Barriers to Reproducibility and Automation

D. A. Brown
National Nuclear Data Center
Brookhaven National Laboratory

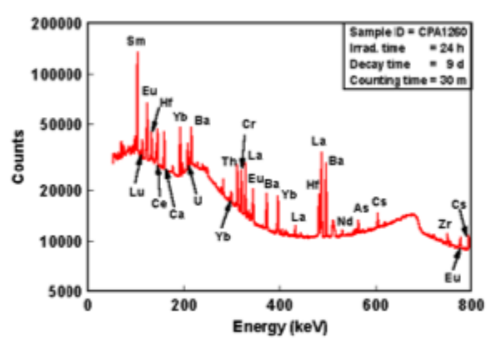
BROOKHAVEN NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

The Nuclear Data Pipeline

Our goal is to get the highest quality data to users

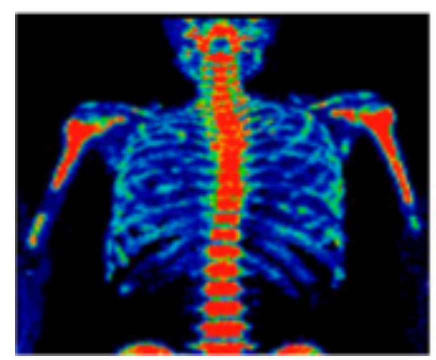
security



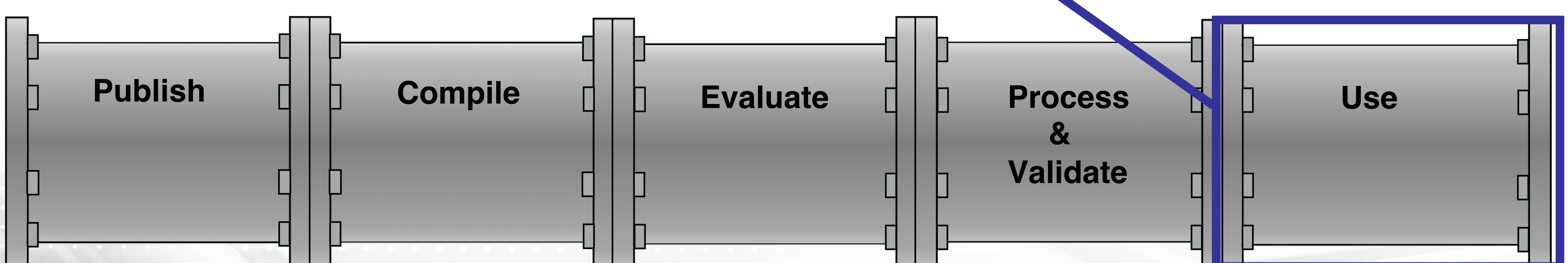
science

FRIB

isotopes

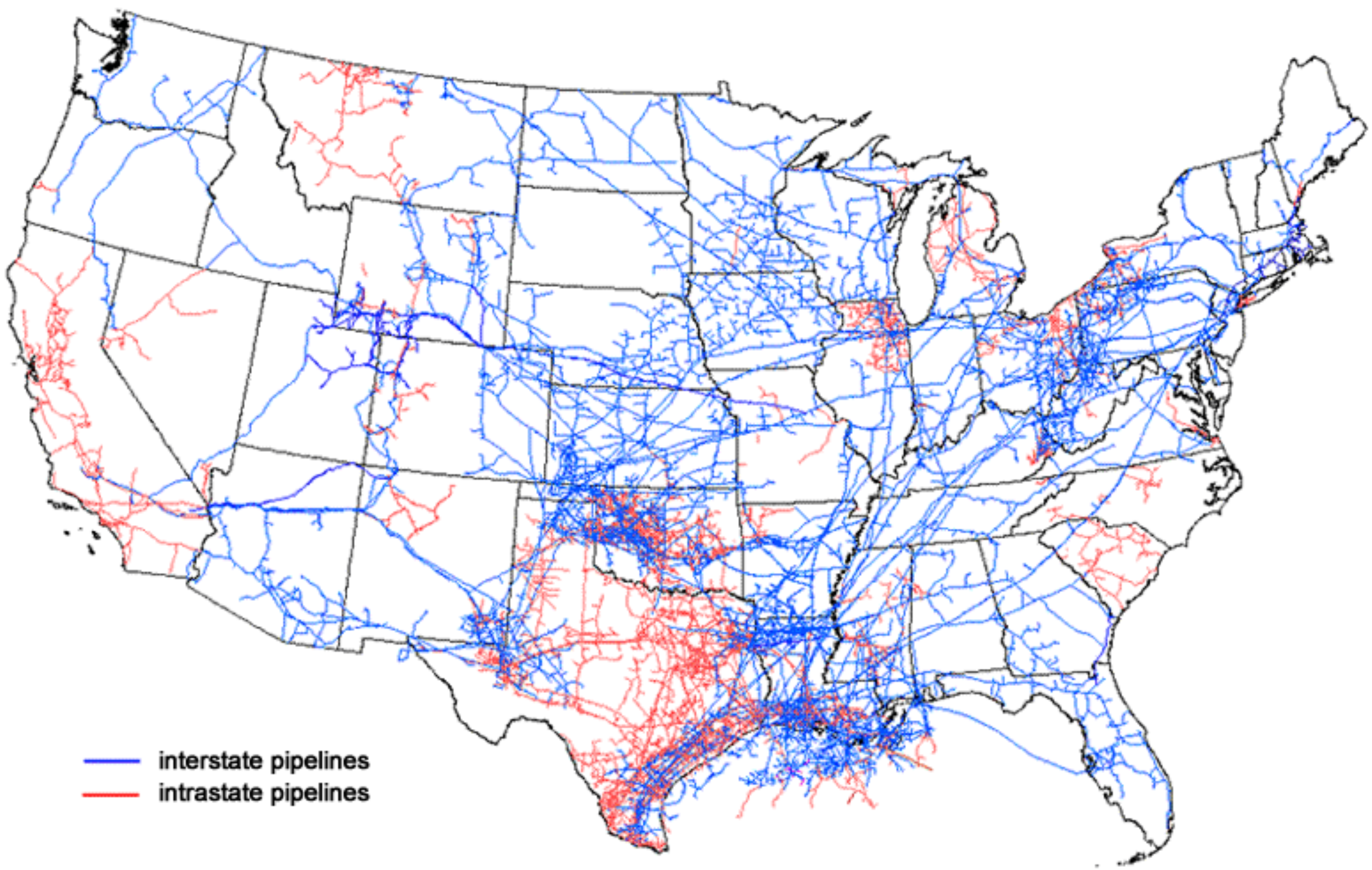


energy



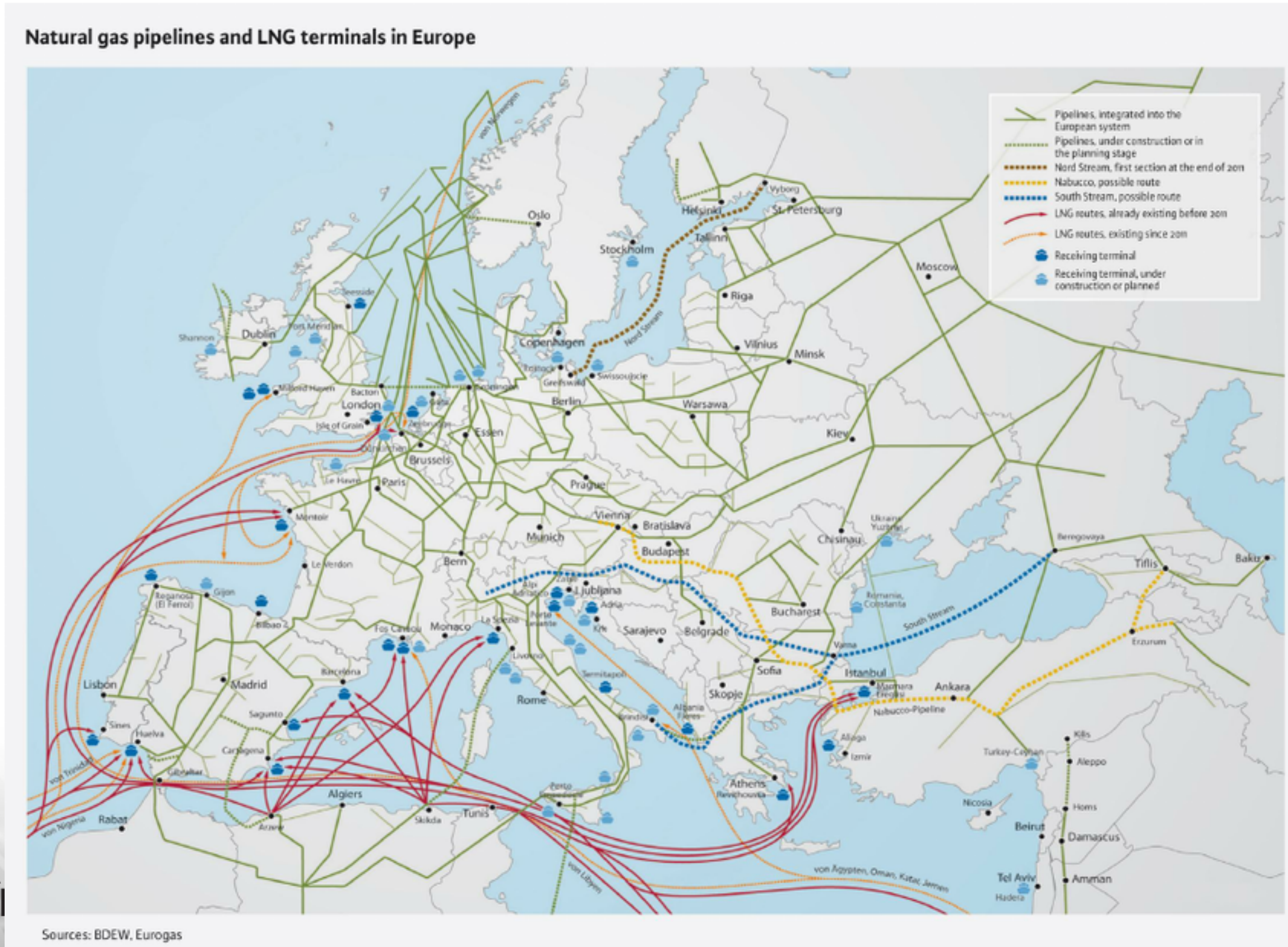
The nuclear data pipeline is more of a network

Map of U.S. interstate and intrastate natural gas pipelines

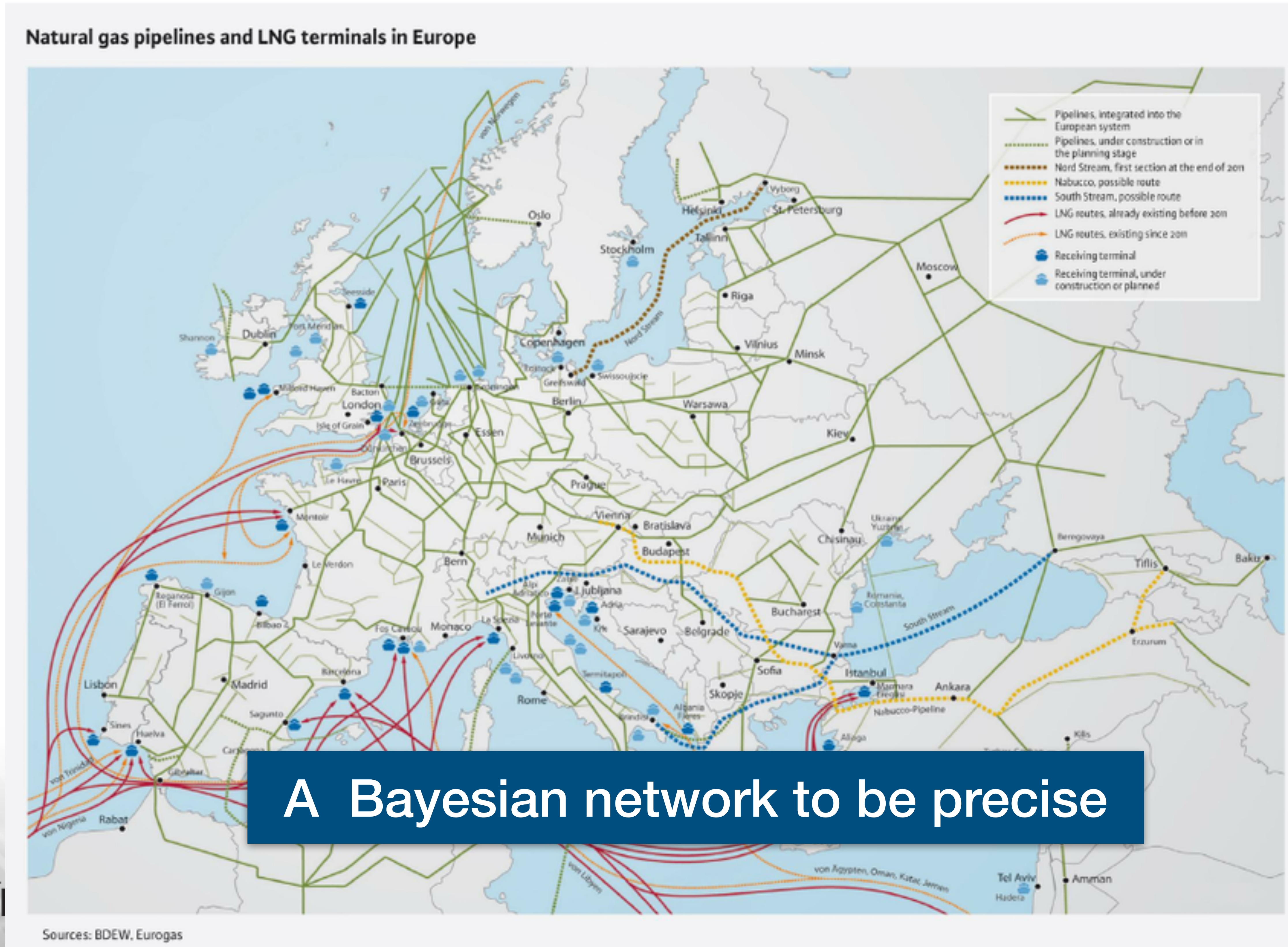


Source: U.S. Energy Information Administration, *About U.S. Natural Gas Pipelines*

The nuclear data pipeline is more of a network

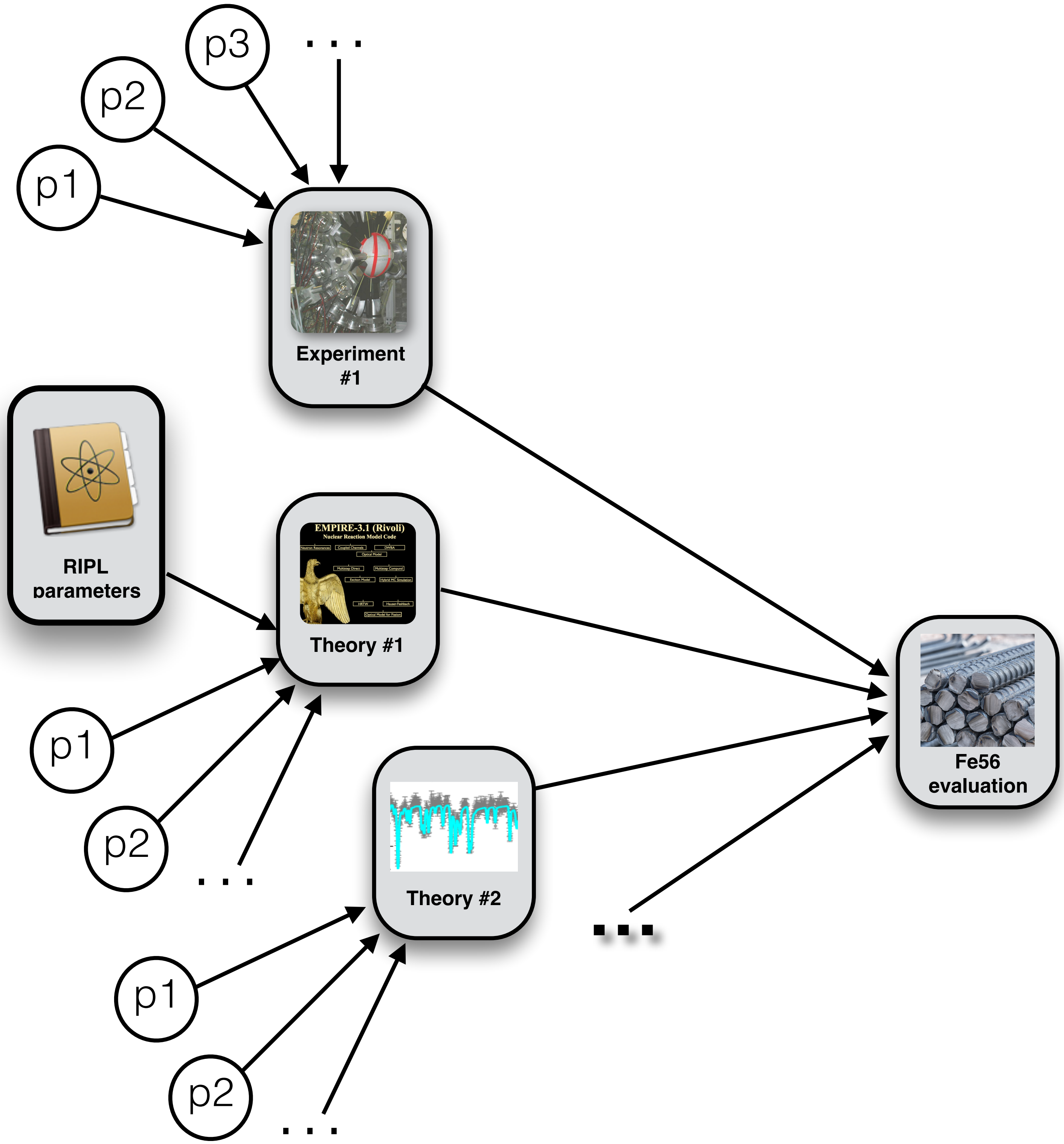


The nuclear data pipeline is more of a network

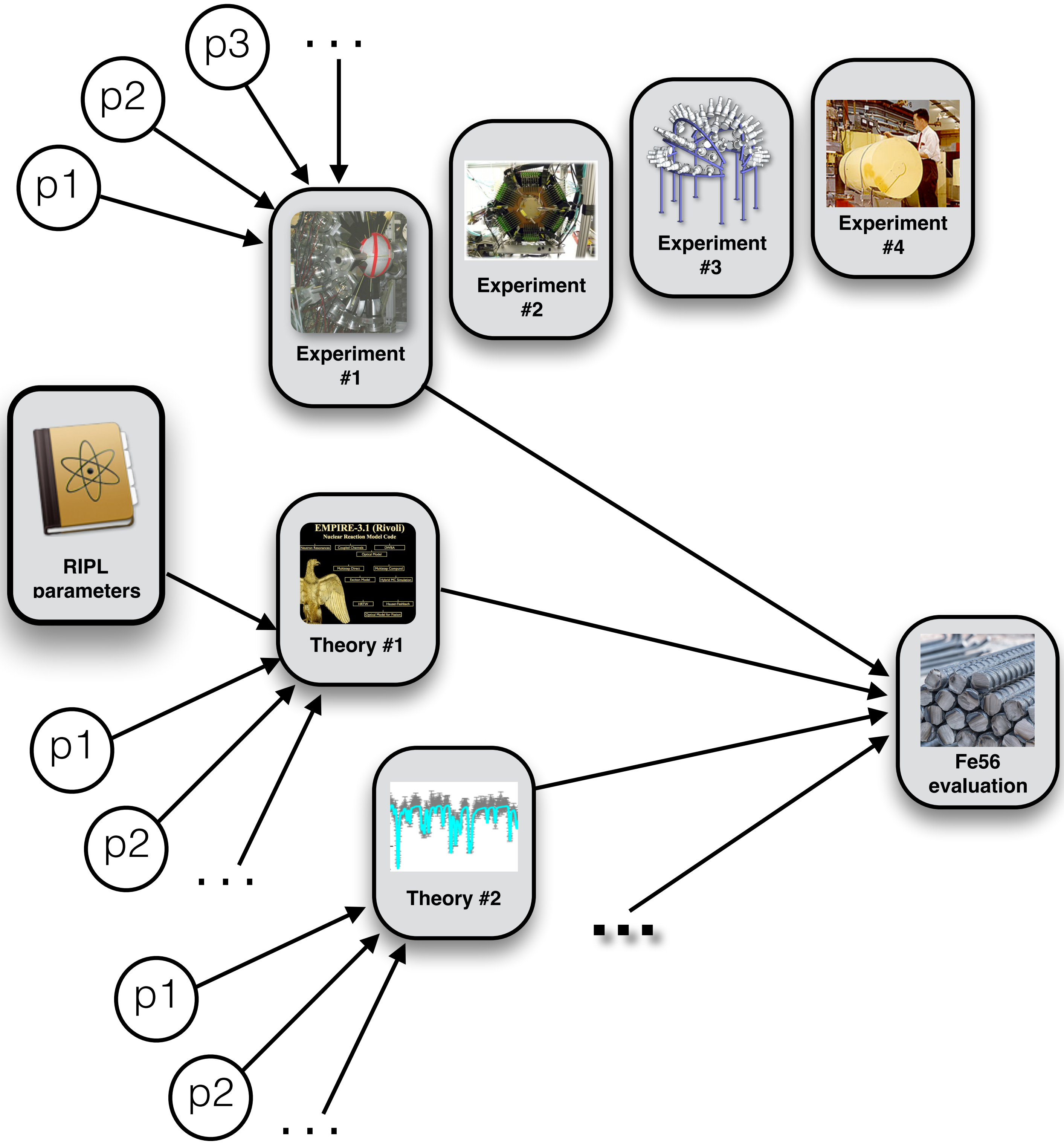


A Bayesian network to be precise

We want to take an evaluation



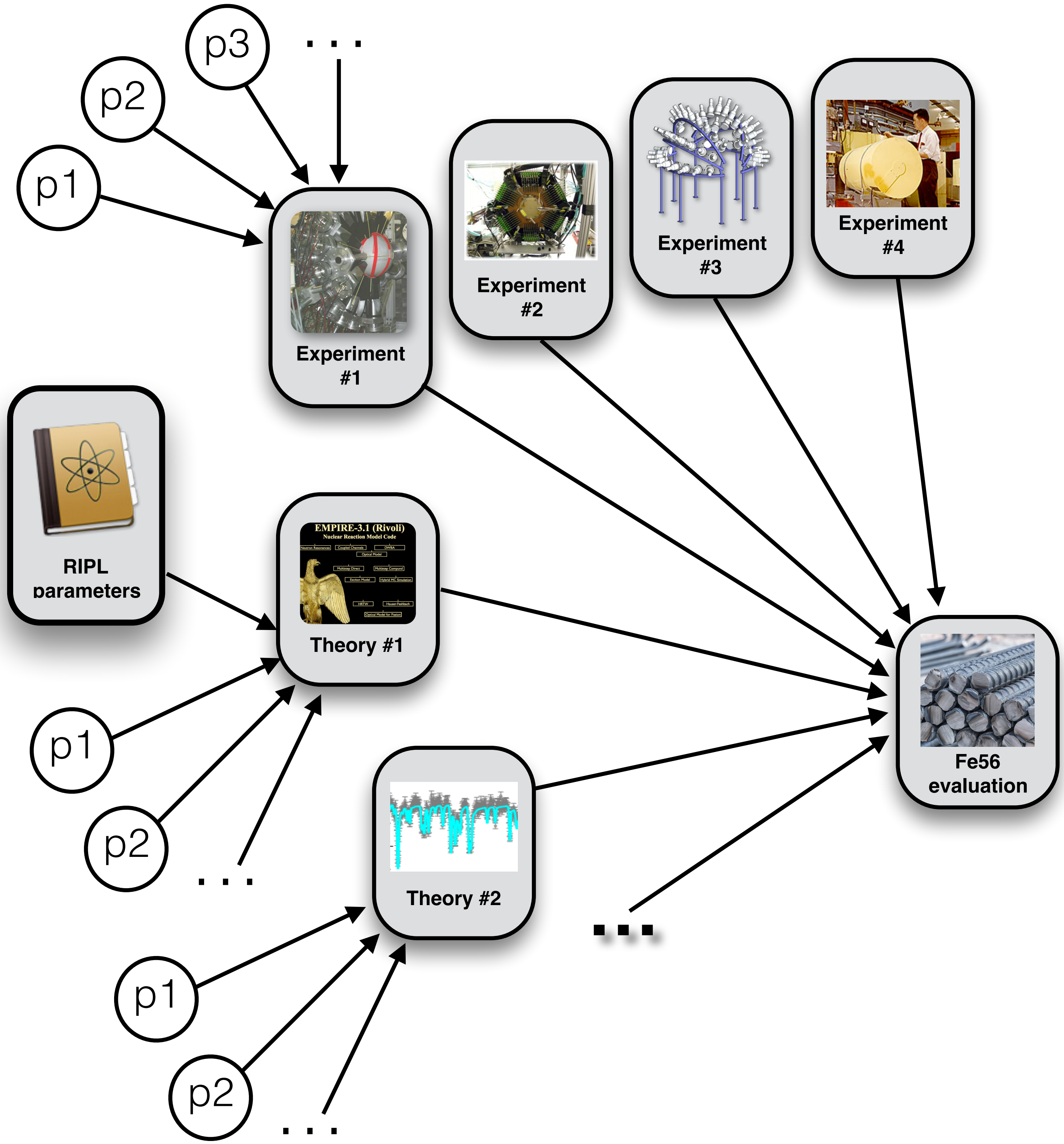
add a few new experimental results



have something magical involving containerized applications happen

docker

… and get a newly updated evaluation



Including integral data too

BROOKHAVEN
NATIONAL LABORATORY

What follows is based on conversations during the 2019 CSEWG meeting with various CSEWG members

There are many things keeping us from this vision

Barriers caused by EXFOR

NRDC

ENDF B-VIII.0

NEA

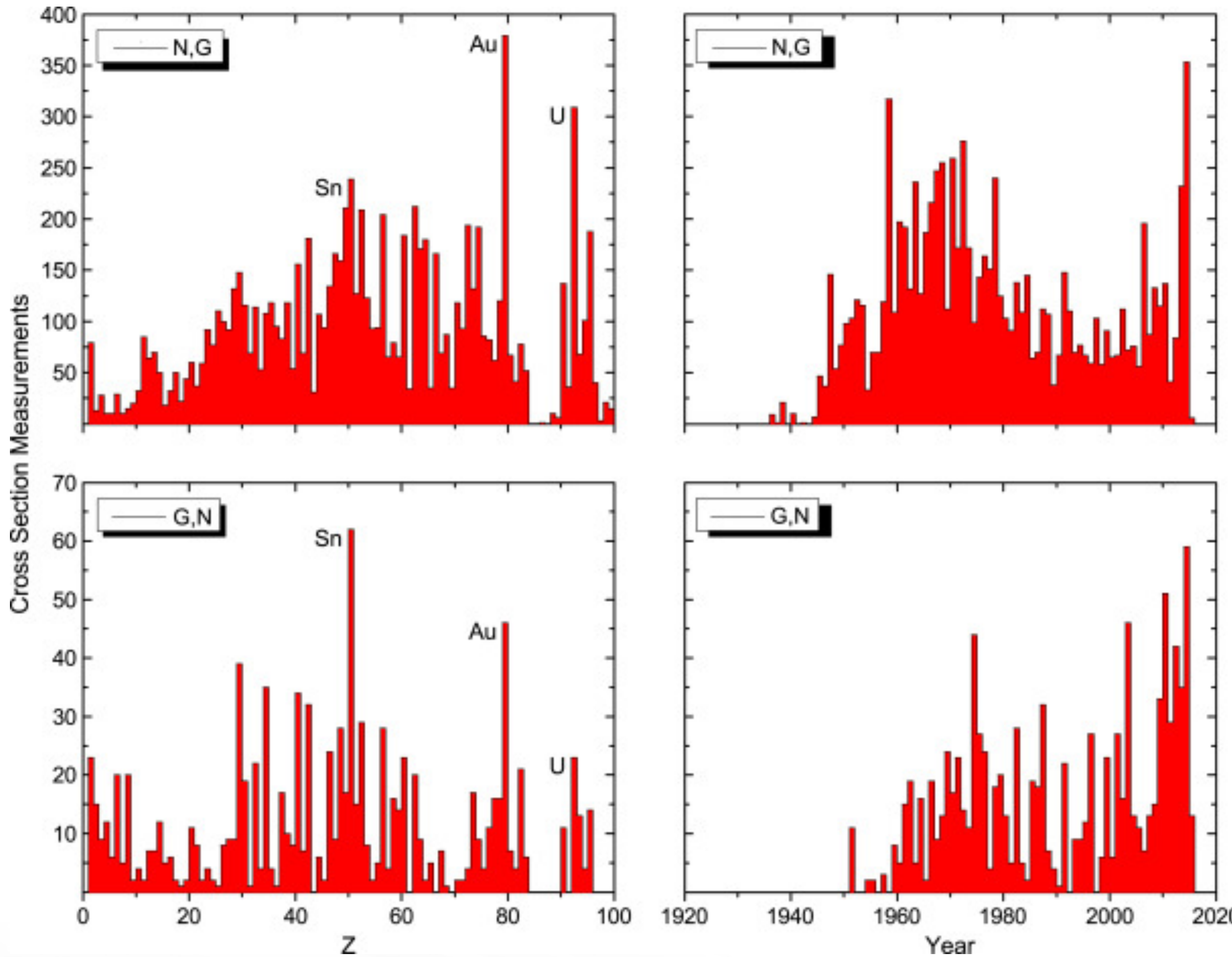
Rsicc

EXFOR and NSR Compilation is On-going

EXFOR compilations worldwide conducted by NRDC network

22,633 experiments compiled

At 1M€ per experiment, this is sizable investment

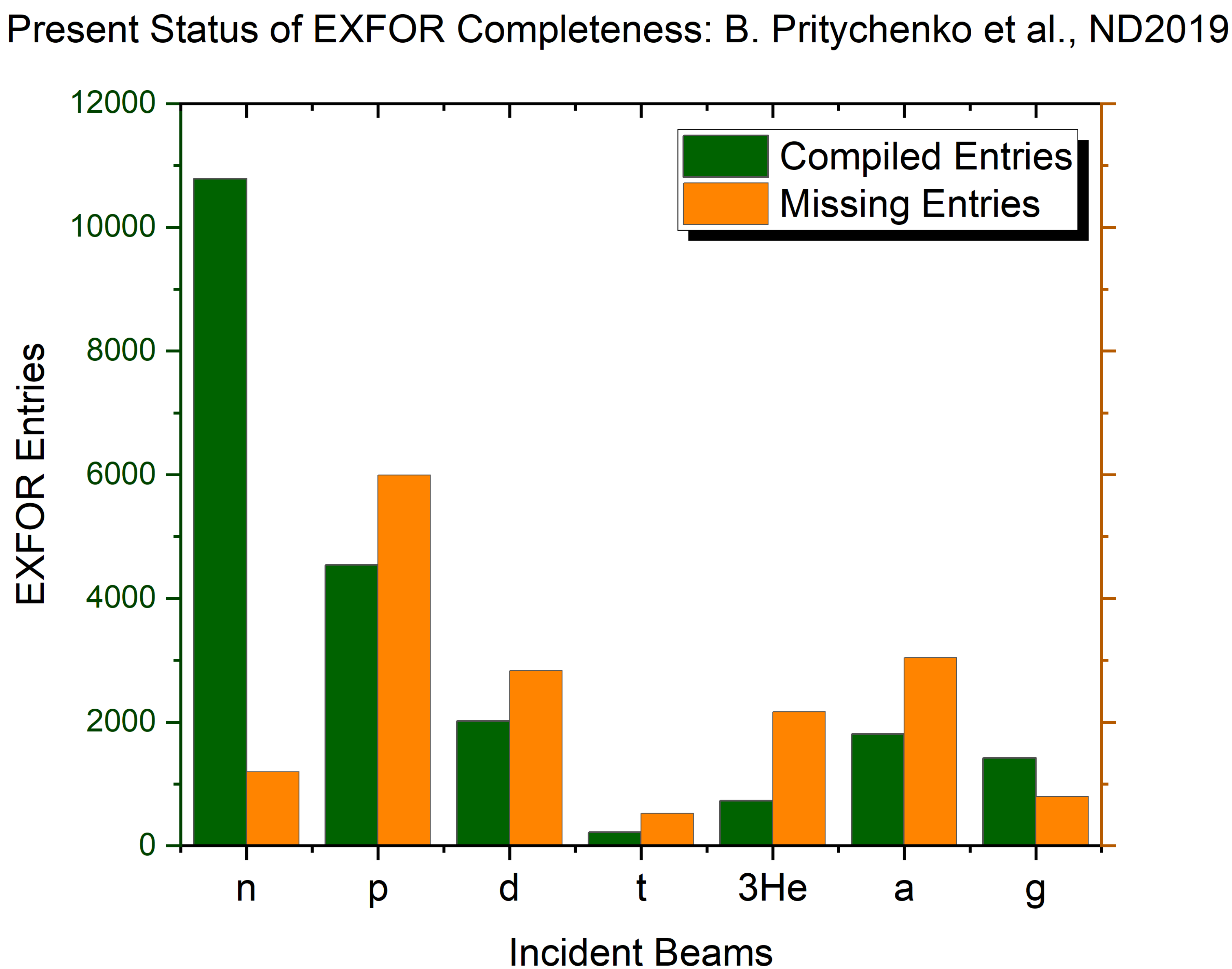


Number of evaluations compiled by Z and year (B. Pritychenko)

U.S. DEPARTMENT OF ENERGY

BROOKHAVEN NATIONAL LABORATORY

EXFOR compilation is incomplete



EXFOR is an invaluable resource for long-term storage of experimental data, and allows users to quickly find data

- It is often the first place evaluators and experimentalists search for experimental data
- EXFOR represents long term storage of experimental data in a consistent format
- Many entries have uncertainty information, so is a great resource for uncertainties presented over time
  - The ERR-ANALYS section does not have a consistent format

Slide from A. Lewis, CSEWG 2019

Extracting uncertainty information from EXFOR isn't always easy

13901 – 2004 Fotiades $^{238}U(n,n'g)$

```
ERR-ANALYS (ERR-S) Statistical uncertainty.
           (ERR-1,5.,11.) Uncertainties in the gamma-ray
           absorption in the sample, finite beam size effects,
           well as detector efficiency uncertainties
           (ERR-2) Uncertainties in the fission foil thickness,
           fission cross section, ionization chamber efficiency
           (ERR-3,1.,2.) The uncertainty in the neutron flux
           (ERR-4) Uncertainty in target thickness
           (ERR-5) Uncertainty in dead time corrections for
             the detection of gamma rays
           (ERR-6,0.15,0.2) Uncertainty in dead time correction
             for neutrons in the fission chamber
```

Fotiades et al Phys. Rev. C. 69 (2004

Slide from A. Lewis, CSEWG 2019

Extracting uncertainty information from EXFOR isn't always easy

	1998 Data (%)
E_γ (keV)	$\delta\epsilon_\gamma$ [a]
100–400 (Planars)	5
600–750 (Coaxials)	10
751–900 (Coaxials)	9
901–1200 (Coaxials)	8
>1200 (Coaxials)	7
E_n (MeV)	$\delta\Phi(E_n)$
1–4	1.0
4–8	1.5
9–19	2.0
20–50	1.5
51–100	1.2
δt	0.3
δ Dead T_γ	0.1
δ Dead T_Φ	0.2
Additional fluence uncertainty[b]	5

Fotiades et al Phys. Rev. C. 69 (2004)

13901 – 2004 Fotiades ^{238}U(n,n'g)

```
RR-ANALYS (ERR-S) Statistical uncertainty.
          (ERR-1,5.,11.) Uncertainties in the gamma-ray
          absorption in the sample, finite beam size effects,
          well as detector efficiency uncertainties
          (ERR-2) Uncertainties in the fission foil thickness,
          fission cross section, ionization chamber efficiency
          (ERR-3,1.,2.) The uncertainty in the neutron flux
          (ERR-4) Uncertainty in target thickness
          (ERR-5) Uncertainty in dead time corrections for
            the detection of gamma rays
          (ERR-6,0.15,0.2) Uncertainty in dead time correction
            for neutrons in the fission chamber
```


Slide from A. Lewis, CSEWG 2019

Barriers caused by the ENDF format

ENDF B-VIII.0

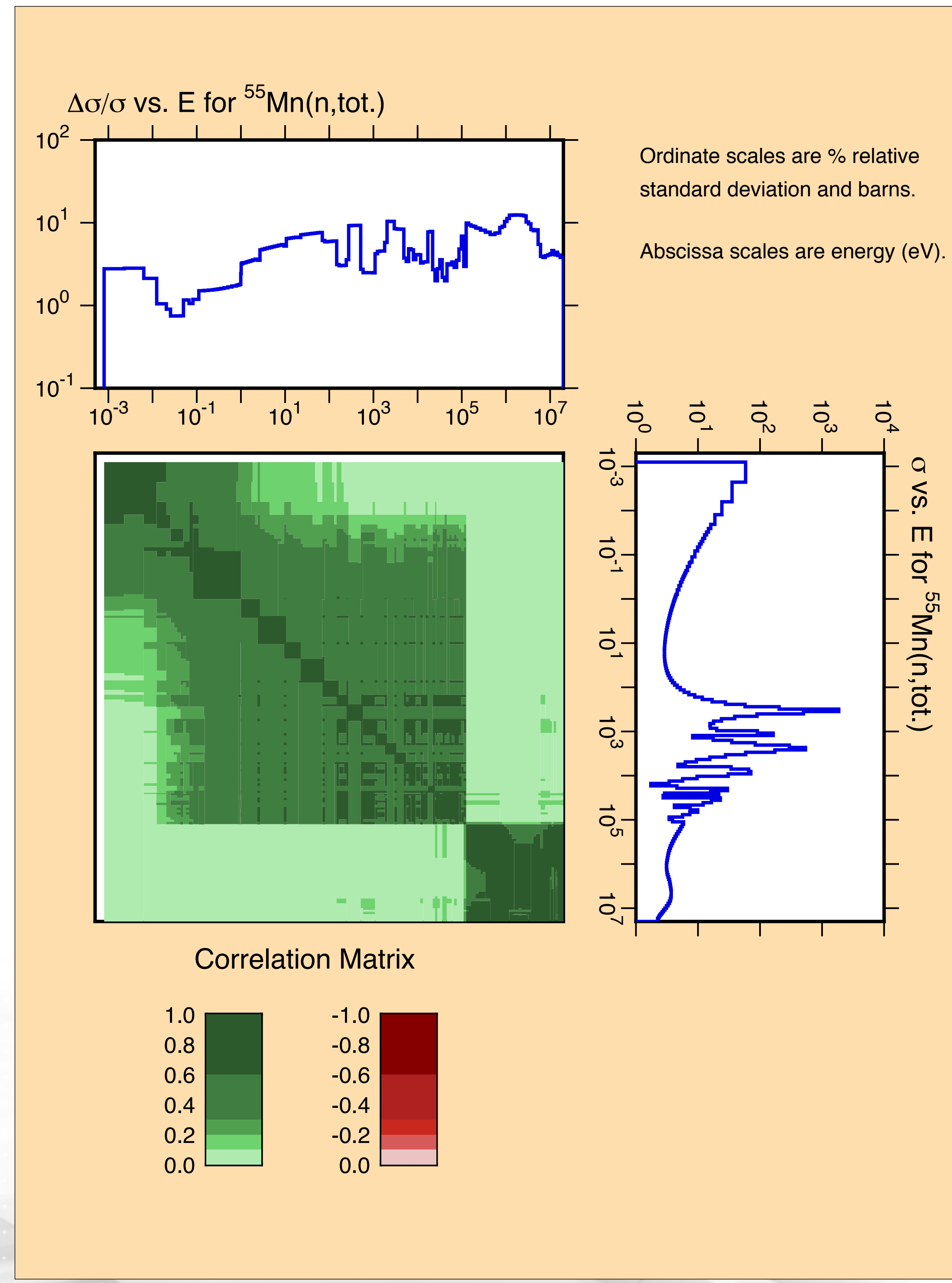
ENDF format strives to provide a Gaussian Process Regression model

- For a given reaction *rxn*, every emitted particle *p*, store

$$\sigma_{rxn}(E)$$

$$P_{rxn,p}(E',\mu\,|\,E)$$

- both as linear interpolatable functions
- and, the covariance matrices for each (what that means is a different question…)



Missing or extremely limiting covariance data format

- **Thermal Scattering Law data**
  - **Nothing!**
- **Fission Product Yield data**
  - Y±ΔY only
- **Decay data**
  - Discrete energies,
  - Q,
  - $T_{1/2}$,
  - Branching ratios,
  - ICC
- **Atomic data**
  - **Nothing!**

} criticality, reactors

} spent fuel, decay heat, etc.

} radiotherapy, shielding

Full format, but (basically) no data

- **protons**
- **deuterons**
- **tritons**
- **helions (^{3}He)**
- **alphas**
- **photonuclear**

} fusion (protons, deuterons, tritons, helions, alphas)

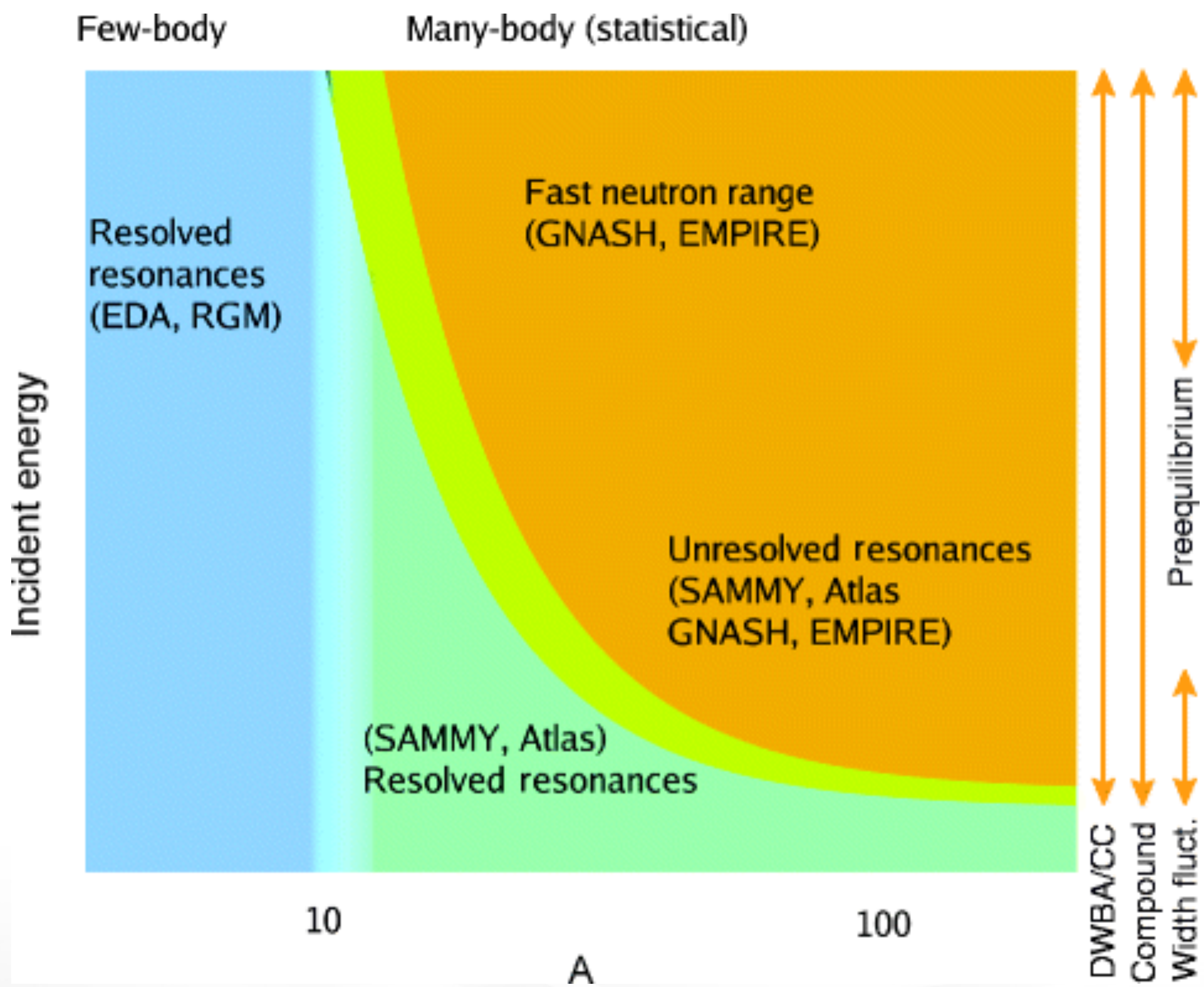
non-proliferation, assay (photonuclear)

Neutron sub library contains nearly all the covariance data in ENDF/B

- 30: any parameters — unused
- 31: nubar — widespread use
- 32: resonance — widespread use
- 33: σ(E) — widespread use
- 34: P(μ|E) — missing correlations; limited use
- 35: P(E'|E) — missing correlations; limited use
- 40: Y(E) — limited use
- energy release in fission — Big 3 only

Models are used to generate the GPR model

- Some are predictive (GNASH, EMPIRE, TALYS)
  - Useful extrapolating beyond experimentally known regions
- Some are "complete" (R matrix models)
  - All observables over all energies
  - Must be fit to data
- Least biased are splines



Selection of known *modeling* issues

- R-matrix is not predictive, is only a fit to data
- RRR - URR connection and fluctuations!
- ENSDF - RIPL/model connection
- Level scheme - Level density connection
- Only “predictive” and “global” OMP is Koning-Deleroche: spherical only
- What gamma ray strength function?
- Fission!

Many parameters have non-Gaussian PDFs

- Fission barriers
- Anything with a threshold
- R-matrix resonances are only Gaussian if well measured

Barriers caused by our existing tool chain

Export control codes

- Within the US, most processing and simulation codes are Export Controlled and distributed through RSICC
- RSICC charges for EVERYTHING and any code under their control is assumed to be Export Controlled
- Hence the nickname "code graveyard"
- Is there analogous problem outside of US?

There are strong barriers to both acquiring and using EC codes and codes that are perceived to be EC

Avoid vendor lock

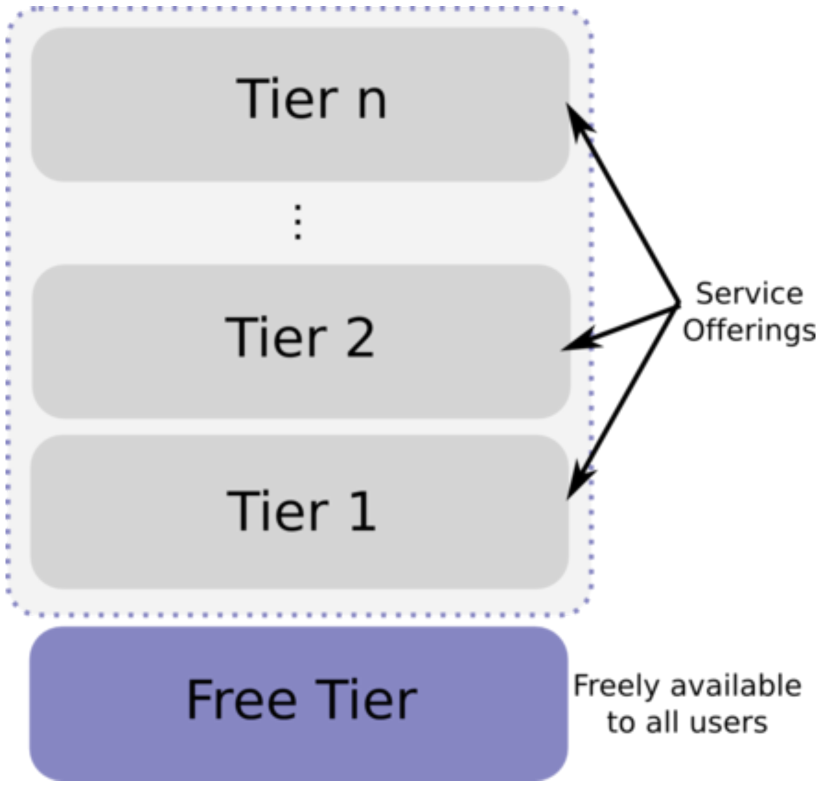
From Wikipedia, the free encyclopedia (https://en.wikipedia.org/wiki/Vendor_lock-in):

> In economics, vendor lock-in, also known as proprietary lock-in or customer lock-in, ***makes a customer dependent on a vendor for products and services***, ***unable to use another vendor without substantial switching costs***. Lock-in costs that create barriers to market entry may result in antitrust action against a monopoly.

This is why the ENDF->GNDS transition is so hard!

Freemium model and ever changing conditions

- Freemium services (e.g. Docker) seem great — their free!
- Terms and conditions change
- “Bait and switch”
- The NNDC has be zapped by this phenomena several times
    - SyBase
    - GForge
    - …

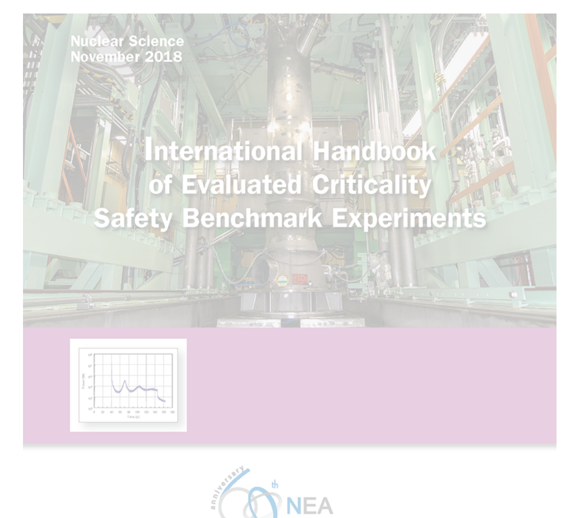


Freemium codes are timebombs

Lessons

- Use openly available, OPEN SOURCE codes only
- Design system to be portable in case we need to switch vendors

Barriers caused by the proposed scheme



Remote execution creates gigantic cybersecurity risks

- Letting someone else run on your computer may **expose your system**
- Running on a cloud system may **expose your data**
- Either requires experience and buy-in from sponsor & institution

Consider Docker containers on open science grids (CERN/DOE/NSF?) rather than private sector service?

Barriers caused by our validation databases

Critical Benchmarks and Sources of Uncertainty

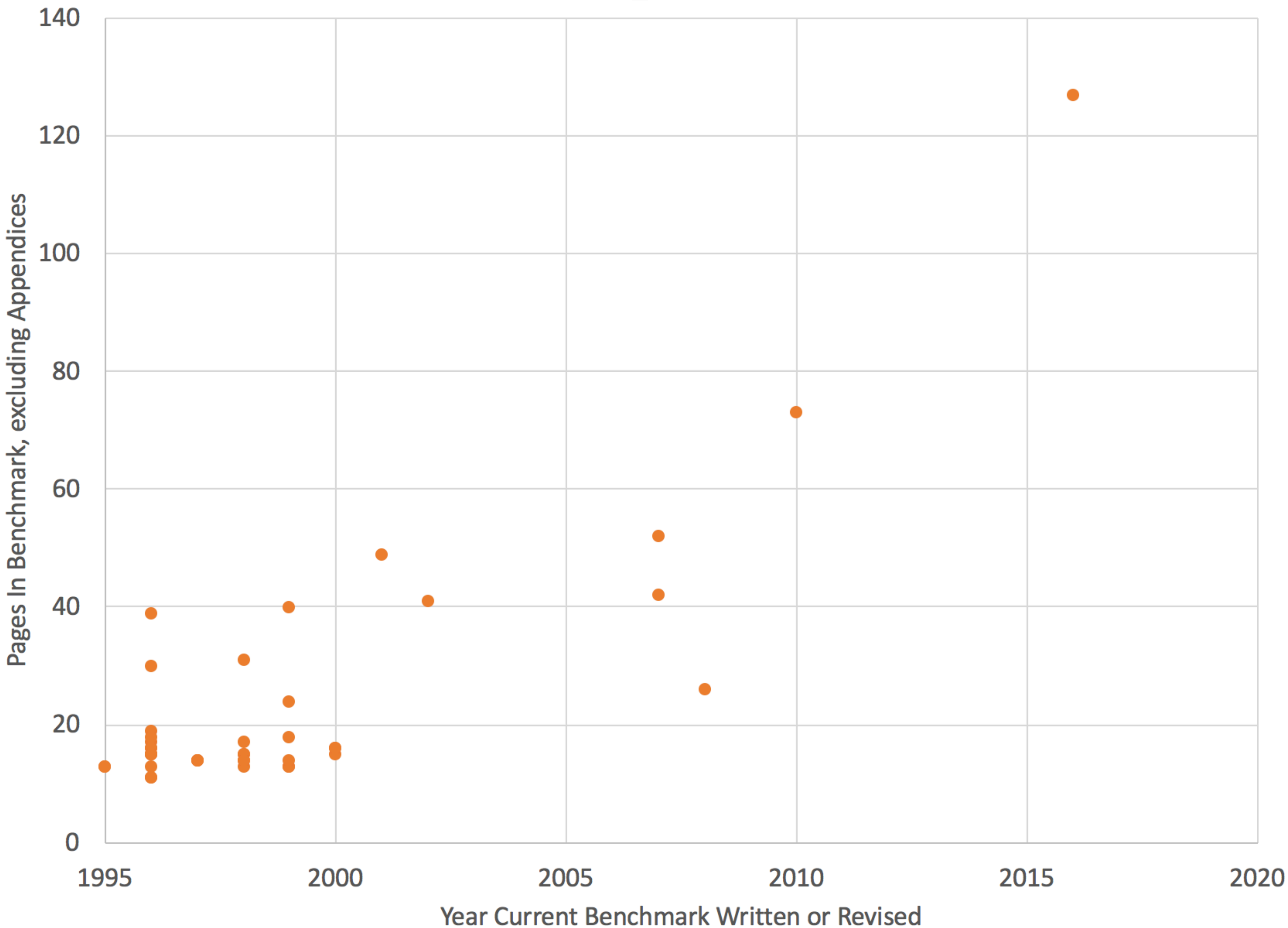
- **Critical Experiments-** Controlled assemblies of nuclear material designed to just achieve the critical point (or slightly lower/higher)
- **Critical Benchmarks-** Computer simulations of the real critical experiment
- **Uncertainties**
  - **Experimental:** How certain are the experimenters of the k_{eff} reported?
    - Uncertainty in measurement technique, reproducibility measurements, etc
    - Usually small contribution
  - **Benchmark Model Uncertainties:** How certain are the evaluators of the benchmark model k_{eff}? Model vs. Reality
    - Uncertainties from mass (are all masses or densities well known?), dimensions (were all parts measured? How do they fit together?), and composition (what are the constituents of all parts, including impurities?)
    - Usually majority contribution

Not All Benchmarks are Created Equal

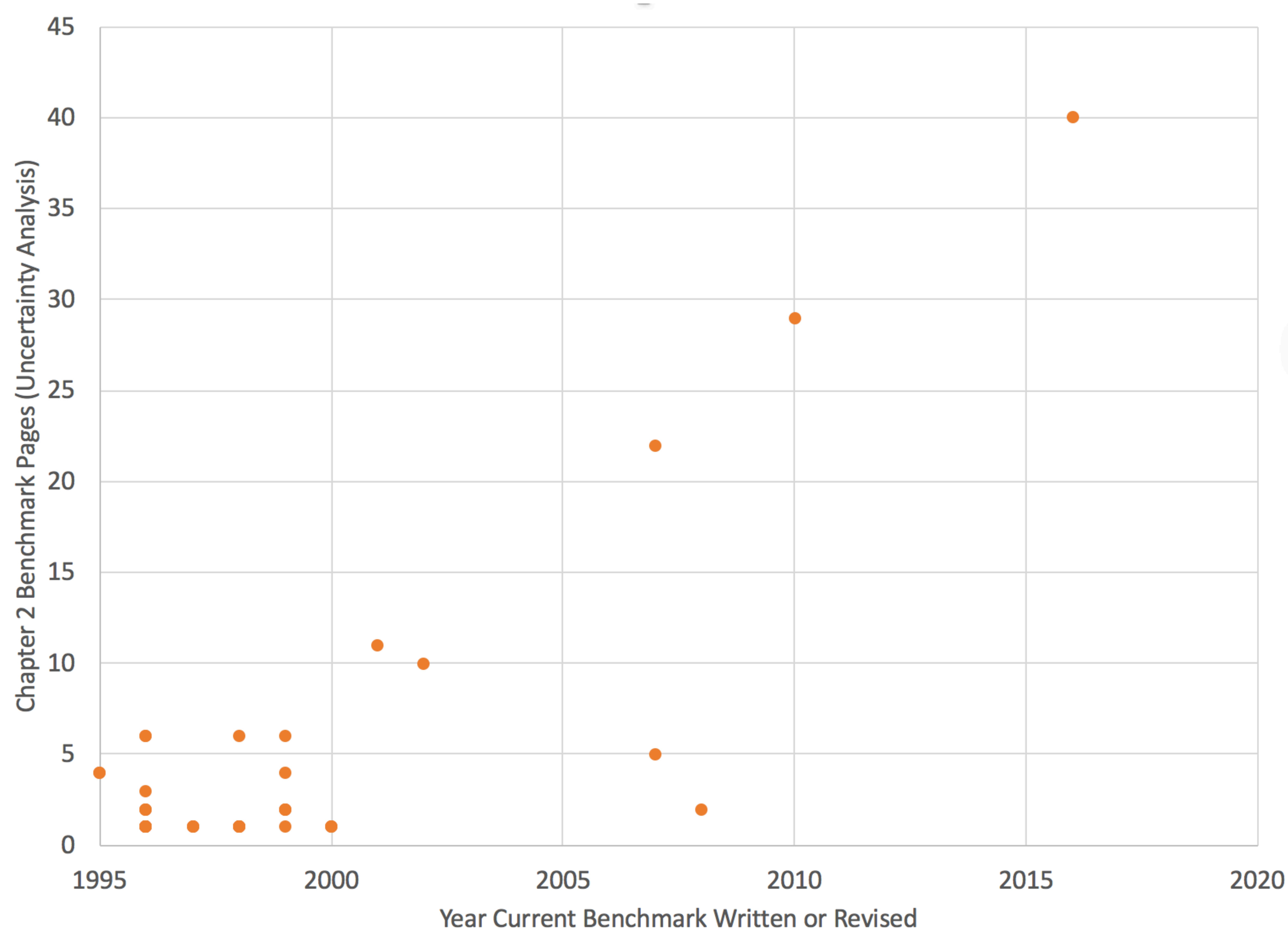
- Criticality safety validation driving force behind evaluations
  - Most experiments evaluated decades later by non-experimentalists
- ICSBEP expectations have evolved over time
  - Earlier evaluated benchmarks tend to be more brief
  - Many evaluated benchmarks are missing major sources of uncertainties
  - Computer power was limited, more reliance on simplified geometries

Example: PU-MET-FAST-001 (Jezebel) Section	Revision 2 pages (2007)	Revision 4 pages (2016)	Increase
1 (Experimental Data)	6	33	x5
2 (Experiment and Uncertainty Evaluation)	< 1	40	x40
3 (Benchmark Model)	3	46	x15
4 (Sample Calculations)	1	8	x8
Appendix (Supporting Documentation)	5	46	x9

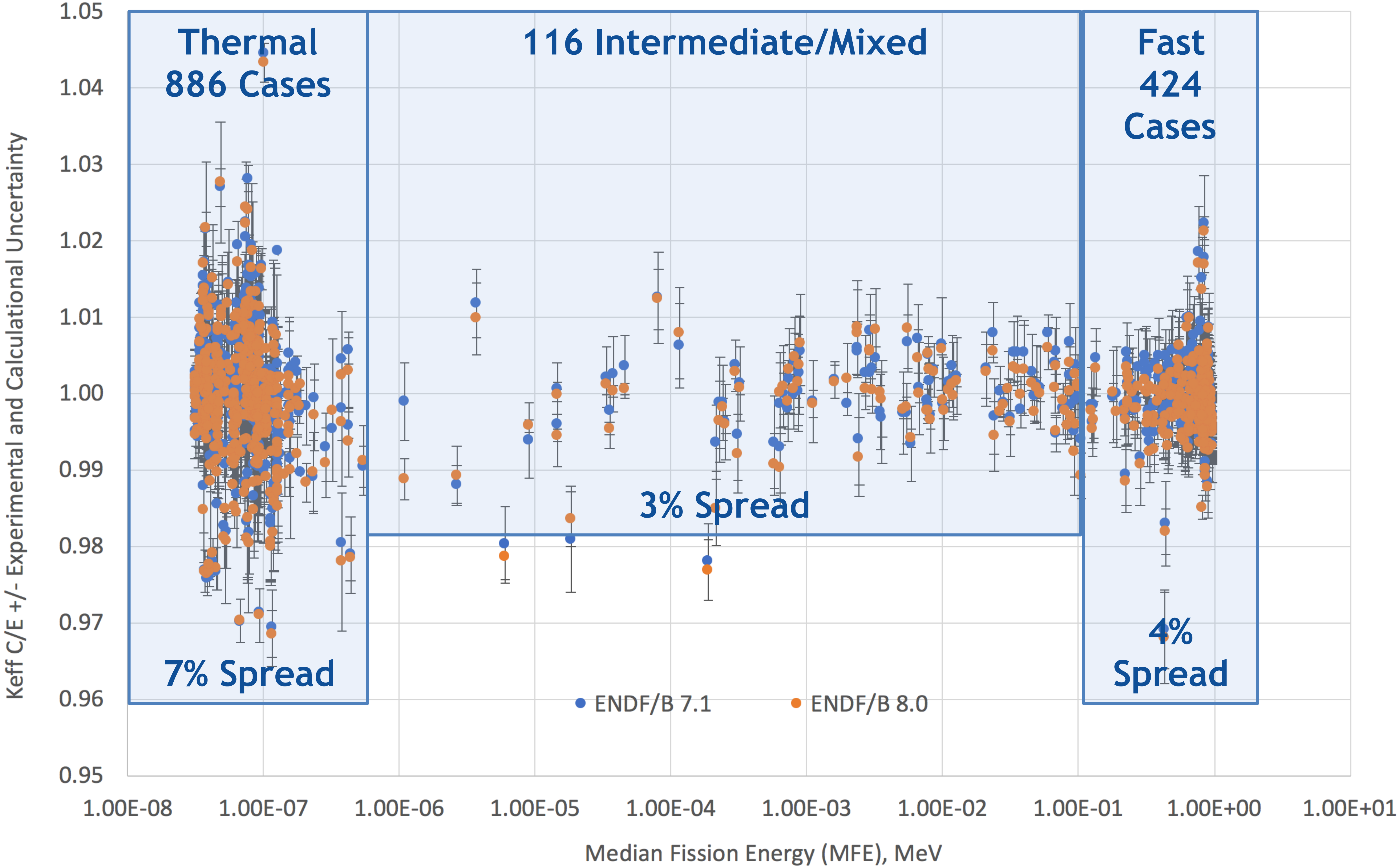
Total Page Count for Fast Pu Metal Cases over Time
(excluding appendices- sample inputs, etc)



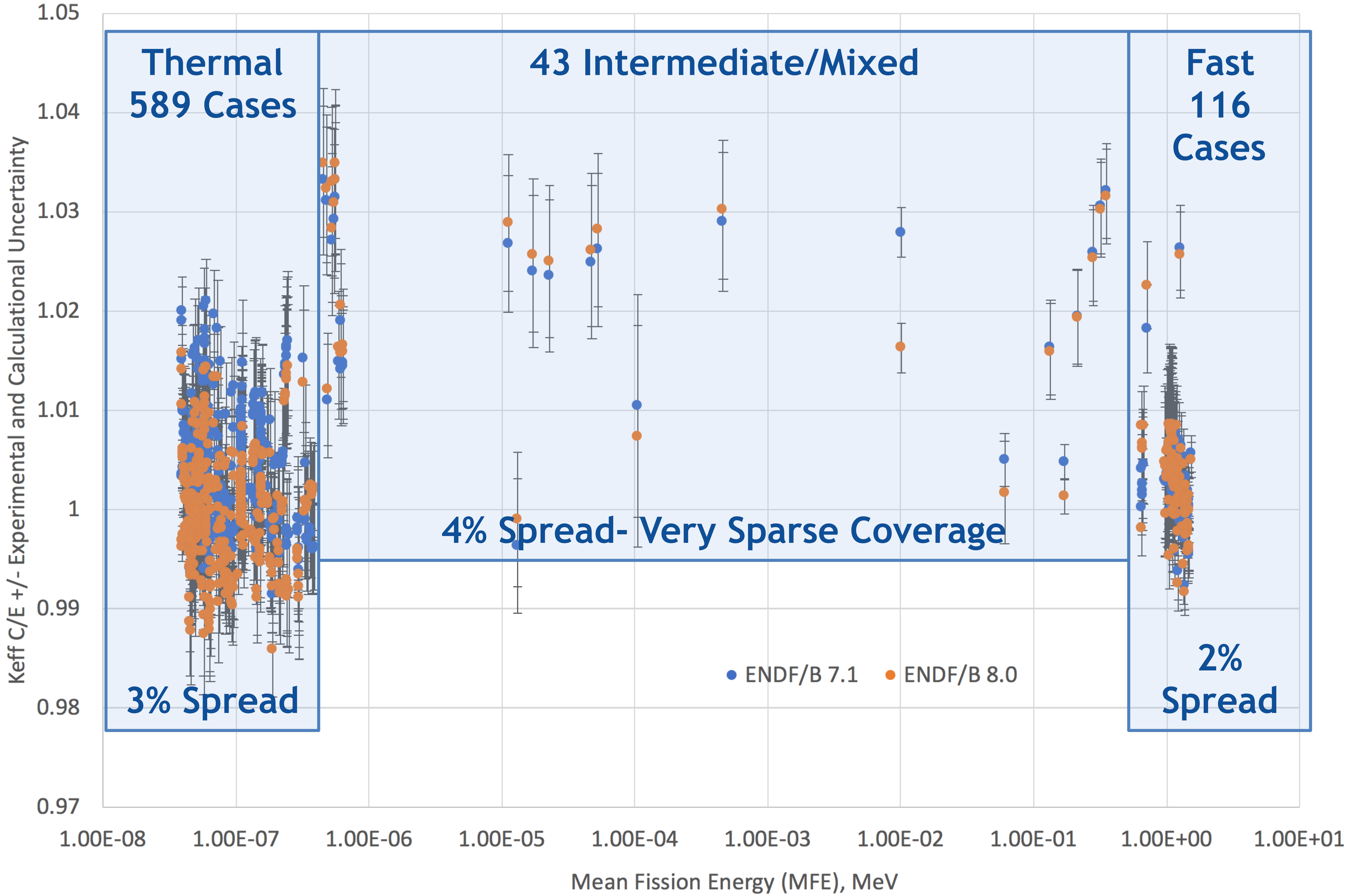
Uncertainty Analysis for Fast Pu Metal Cases over Time
(Length of Section 2)



ICSBEP HEU Benchmarks Overview



ICSBEP Pu Benchmarks Overview



Slide from C. Percher, CSEWG 2019

HMF-001, "Godiva"

- Benchmark is subcritical shell experiments completed to inform Lady Godiva design
- "Uncertainties" are only experimental- from extrapolation to idealized critical sphere from subcritical shells
  - Shell radii were not well known!
- Missing **MAJOR** Uncertainties:
  - Uranium Mass
  - Dimensions of shells
  - Uranium composition
  - 100 pcm uncertainty is likely not right

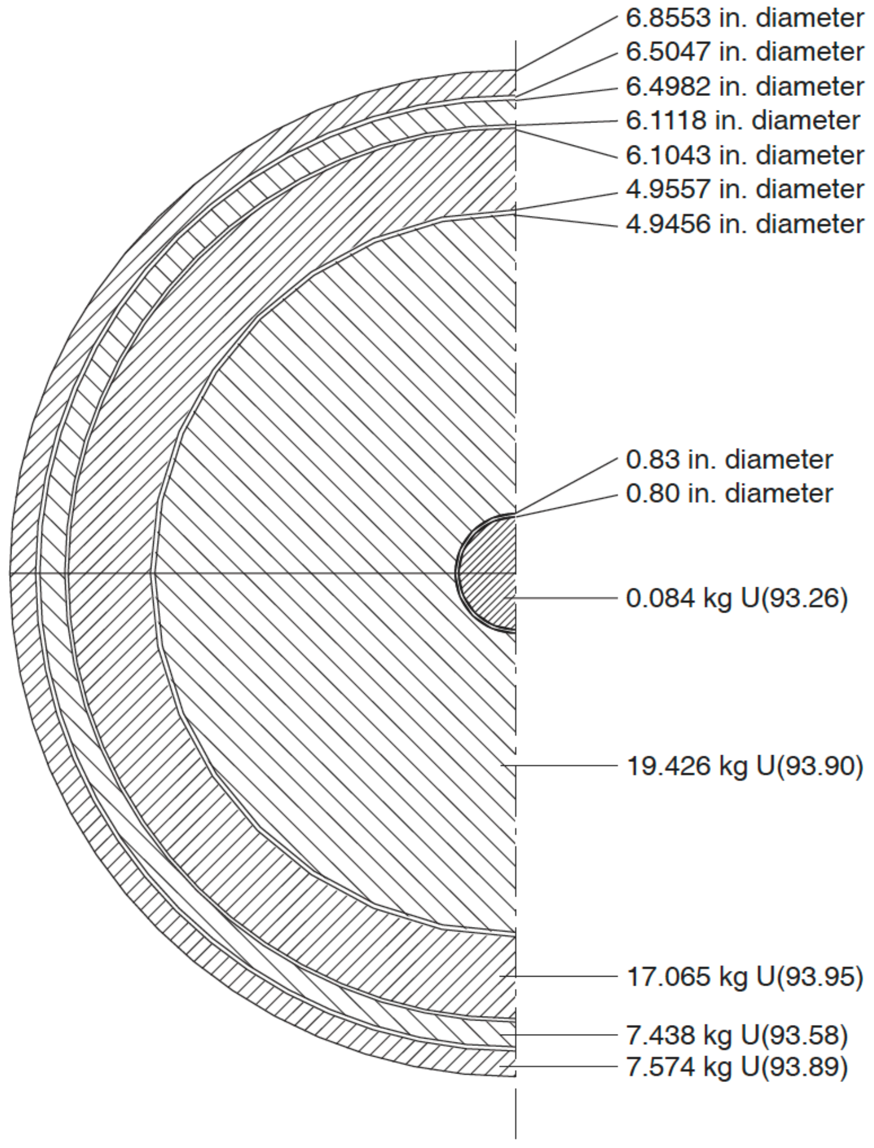
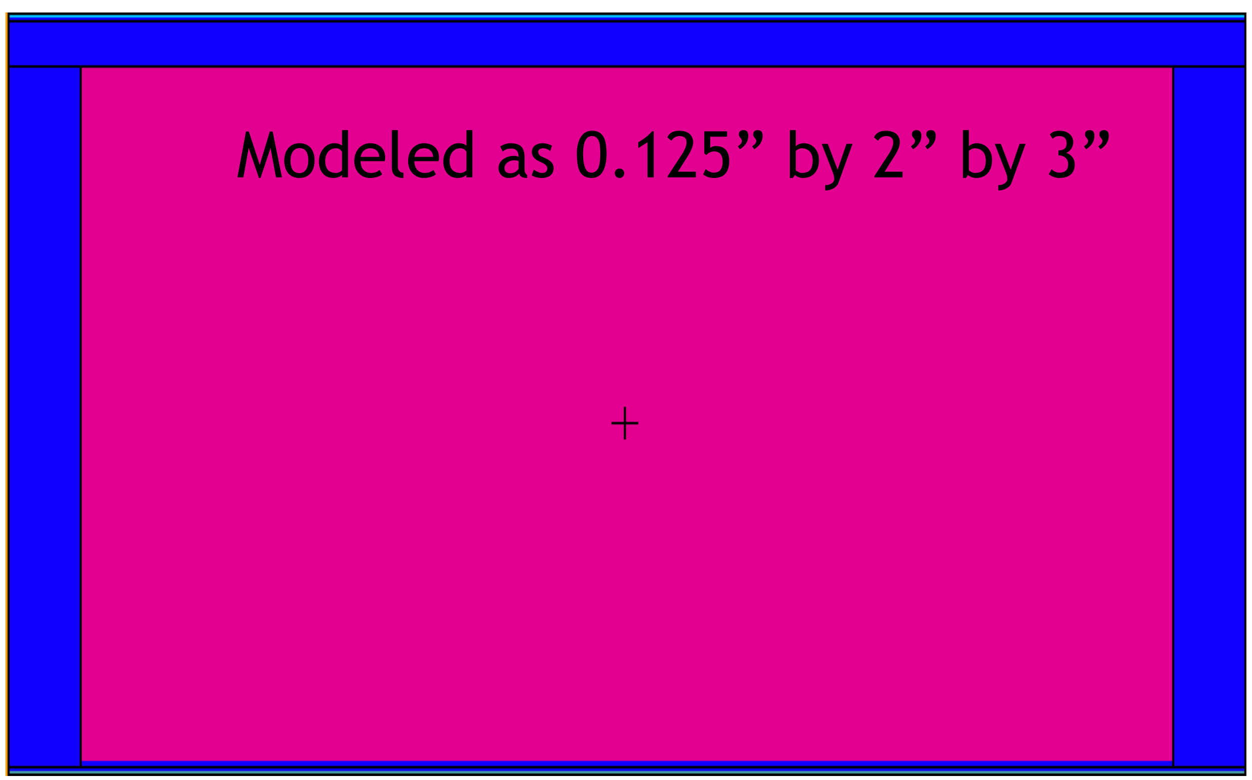


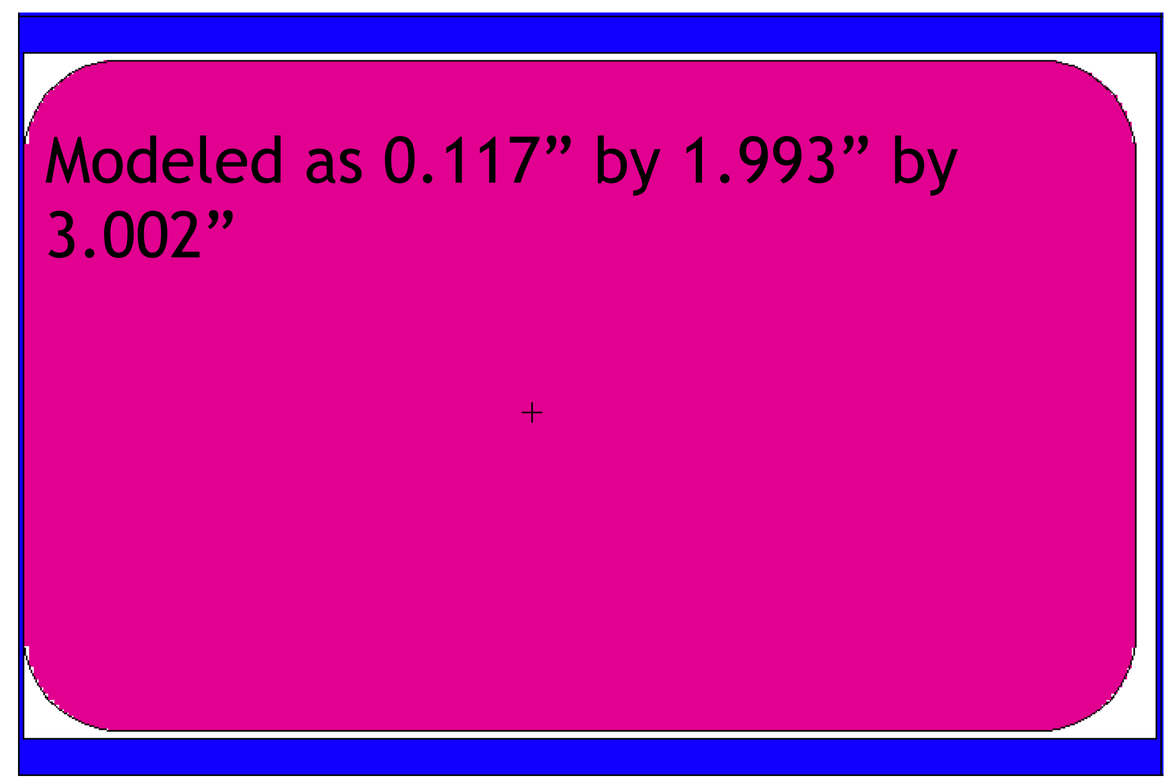
Figure 2. Idealized Final Configuration of Subcritical U(94) Spherical Shells.

ZPR/ZPPR Examples

- Many examples of ZPR/ZPPR benchmarks, very complicated honeycomb drawer configurations comprising thousands of fissile, diluent, and reflector plates
  - Benchmarks use **nominal** plate sizes with simplified geometries (smeared densities, no streaming paths)
  - No analysis of dimensional uncertainties



PANN Plate From PU-MET-INTER-004



PANN Plate From PU-MET-MIX-002

Majority of Intermediate/Mixed Cases are Suspect

- Pu-Comp-Mixed-001 and -002, *Unreflected and Plexiglass-Reflected Slabs of Polystyrene-Moderated Plutonium Oxide*
  - 34 cases, 1960's experiments evaluated in 1999
  - Only loose Pu oxide benchmarks- considered important for criticality safety
- These benchmarks calculate very poorly (C/E values of 1.02-1.04)
- Likely due to unquantified uncertainties- tape around boxes, compact densities, heterogeneity, thermal expansion, loss of hydrogen due to radiolysis

Table 2. Plutonium-Oxide Polystyrene Cube Parameters.

Cases	H/Pu	Bare Dimensions (X - Y - Z, cm)	Contamination Control	Clad Dimensions (cm)
1-5	0.04	5.13 x 5.13 x 3.81 (±0.02)	2 layers of shrink wrap	5.54 x 5.54 x 4.09 (±0.04)
6-9	5	5.12 x 5.12 x 3.81	Tape cladding	5.21 x 5.16 x 3.88
10-16	15	5.116 x 5.116 x 5.116[(a)] (±0.03)	Aluminum paint & rubberized plastic	5.171 x 5.171 x 5.171 (±0.03)
17-22	15	5.080 x 5.080 x 5.080[(a)] (±0.025)	Aluminum paint & rubberized plastic	5.121 x 5.121 x 5.121 (±0.025)
23-29	49.6	5.09 x 5.09 x 5.09 (±0.01)	MM&M # 471 Tape	5.19 x 5.19 x 5.16 (±0.04)

(a) These dimensions include a ~1-mil (0.0025-cm) coat of aluminum paint.

What Does This Mean for Nuclear Data Users?

- **Use caution when relying on a benchmark to inform nuclear data-** Read the evaluation and use your judgement

- **New OECD Working Party for Nuclear Criticality Safety (WPNCS) Subgroup 8-** Preservation of Expert Knowledge and Judgement Applied to Criticality Benchmarks
  - New Subgroup approved in Sept 2019
  - Capture historical and tribal knowledge of benchmark issues
  - Ultimate goal to “grade” ICSBEP benchmarks, similar to past efforts related to differential measurements
  - Will Wieselquist (ORNL), Chair
  - Will help identify candidates for re-evaluation

- **Prioritize and complete new, modern experiments that can undergo full uncertainty and correlation assessment**

Barriers to putting it all together

NRDC

ENDF B-VIII.0

NEA

RSICC

Application to U-233

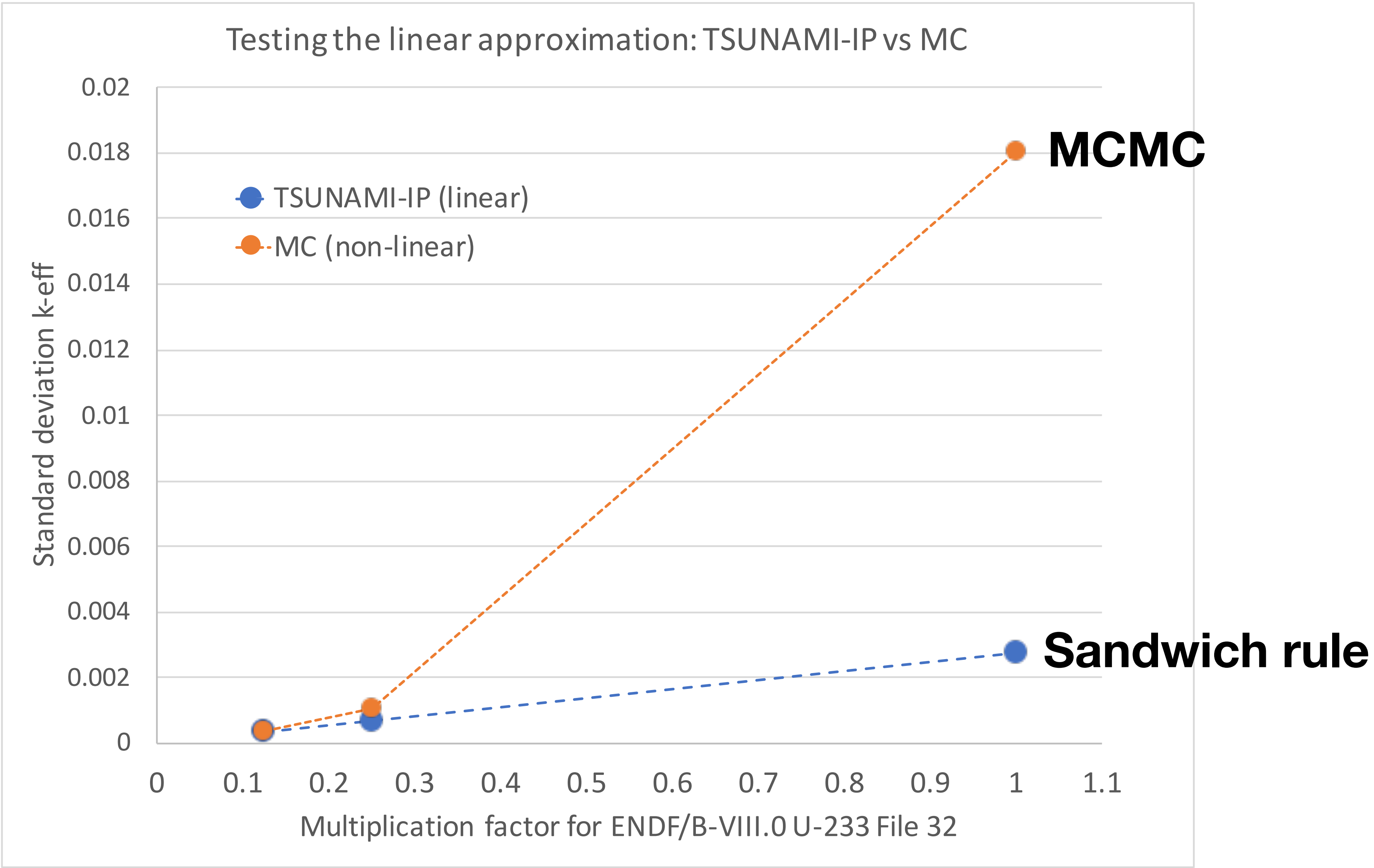
- 1000 randomly perturbed resonance parameter sets created by sampling File 32
- For each set calculate k_{eff} for U233-SOL-{THERM, INTER}-001-001 (KENO code), and then calculate k_{eff} mean values and uncertainties,
  - compare to corresponding TSUNAMI-IP's
  - Compare to measured IBE data
- For each set calculate differential cross sections using the SAMMY and then calculate mean values and uncertainties (transmission, fission)
  - Compare to SAMMY File 32 calculation, assuming it can be done
  - Compare to differential data (transmission, total, fission) by K. Guber (ORNL)

OAK RIDGE National Laboratory

Slide from G. Arbanas, CSEWG 2019

MC *vs.* linear approx. for Δk_{eff} of U233-SOL-INTER-001-001

- MC reveals large deviation from non-linearity for ENDF/B-VIII.0 U-233 File 32



Slide from G. Arbanas, CSEWG 2019

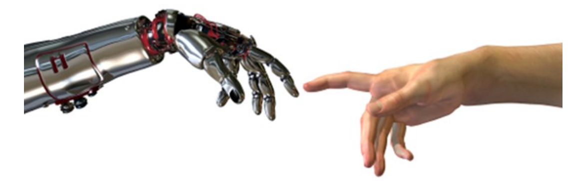
Take away message from this & related studies

- We are forced to use Total Monte Carlo for large parts of calculation

- However, CEA Bruyères-le-Châtel & PSI demonstrated that full variation feasible

We need to work up resource requirements for full network

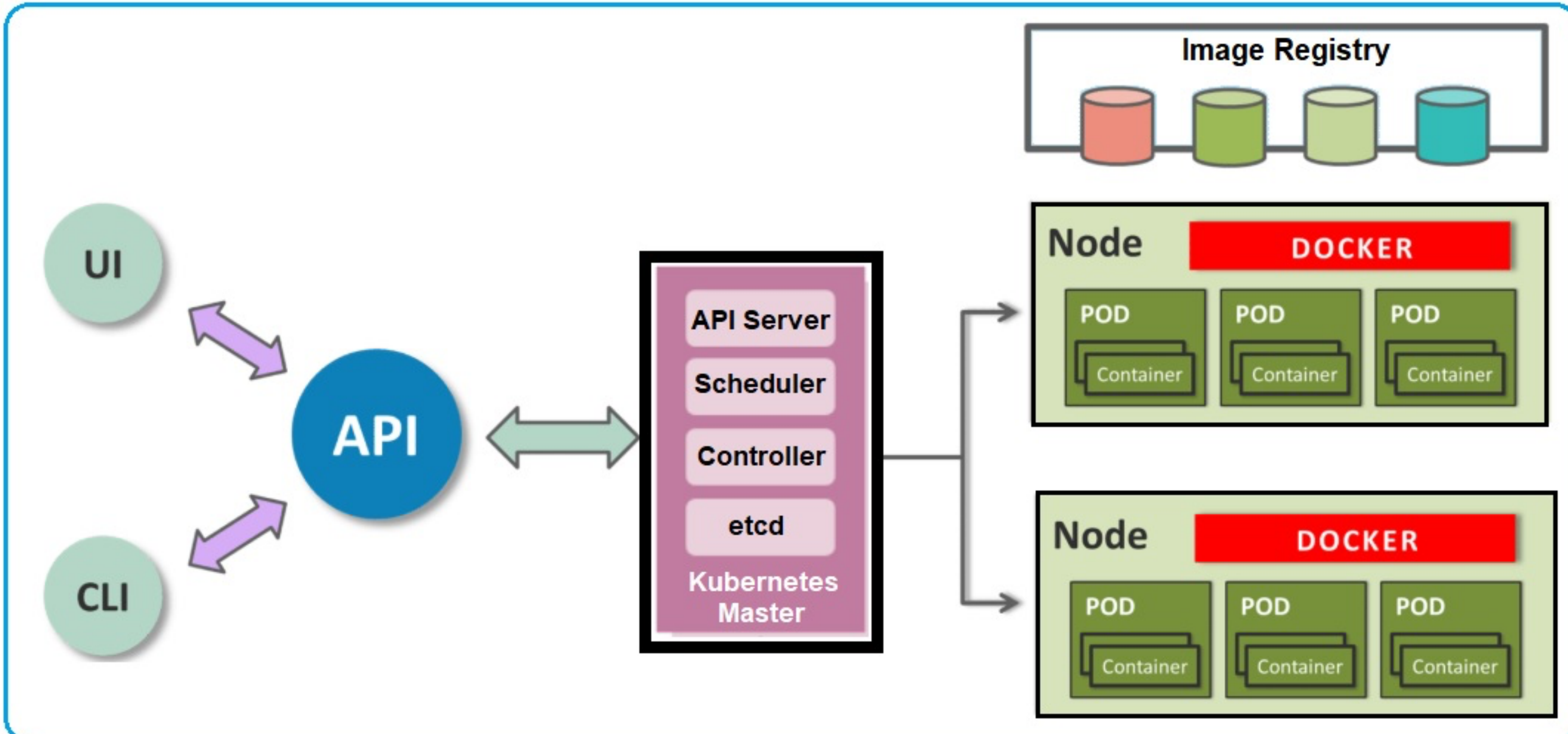
There are many things keeping us from the SG-49 vision



- **EXFOR issues**:
  - EXFOR incomplete scope
  - Incomplete information in entry
  - EXFOR not machine readable enough (ignoring EXFOR pointers even!)
- **ENDF issues**:
  - RRR evaluations “not automateable”
  - Model defect in reaction codes
  - ENDF not a fully formed GPR
  - G in ENDF GPR problematic in some cases (thresholds, fission)
  - Imperfect fundamental data (RIPL, esp. levels)
- **Issues with our Tools:**
  - Export control issues
  - Freemium model + dealing with changing terms and conditions
  - Vendor lock
- **Cyber security issues:**
  - Docker & Kubernetes
  - Remote execution
- **ICSBEP issues:**
  - Bad benchmarks
  - Incomplete model specification
  - “redundant” tests (neutronically similar)
- **Issues putting it all together:**
  - Not enough CPUs?
  - Overly simplified network
  - Can’t vary everything

Extra slides

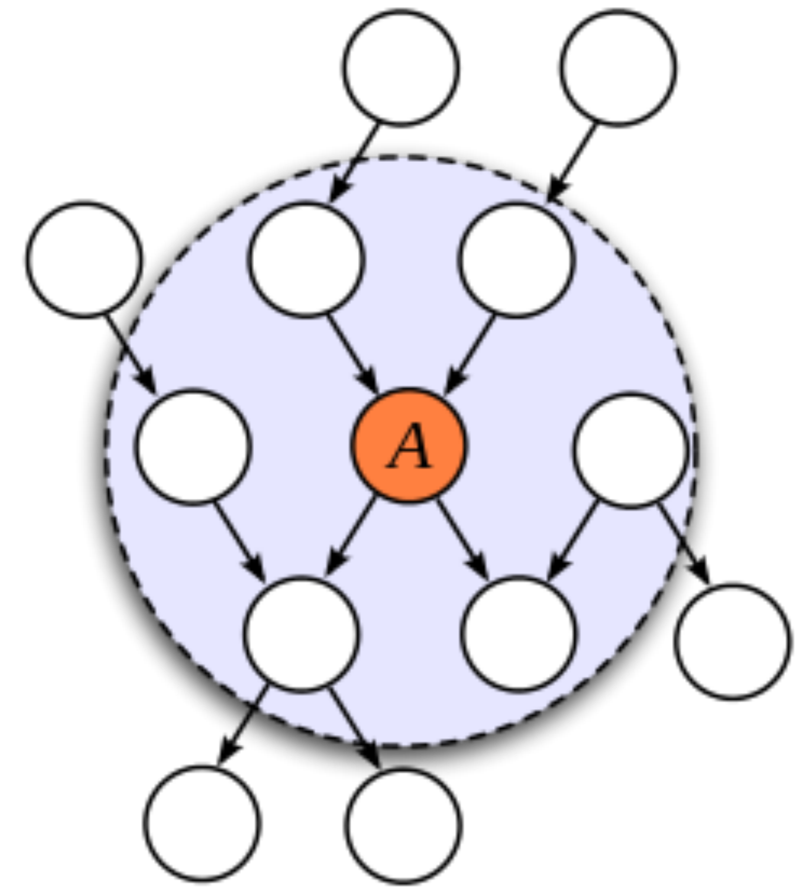
have something magical involving containerized applications happen



An observation: we don't need to eat the whole Bayesian network at once

- ***Markov blanket*** for a node are all the variables that shield node from rest of the network
- A Markov blanket of a node is the only knowledge needed to predict the behavior of that node and its children
- Mathematically:

P(A | MB(A), B) = P(A | MB(A))



Translation: We need cross-covariances,
but not all of the cross-covariances,
the network takes care of the rest

ICSBEP- September 2019 Edition

- Even with 748 Plutonium configurations, only fast and thermal regimes are well represented
  – 589 are thermal plutonium solution configurations, 530 cases are >80% thermal
  – 116 fast metal cases, 82 are >80% fast
  – Only 35 have intermediate fission fractions > 30%
    - **4 modern BFS configurations (Russia, IPPE)**
    - 1 k_{inf} measurement (UK, small sample reactivity measurement)
    - 3 Zero Power Reactor (ZPR) configurations (USA, Argonne Reactor Mock-ups)
    - 27 are plutonium oxide polystyrene compacts (USA, Hanford Poly Block experiments)