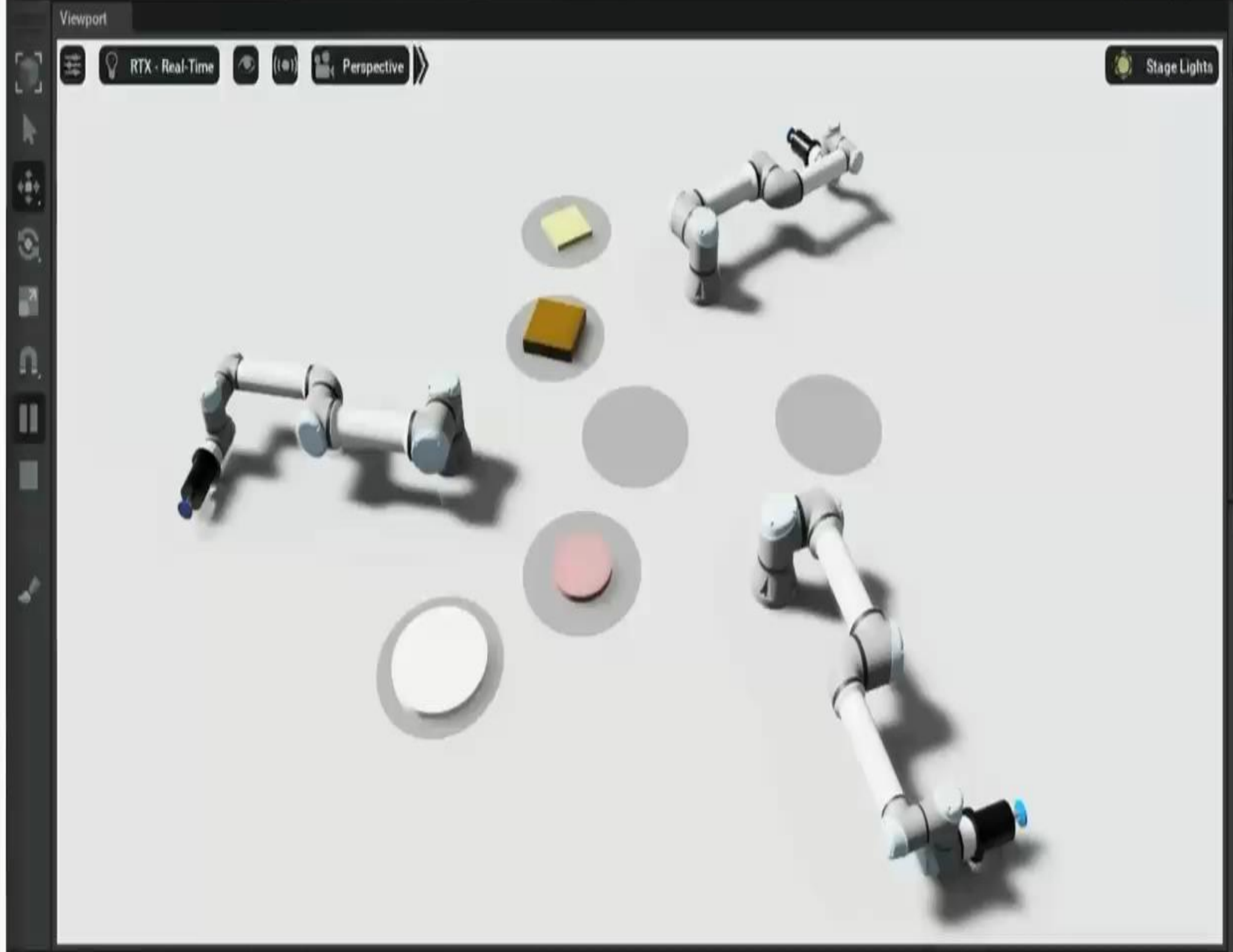


Candidates for experimental evaluation

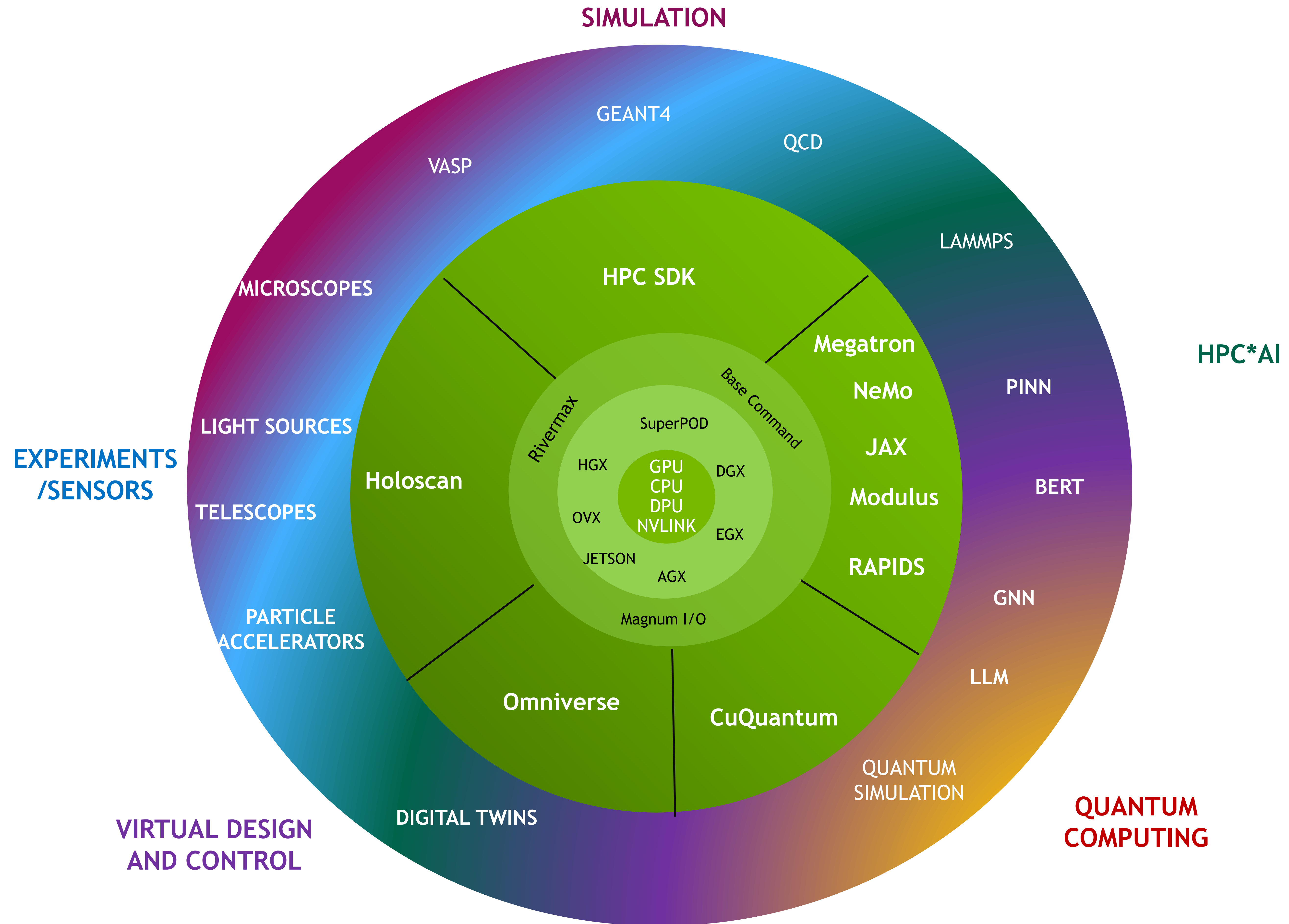
2



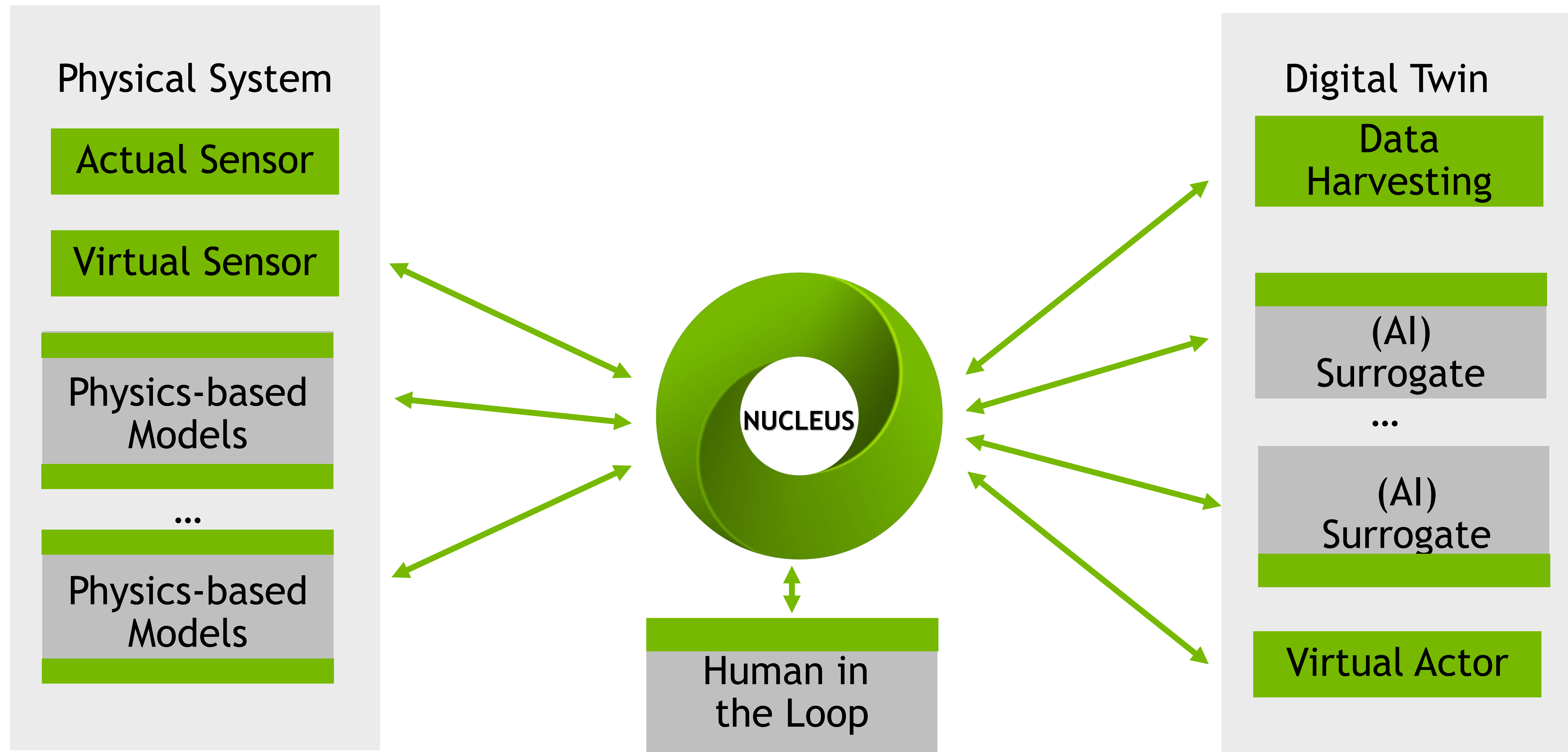
Self-driving lab performs experiments



NVIDIA PLATFORM EVOLVING TO MEET THE CHALLENGE



OMNIVERSE: PLATFORM FOR BUILDING DIGITAL TWINS



ADVANCED TOOLS AND TECHNOLOGIES

Foundational Platform Components

NUCLEUS



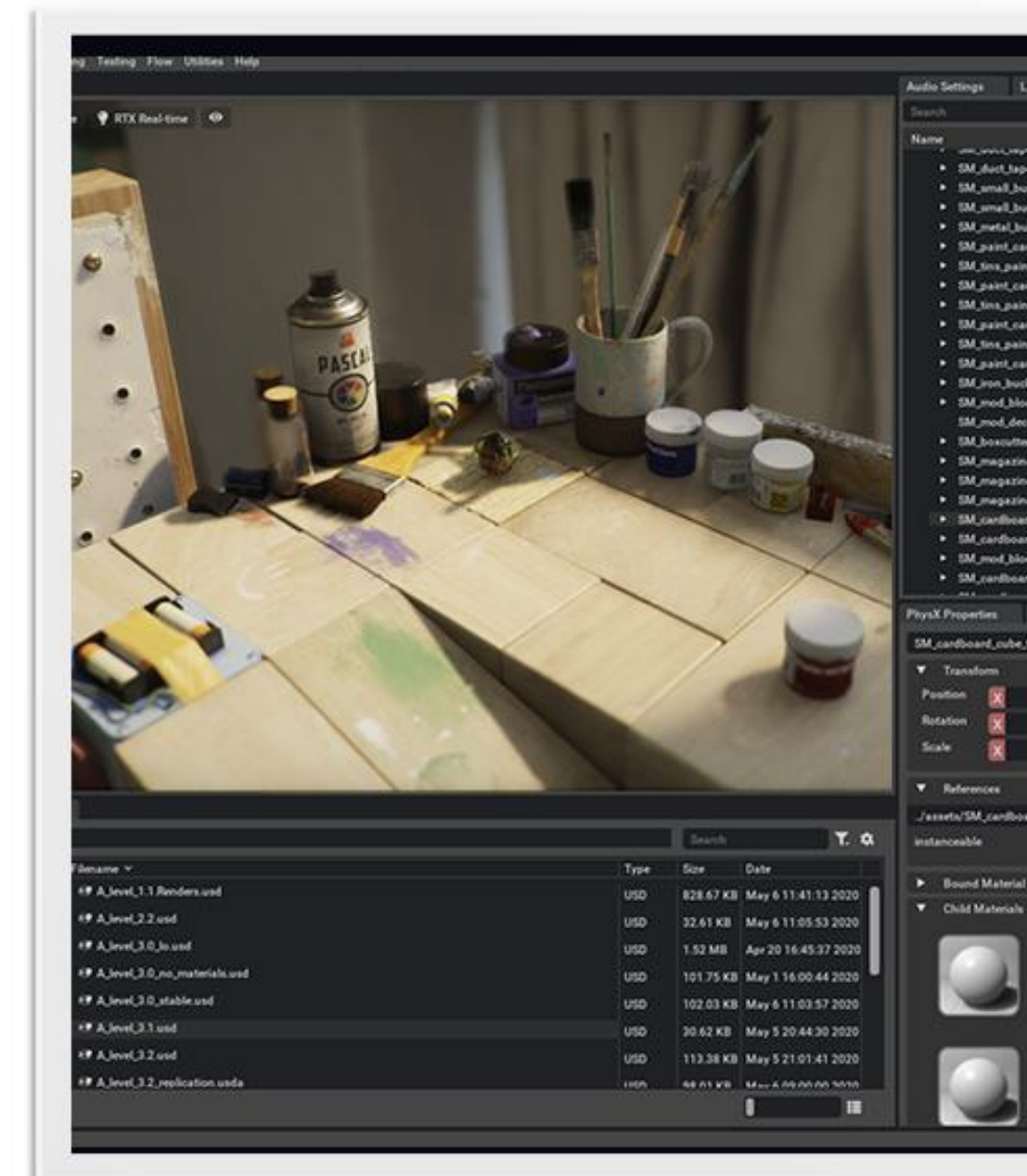
Source of truth

CONNECT



Coupling

KIT



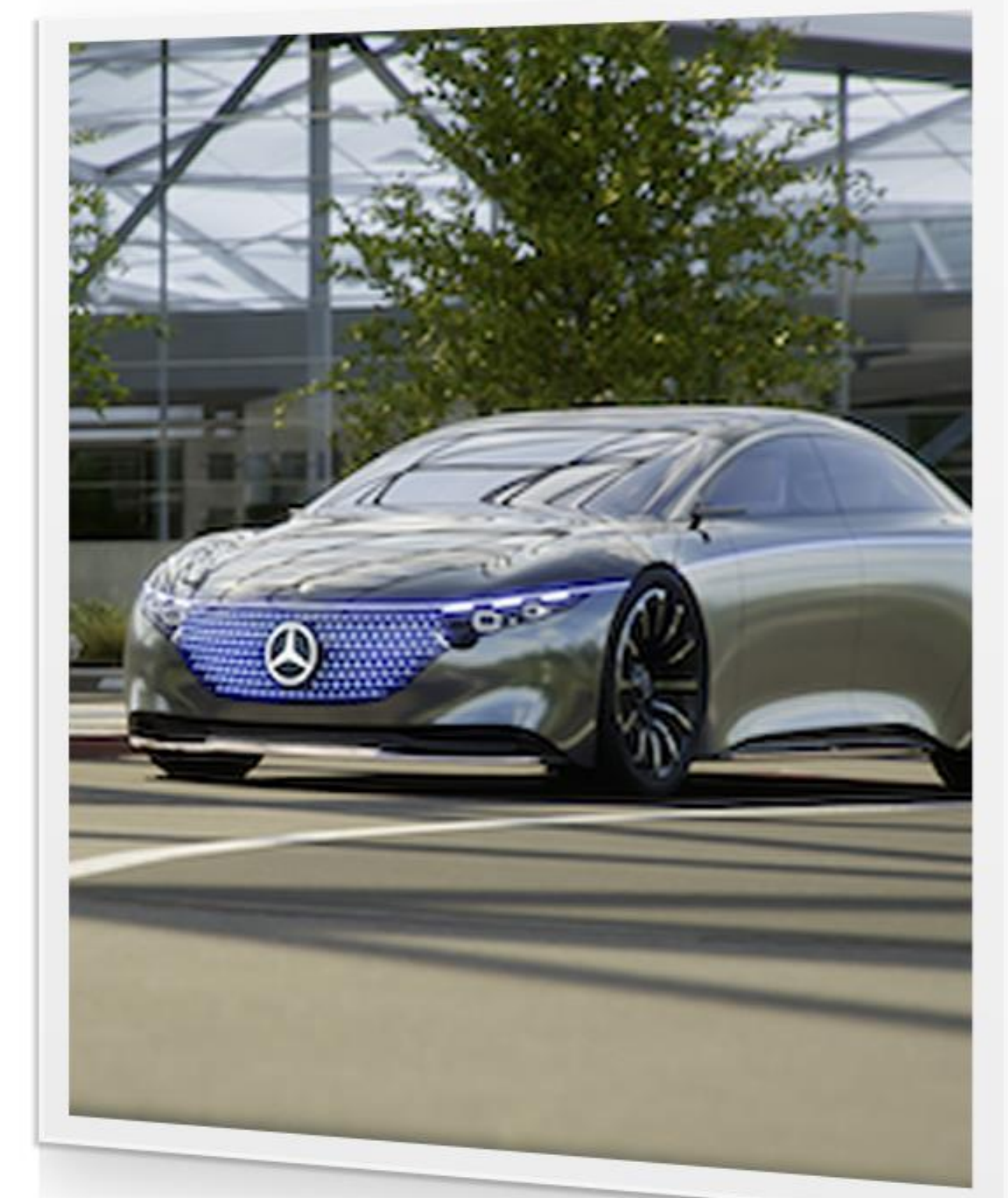
Application API
User experience

SIMULATION



Virtual Actor

RTX RENDERER



Virtual Sensor



DATA AGGREGATION AND COORDINATION VIA USD



UNIVERSAL SCENE DESCRIPTION

The “HTML” of 3D Virtual Worlds

- ▶ Developed by Pixar
- ▶ Foundation for NVIDIA Omniverse
- ▶ Open-sourced API and file framework for complex scene graphs
- ▶ Easily extensible, simplifies interchange of assets between industry software
- ▶ Introduces novel concept of layering
- ▶ Enables simultaneous collaboration for large teams in different department working on the same scene
- ▶ Originated in M&E, now becoming a standard across industries including AEC, Manufacturing, Product Design, Robotics





BACKUP

EXAMPLE OF DIGITAL TWIN FOR ASTROPHYSICS

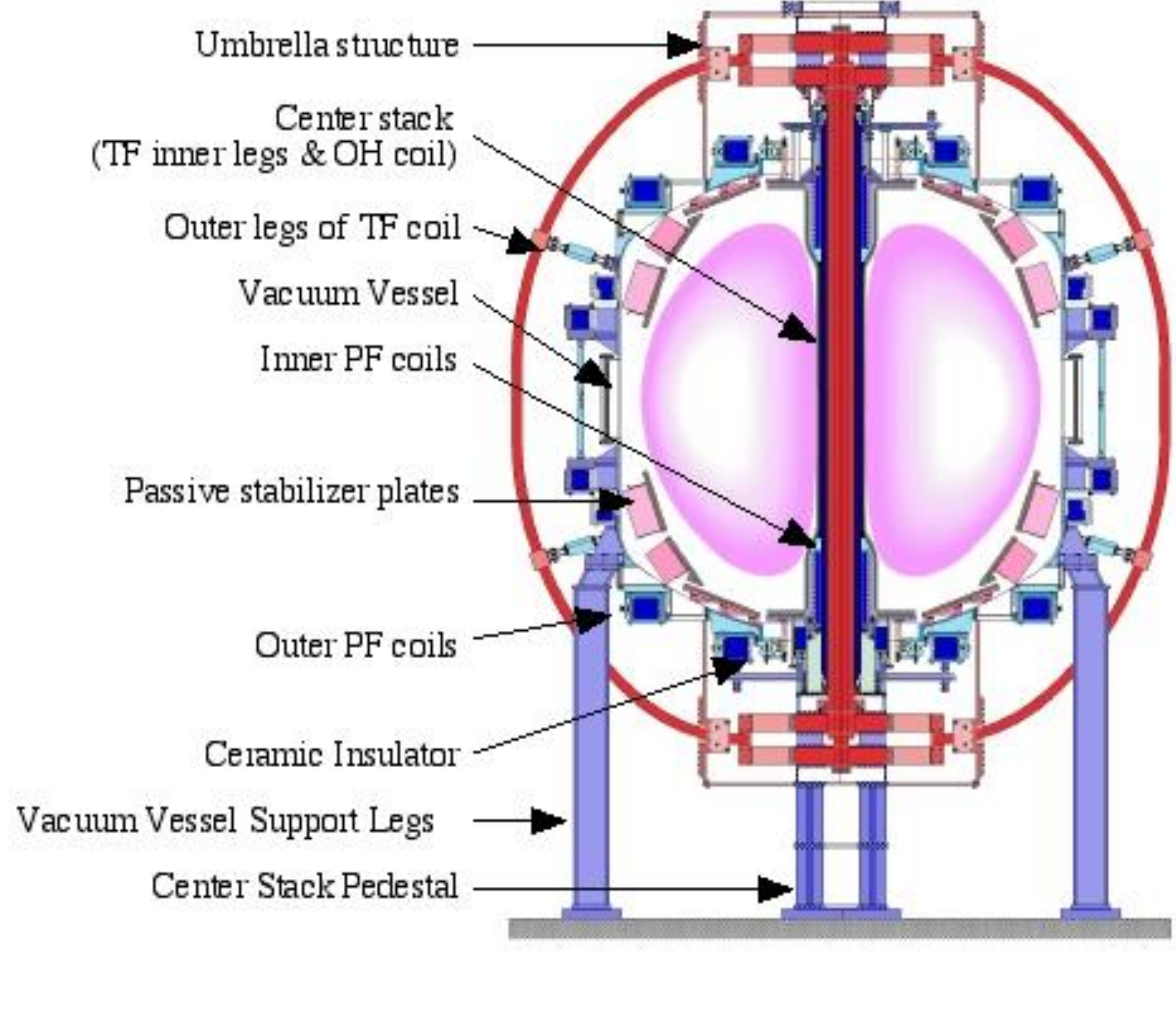
Moonwalker Digital Twin of the South Pole of The Moon

<https://www.youtube.com/watch?v=E0Rz0ZbwhJY>

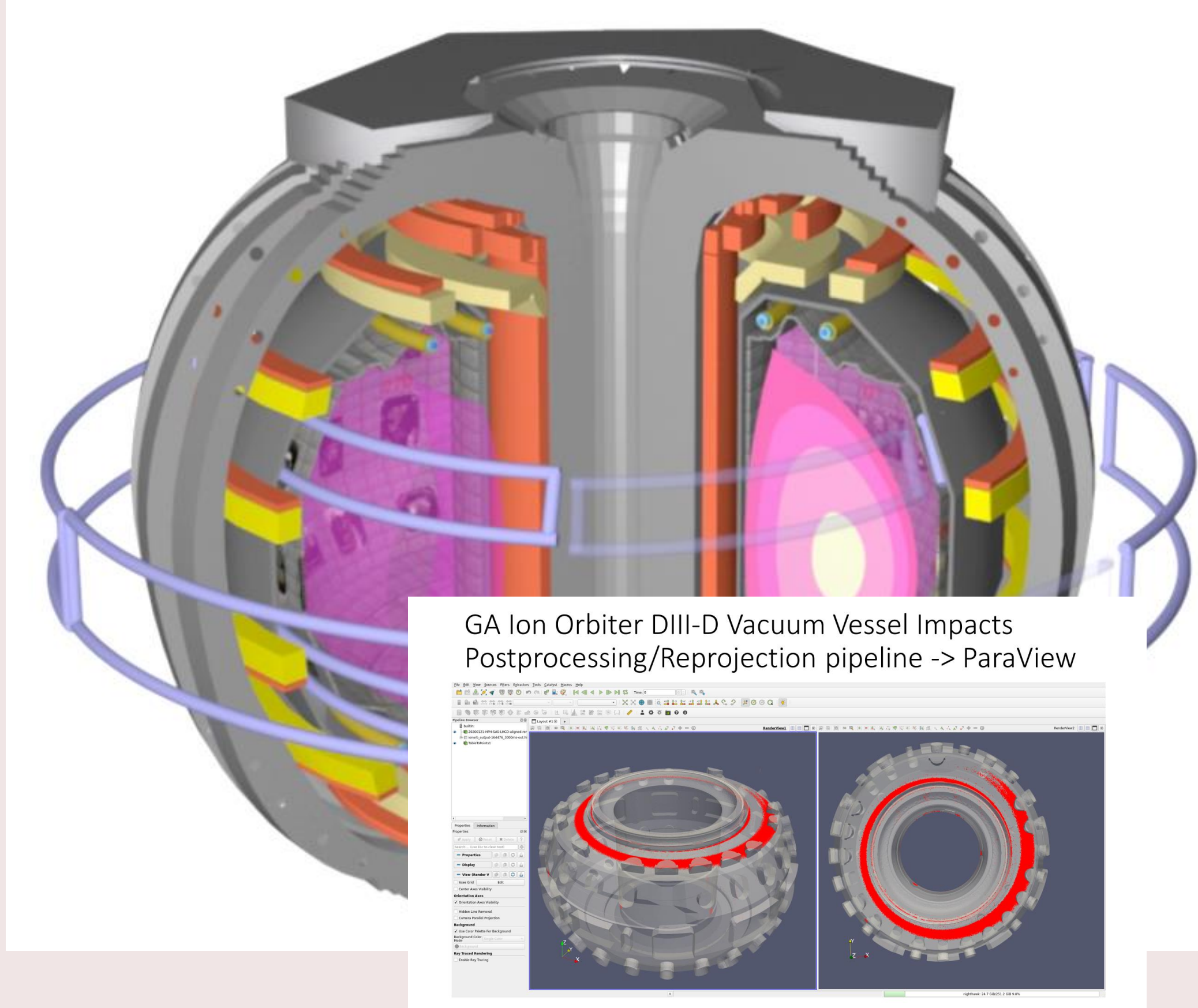


SUPERFACILITY WITH GA AND NERSC/ANL: VISION FOR FUTURE

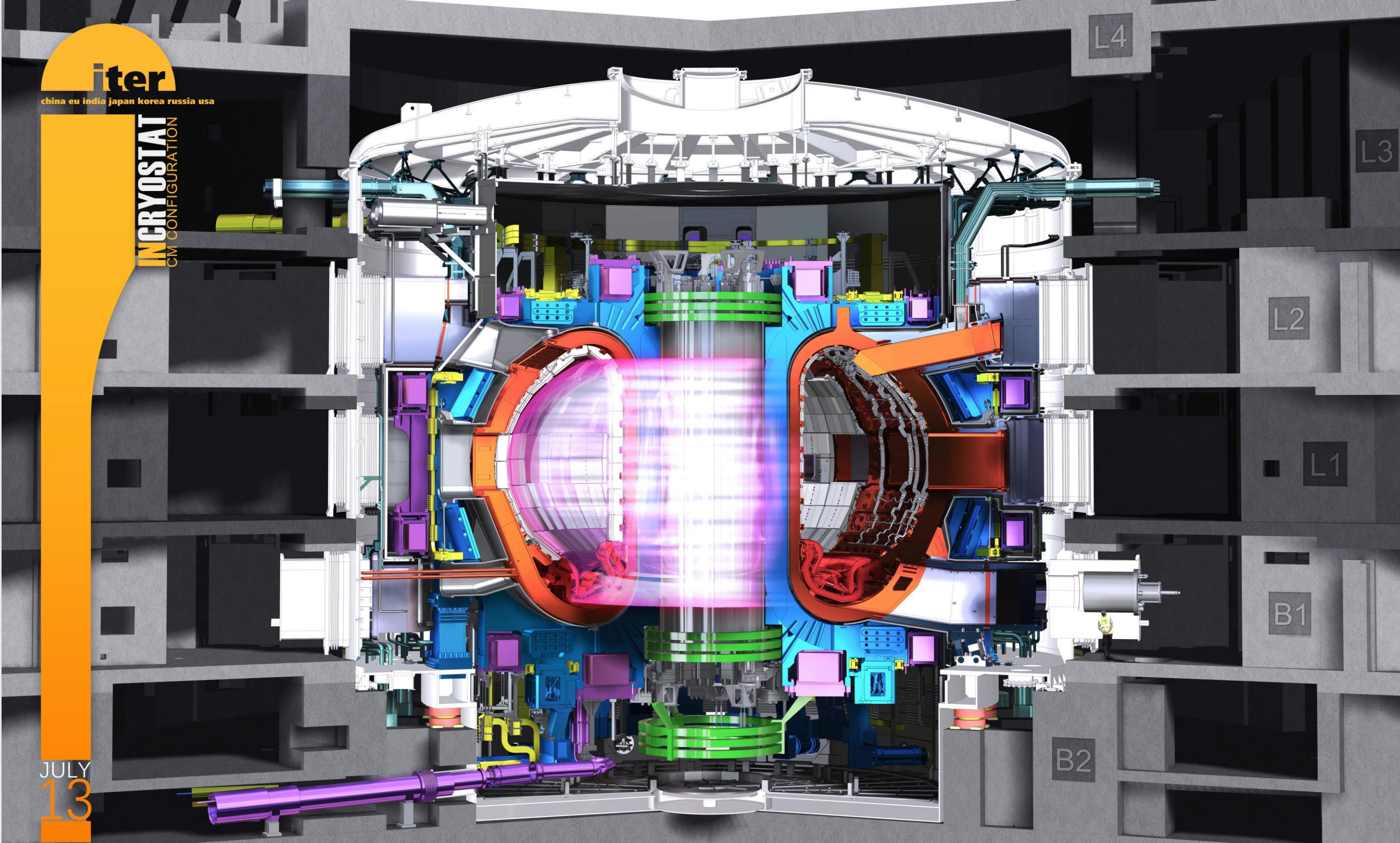
NSTX-U



DIII-D National Fusion Facility

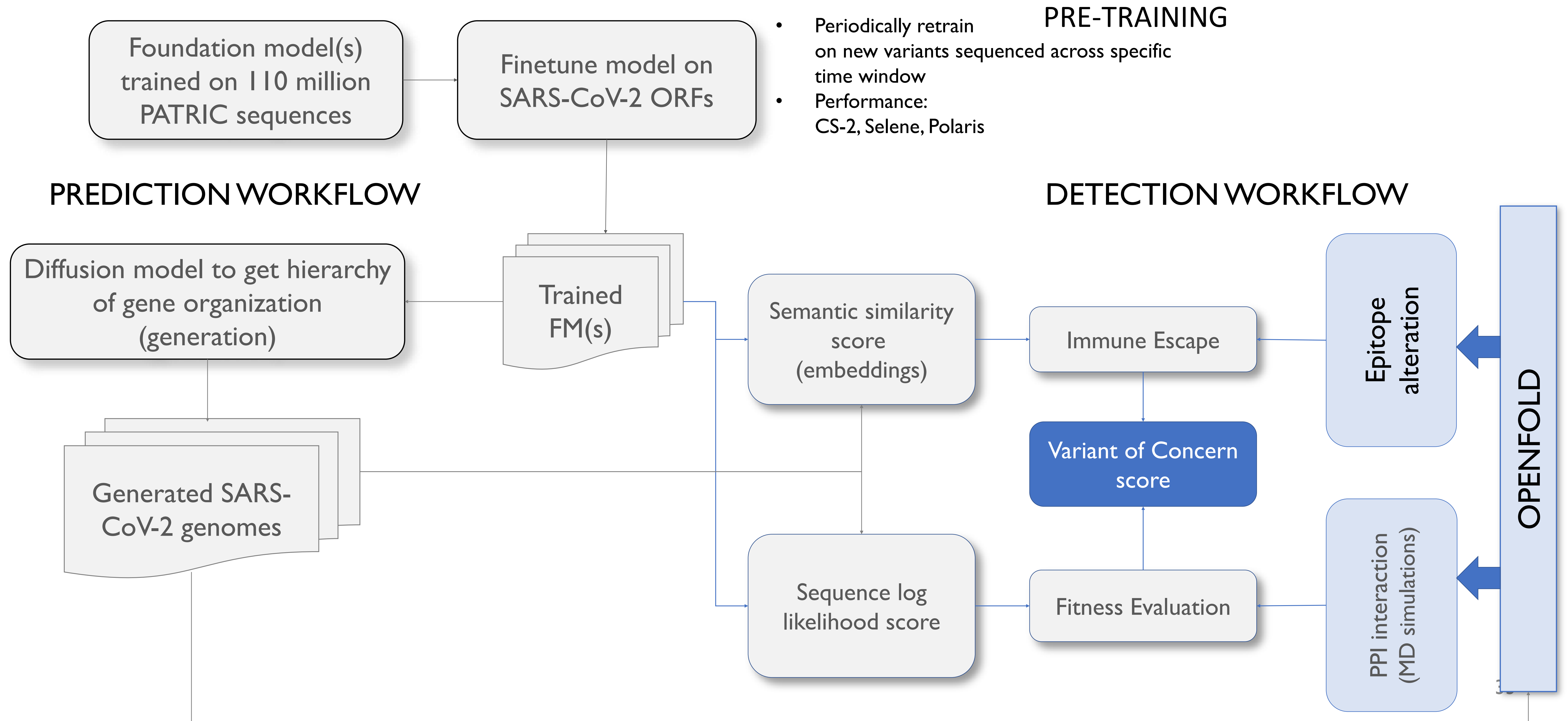


ITER

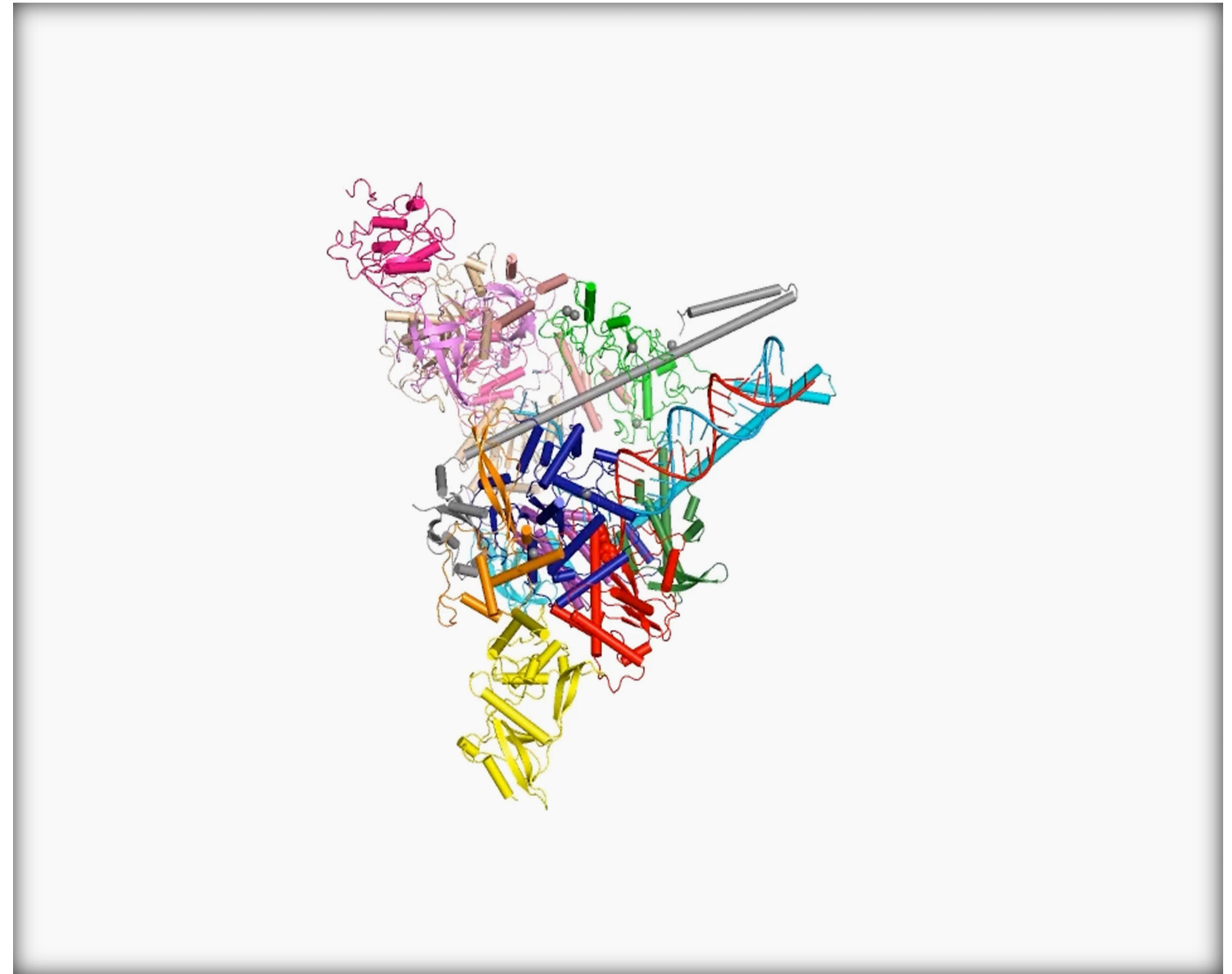
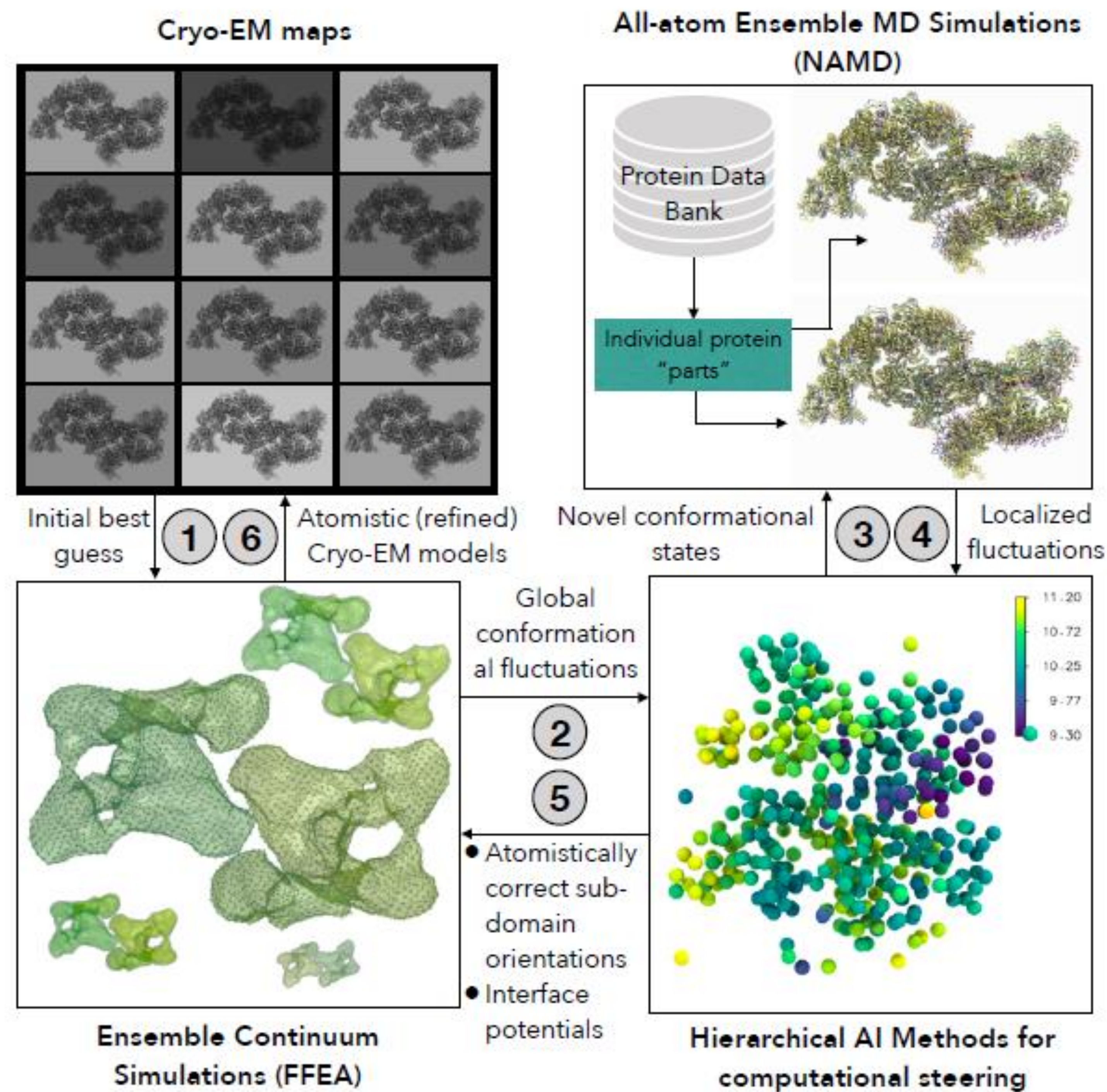


FUTURE POST Exascale

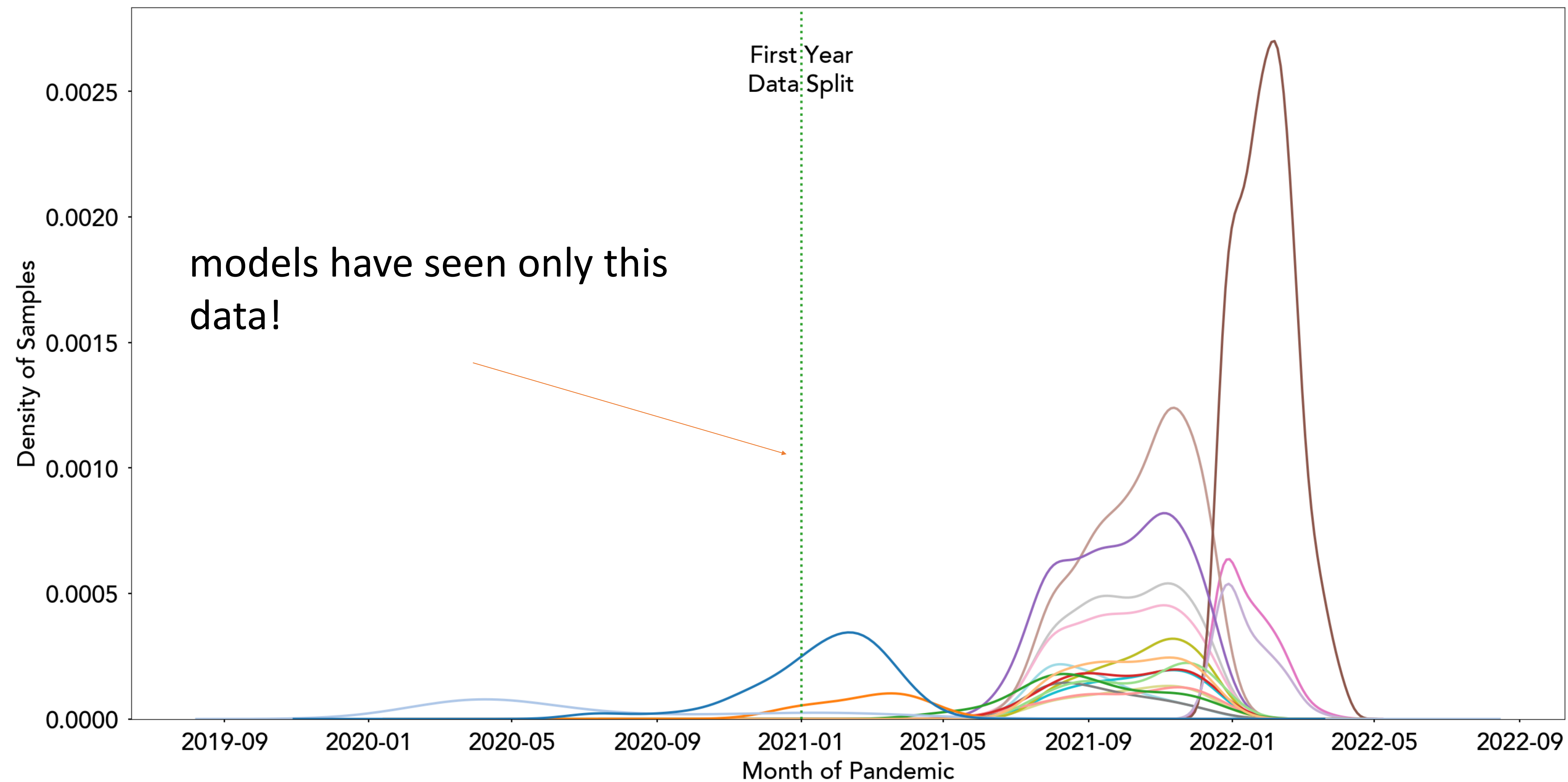
Using foundation models to predict SARS-CoV-2 evolution



Science goal: Understand how to “trap” the SARS-CoV-2 replication-transcription complex (RTC) using Low Res CryoEM



GenSLMs finetuned on SARS-CoV-2 genomes can distinguish variants



- | | | | | | | | | | |
|---------|-----------|-------------|----------|---------|----------|-----------|---------|----------|---------|
| — B.1.2 | — B.1.429 | — B.1.617.2 | — AY.122 | — AY.44 | — BA.1.1 | — BA.1.15 | — AY.26 | — AY.100 | — AY.39 |
| — B.1 | — AY.25.1 | — AY.43 | — AY.4 | — BA.1 | — AY.103 | — AY.25 | — AY.3 | — AY.119 | — AY.47 |