Frontier of Scientific Innovation

& Our Energy Future

Dr. Dimitri Kusnezov

U.S. Department of Energy

March 13, 2021

NETL

Disclaimer: These views are mine and not of the DOE

1. Who are we

DOE: Department of "Everything"

We are a mission agency: *we have problems to solve, often on schedules*

We have tough requirements*: US Energy Sector, Cyber over many unclassified and classified networks, Nuclear Security, Emergency Preparedness and Response, Intelligence, Strategic Petroleum Reserve, Loan Program, Environmental clean-up*

We create and maintain tools and people to work these: *17 national labs, world class user facilities, worlds fastest supercomputers,..*

We solve problems no one else can

We know how to protect data and information

We are a go-to agency for informing crises and urgent decisions

We have turned to big science and big compute as unique means to approach problems that cannot be instrumented: *Built around Uncertainty Quantification, Verification & Validation; Many successes. Billions in deferred costs. Likely more in the future; Innovation in our missions will require virtual tools to explore ideas; Problems typically have a cost for inaction.*

Today AI based methods, which are still nascent and narrowly applied, are providing means to innovate and impact everything including the science, environment, energy, health and climate.

DOE

Selected DOE Science Accomplishments

- **Largest funder of the physical sciences in the US**
- **Largest generator of Nobel-prize winning scientists in the world**
  - 40% of all winners for Physics, 25% for Chemistry
- **Global Leader in development of supercomputers**
- **Human Genome Initiative & Project**
  - Transformative across domains from the economy to precision medicine.
- **Led the development of many important US technologies:**
  - **LED Lighting**
  - **Fracking**
  - **Microelectronics advances**
  - **Batteries & Storage**
  - **Wide-band radar**
  - **Developed AI capabilities**
  - **Nuclear Power**
  - **Nuclear Medicine/ Radioisotopes**
  - **CRISPR/Cas9**
  - **....**

2. How we make predictions today

Prediction is part of our day to day lives...

- First prediction of solar eclipse by Thales of Miletus *(585 BC)*
- Over the past year we have turned to predictions to understand our daily lives due to COVID
- Tide tables
- Lifetime of first excited state of hydrogen
- Performance prediction for a new aircraft
- Coal boilers coevolving with renewable resources
- Astrology
- Energy/water nexus – planning water availability, its transportation, its use to energy generation, fuel supply, customer demand...
- Weather this weekend

... But consequences of poor predictions are not all the same.

Today we are turning to science based prediction to inform increasingly serious problems.

Science/Policy Interface – a few observations on the subtleties of trust

How we implement science to inform decisions and policy is inherently challenging

More scientists is not necessarily the answer either

- *The pace of science is not typically commensurate with the pace of decisions that are needed*

Observe that 'trust' in the role and value of science is personal and experientially rooted

- *Scientific divisions were captured centuries ago, based on beliefs of how we reason*
- *We see these schools of thought today not only among scientists; Its not scientist vs non-scientist*

Today I hope to highlight challenges and approaches to being responsive to urgent issues.

How do we understand complex systems?

Historically:

- Induction/Experiment/Empiricism– if you measure it or test what you are looking for, for example – 'Pillar 1'
- Deduction/Theory/Models – if you understand the principles of behaviors, you can make predictions on what could happen – 'Pillar 2'

More recently, computational science has been promoted as a 'third pillar' of science, complementary to experiment and theory.

I view it rather as an enabler of theory and deduction and not as something separate. More than semantics, because it can shape where you think innovation and discovery are heading.

My view on where discovery and innovation are heading from the vantage point of HPC

High performance computing (HPC) and computational science tools enable exploration of theories and models well beyond what humans can analytically compute

HPC serves as enablers to extend our reach into nature's most complex systems or phenomena

Induction/ Theory/Models → Computational science enables solving far more complex theories

But it also has allowed asking ever more complex questions

The tool for prediction today is largely simulation – but there are two distinct classes of problems

Two categories of problems

Category 1: Well defined, scientifically processed questions.

- *Those for which a mathematical theory of error exists*
- *For example, scaling a "single-physics" application to the maximum scale the computer can handle; well defined problems in controlled approximations.*
- *The most comfortable and conventional approach – typical for academic research*

Category 2: Outcome focused, often technically imprecise.

- *Those which we are applying to high-leverage decisions; where the promise of supercomputing is driving us to address issues of national importance, but where approach to prediction requires significant development.*
- *These are typically multi-scale problems and multi-disciplinary*

Innovation and discovery are enabled by how far we are willing to push ourselves into this second category

COMPUTING IN SCIENCE & ENGINEERING, Sep/Oct 07, pp 62-67

Many examples of Category 2 successes where we are called upon to help address questions that are time urgent:

Columbia accident

Wikileaks

Ebola

Covid-19

Underwear bomber and aviation security

Burnt Frost – satellite shootdown

Fukushima Daiichi

Macondo – Gulf Oil Spill

Iran deal

Veterans suicide

…..

We have to take pressing questions and translate as best we can:

Predictions Today

(n.b. Not everything is suited to be scientifically-informed.)

What actions are needed and when?

What is your confidence?

How do bring science to bear into the decision process?

What does it mean?

What are the risks?

What happened?

Can it happen again?

…

What are the right questions?

Are the right people asking?

Are we positioned to answer them?

3. AI & Prediction

So what is AI?
My view on some of its elements

- **AI is not a thing, or a technology. It is a category of empowering ideas that are still nascent.**
- **It can be a collection of the latest ideas combined with 50 year old approaches**
- **If you tell me "These things are part of AI", I will say "sure, why not". In the end, time will tell what is part of AI and what is not, based on what you are able to do.**

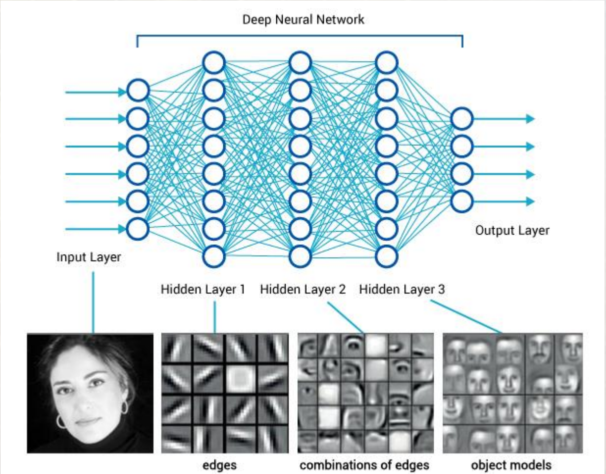
AI today includes:

- **Data Ingestion/Sensors**
- **Learning from data**
  - *Machine learning, including deep learning, supervised/unsupervised/ continuous learning,…*
  - *Surfacing questions from the data we did not know to ask*
- **Deciding/Proving**
  - *Understanding whether learned information is actionable for consequential decisions*
  - *Smart inference*
  - *Uncertainty quantification*
- **Autonomy/Automation**
- **Human/AI interface**
- **Lives in rich data environments**
- **Applied at source of data creation**
- **"+ anything you want"**

Deep Learning is a part of AI

We hear a lot about Deep Learning. Why do we turn to ML/DL?

- When patterns exists in our data - *Even if we don't know what they are*
- When we don't know the underlying mathematical relationships
- Data is of high-dimension and we want to discover a lower-dimension representations – "Latent Space"
- When we have lots of data



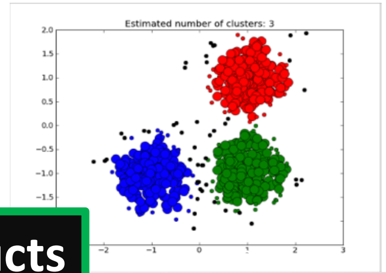
'Opposite' to our traditional model-based prediction

Supervised Learning

- There is some objective we expect machine to accomplish
- We label data to tell machines what to look for
- Tune algorithms until some performance threshold is reached

Unsupervised Learning

- Algorithms are discovering some sub-structures in data
- Power usage spikes are happing at 1 am at zip codes x, y and z.



AI problems wont be solved with only off-the-shelf products

Why is now the time to drive AI?

AI is a roughly 70 year old discipline

Two main factors are driving innovation today in AI

We can create tidal waves of data & information

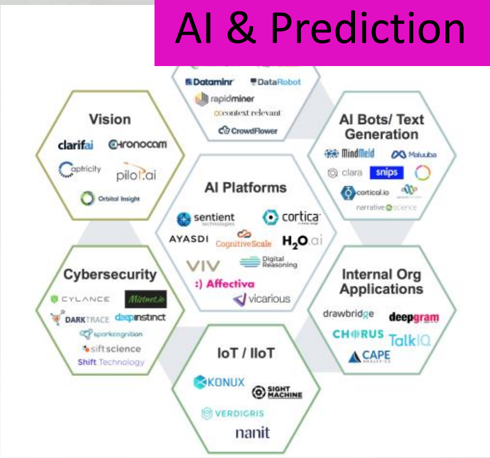
Today we are creating, sensing, measuring, computing, imaging,… more data than we can humanly deal with. Decision support using all this data is needed in all mission, business and operational areas.

We can create altogether new computer architectures

Designed specifically to support large-scale data analytics and machine learning. These systems integrate stream processing at the network-sensor interface with high-performance analytics for integrated data spaces.

We are finally in a position to test ideas that until now we could only dream about.

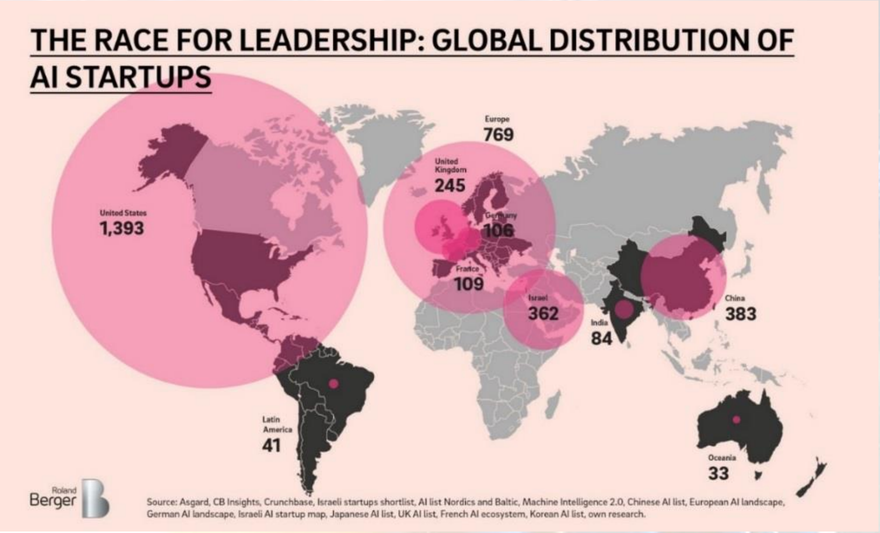
The Race for Global Leadership will Require Partnerships



- Partnerships for tools, hardware, data, etc, combined with federal labs, data and problems

Challenges:

- Frontier technology is currently driven by the private sector
- Leadership forming in competitive and adversarial countries
- Growing technology gap in AI expertise in government
  - Government is a consumer and does not have trusted sources



AI is a foundational technology that is disrupting the current landscape and will lead to decades of innovation

So how can we lead this?

AI and ML tools and approaches are enabling exploration of data and information well beyond what humans can comprehend

Deduction → Experiment/Empiricism

AI based methods now enable richer understanding of complex data

Induction → Theory/Models

HPC & Computational science enables solving far more complex theories

Latent space models to supplement theory

Model based predictions to fill gaps in experimental knowledge

UQ should live here

Just two pillars. HPC and AI are enablers to extend human abilities to deal with complexities of big data and big models. But the need to make predictions that matter will drive these two towards each other.

More challenging as AI is becoming an integral part of how we address urgent issues:

Two approaches, but they are evolving and becoming more interdependent:

Deduction → Experiment/Empiricism

Data, sensors, storage, measurements →

AI based methods now enable richer understanding of complex data

Induction → Theory/Models

High-performance computers →

Computational science enables solving far more complex theories

Future:

Active learning workflows

Actionable Predictions (with confidence margins)

AI tech & hybrid HPC/AI systems

We need trustworthy AI that is accurate with high confidence, proven to be unbiased and reliable. This is where DOE needs to push hard and think big.

3. DOE and AI

Some mission related areas:

- ***Clean energy & Climate resilience***
- ***Effective all-hazard response to energy sector disruptions***
- ***Potential Fraud Detection applications***
- ***Methane leak detection***
- ***Infrastructure: Surge, line slack, security***
- ***Future of Scientific Discovery***
- ***Assuring nuclear deterrent capability***
- ***Grid reliability and resilience***
- ***Oil & Gas***
- ***Identifying and diverting hackers***
- ***Predictive Models for Grid Storage***
- ***Potential sorting of multimodal data sources***

AI is core to DOE's future

AI impacts DOE across all its missions, business and operations

"Standing up to meet the Administration's goals from climate to clean energy will require teaming and focused scientific advancements at far different scales. And AI provides the organizing principle to creating new discovery and technology workflows."

Secretary Granholm (2021)

AI is affecting many technologies used DOE wide

Intelligent Sensors, Machine Learning, Data Sciences & Data Analytics, Robotics, Autonomous Systems, Data structure and management, Information and Business Management Systems, Edge Devices/Distributed Systems, Natural Language Processing, Human-machine Interface & Biometrics.

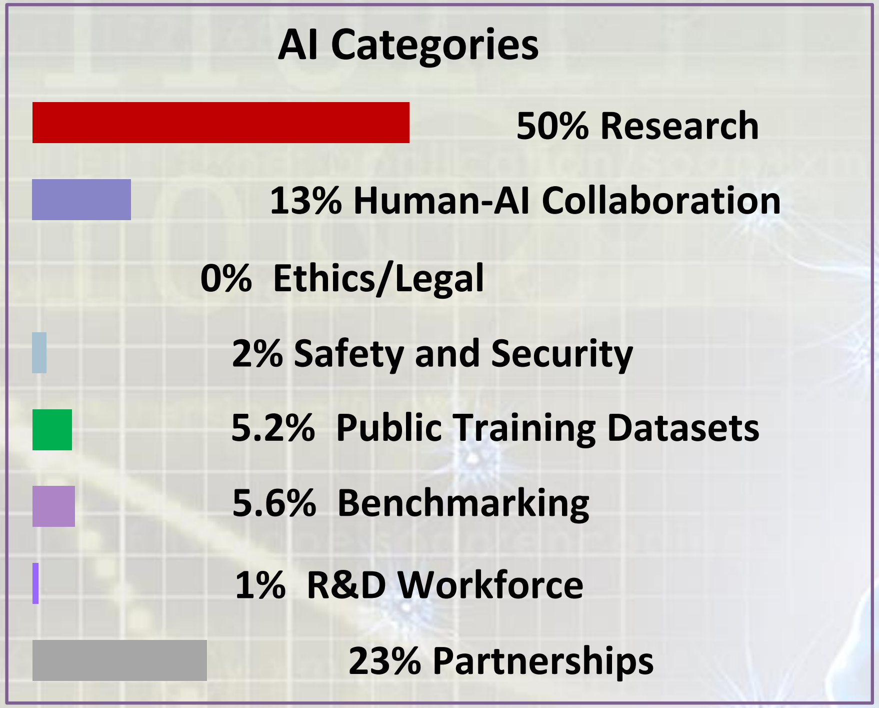
AI has implications for high consequence areas with little room for failure

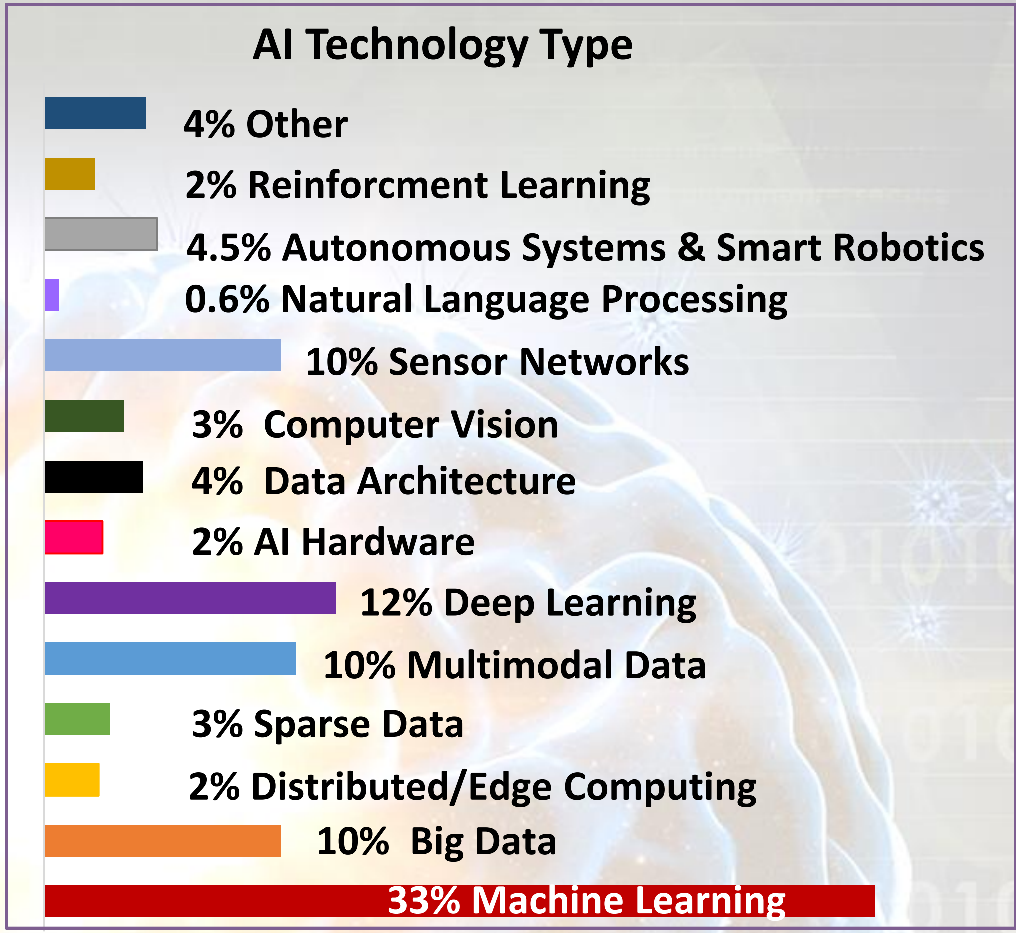
Energy Security, Cybersecurity, Physical Security, National Security, Economic Security & Science, Workforce Development…

We host world-class suite of User Facilities that span HPC/AI to those anchored in the physical, materials, quantum and life sciences and we are rethinking how knowledge discovery is done in fundamentals ways

DOE & AI

A rough snapshot of where we are today





We look at many categories of barriers today as we look to advance AI

- ❑ Insufficient collaborations and coordination between programs
- ❑ Resources/funding
- ❑ Lack of strategy for roadmaps, roles/responsibilities undefined, or top level decisions
- ❑ Data access, trustworthiness, & suitability
- ❑ Legal/regulatory (NDA), etc.
- ❑ Insufficient benchmarking and standards
- ❑ Other: e.g. Nascent capability; Need multidisciplinary skills to apply, deploy, and use AI

AI for Science and the Future of Discovery

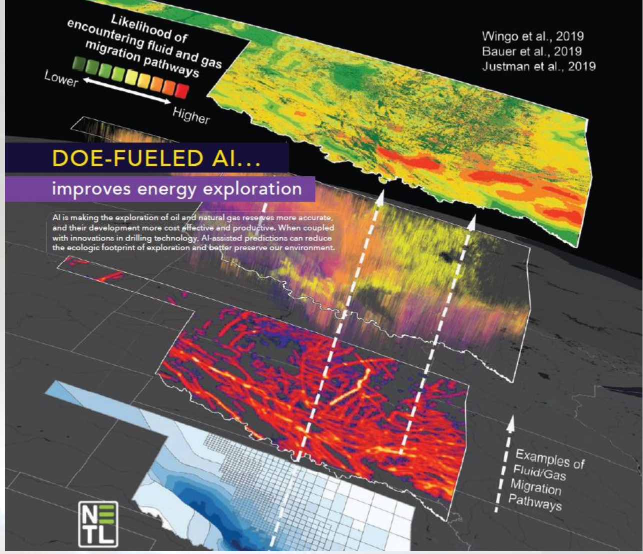
An end-to-end vision starts with the acceleration of discovery rates due to faster experimental cycles, but must be much more:

- Accelerated Discovery
  - Anomaly detection in events, images…, source/transient classification
- Interpretation of Measurements
- Using learned models and the latent space to augment data
  - *ML based models with UQ – observations, measurements, interpretation of simulations and experimental outcomes*
- Acceleration of discovery rates due to faster experimental cycles
  - *Targeted search, optimization, automation*
- Smart facilities and instruments
  - *Semi-autonomous science driven by active learning loops*
- Simulation + AI hybrids data
- Accessible and Integrated Knowledge bases
  - *New interfaces to the literature, data and models*
- AI everywhere, smart processes, smart data, smart simulations

Our job is to think big, beyond projects to transformational challenges.

4. Opportunities and Challenges

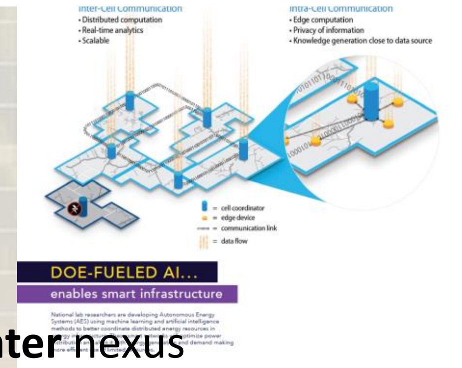
AI for Energy, Climate and Environment



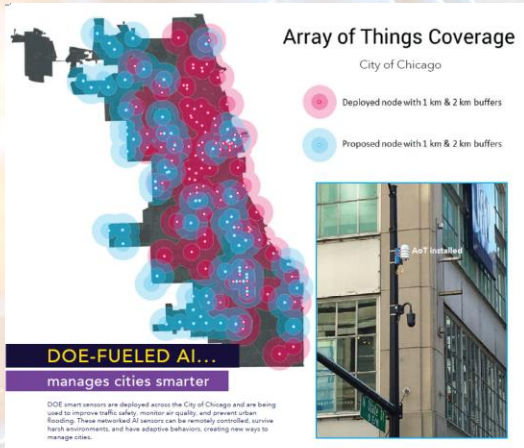
Challenging problems are just about everywhere. For example:

- Higher efficiency for wells using with smarter subsurface
- Converting biomass into energy
- Vegetation management/fires
- Oil spill prevention
- Micro/nano-plastics in the environment

- Traffic, Transportation, Vehicles
- Improving commuting, shipping
- Fuels, fuel economy, engine design…



- Smarter buildings
- Smarter homes
- Smarter microgrids and load management for changing demands due to renewables
- Grid Resilience
- Next gen networks (5G,…)



- **Energy/water nexus**
  including water availability, its transportation, its use to energy generation, fuel supply,…
- **Smart farming**, soil management, more efficient use of phosphates/fertilizers, more direct use of water
- Genetically engineered plants to be more **climate resilient** (weather variations, drought/flood/hot/cold)

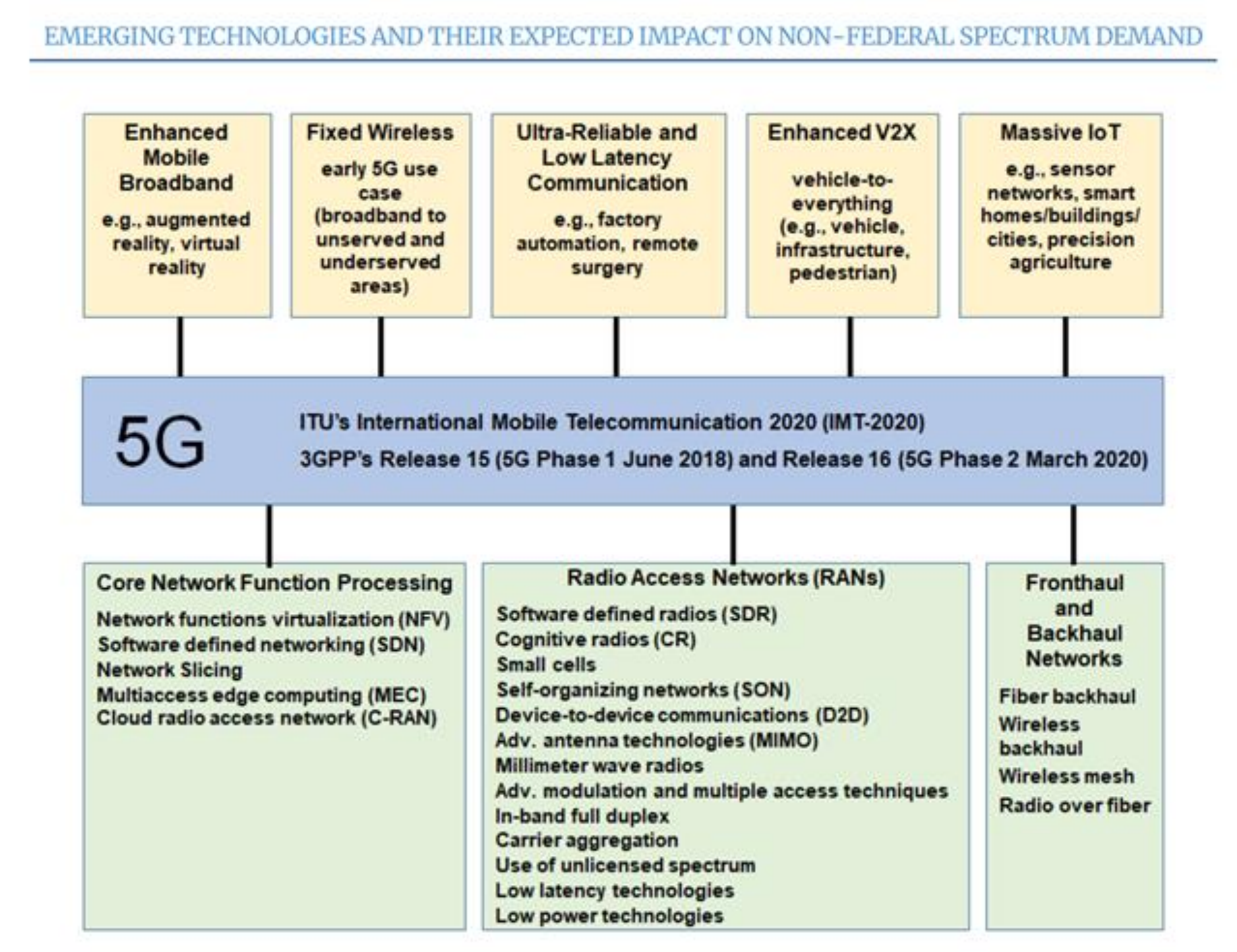
5G and future networks

I like 5G and the future networks because of the opportunities they will provide to innovate. Why is this an opportunity?

A few things to consider:

Many more devices will be able to

- Talk to each other;
- Measure, store, process/learn from information on the fly in ways not previously possible;
- Make decisions and take actions;
- Create a distributed intelligence;
- Communicate more data and much faster;



5G, 5G-Enabled Technologies (top), and 5G-Enabling Technologies (bottom)

https://www.whitehouse.gov/wp-content/uploads/2019/05/Emerging-Technologies-and-Impact-on-Non-Federal-Spectrum-Demand-Report-May-2019.pdf

Adversarial AI

AI today is fragile

Think about it like the internet before we worried about cybersecurity

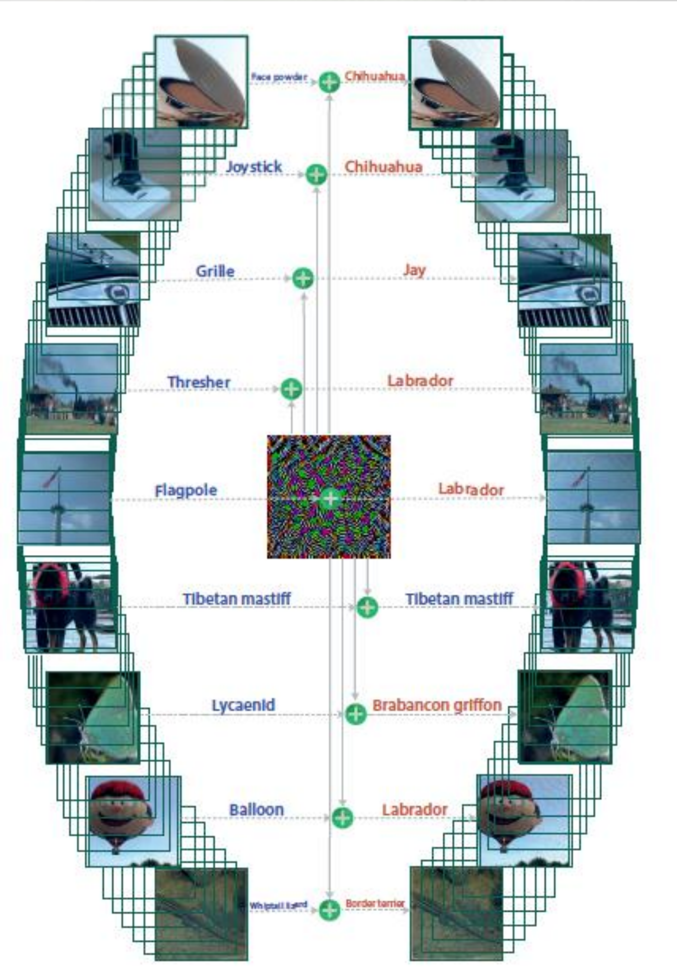
Tools – new chips and methods – are being developed today and we have the opportunity to build in what we really need

Today it can be easily fooled

- *Single pixel defeats of methods – identical to humans, distinct to AI*
- *Miss-identification/Impersonation/Dodging*
- *Ignoring visual objects/events (e.g. stop signs,…)*
- *Data poisoning/reverse engineering models…*

Many new weaknesses being surfaced to help improve AI

For decision support, we need measures of certainty in predictions



We cannot just be a consumer, but need to be fully invested in the intellectual development

There are broad classes of opportunities important to us today including:

I. Applying AI/ML: Learning today across missions
- *Applications of optimization, statistics,… to existing data sets*
- *Pushing the scale of learning and graphs on large GPU systems*
- *….*

II. 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*

III. 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,…*

IV. 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*

Hard problems at DOE scales are drivers to put solutions into practice

Deep partnerships with Industry will be needed

AI today is global and innovation is emerging everywhere. This is something we must build on and where public-private and international partnerships could play a role.

Markets will also need to be created that are foundational to progress on the energy and climate goals. It is likely pre-competitive partnerships that can foster spin-off technologies could be part of this.

It must be more than transactional or progress will be tied to private sector interests that are not necessarily aligned with DOE goals.

Some AI Drivers:
Uncertainty Quantification (UQ)

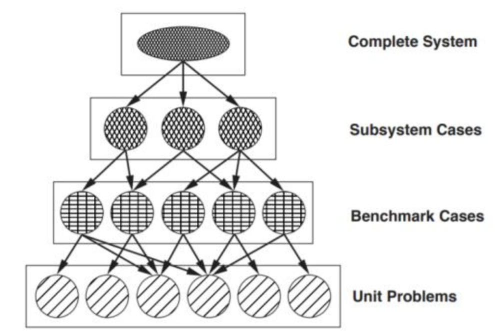
Basic issue: What is the mathematical confidence we can ascribe to complex simulations we perform?

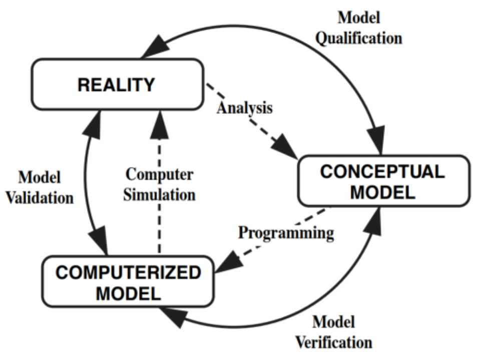
We need pragmatic solutions since we have missions to deliver. Need mathematical frameworks that are implementable

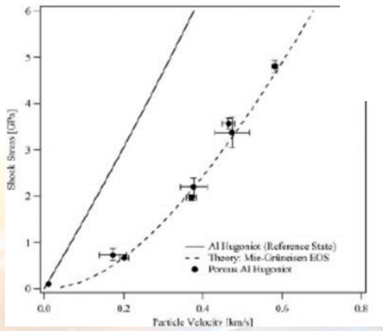
But today I believe that our approach to UQ is dated

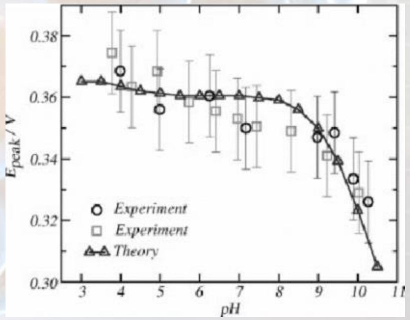
- *Usually start in a space of discretized space of guiding equations*
- *Run a deterministic code end to end*
- *Bayesian approaches, Latin hypercube,…*
- *Post-hoc add-on of tools to: accrue uncertainty, better sample initial conditions, parameters,…*

UQ is also a hard problem. Identifying the underlying dynamical equation from any amount of experimental data, however precise, is a provably computationally hard problem (it is NP hard), both for classical and quantum mechanical systems.









5. Summary

Summary

- The AI landscape is a game changer – we are in a unique position to make an impact
- As with HPC, AI will be part of our decision cycles on critical parts of our missions.
- We cannot afford to be simply a customer - we need to stress the tools and understand how they fail so we know how to trust them
- The impacts of driving deeper applications of AI to hard problems will be broadly beneficial to DOE and all of our mission, business and operational functions
- It should not be business as usual in how we approach problems
- The benefits are societal