



# Data preservation at the National Nuclear Data Center

David Brown

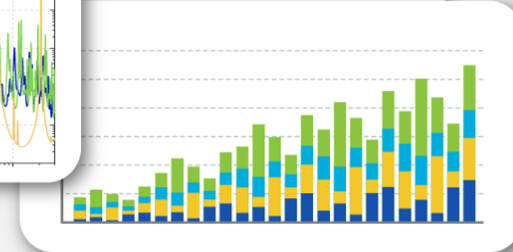
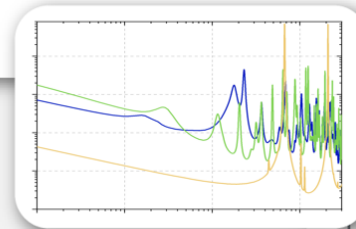
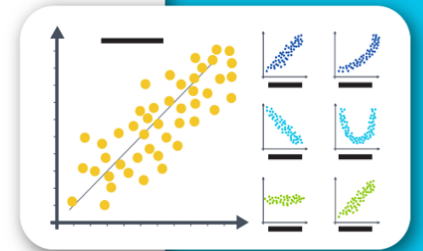
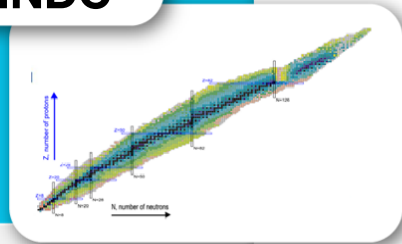
4 June, 2026



# NATIONAL NUCLEAR DATA CENTER

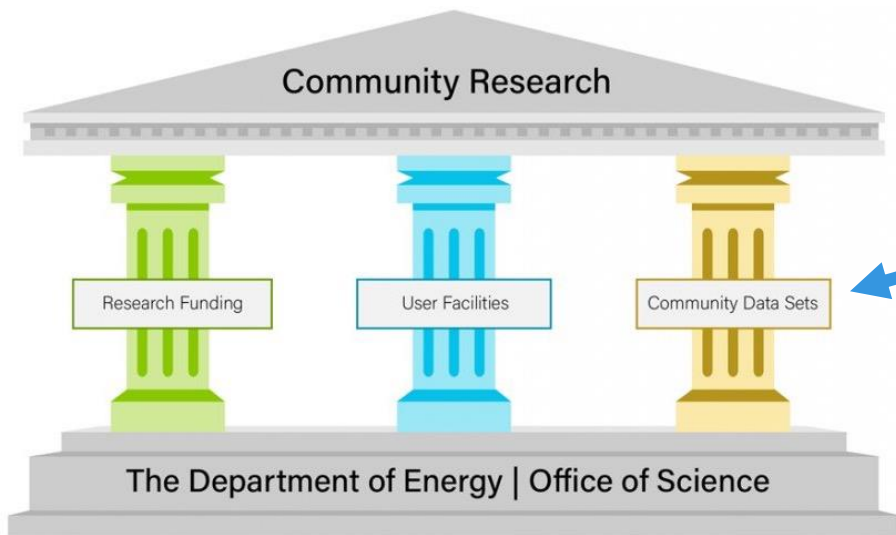


Current, accurate, authoritative  
data in areas of nuclear  
science and engineering



Office of Science designated PuRe Data Resource | open data | data repositories, knowledgebases, analysis platforms

# DOE PuRe Data Resources

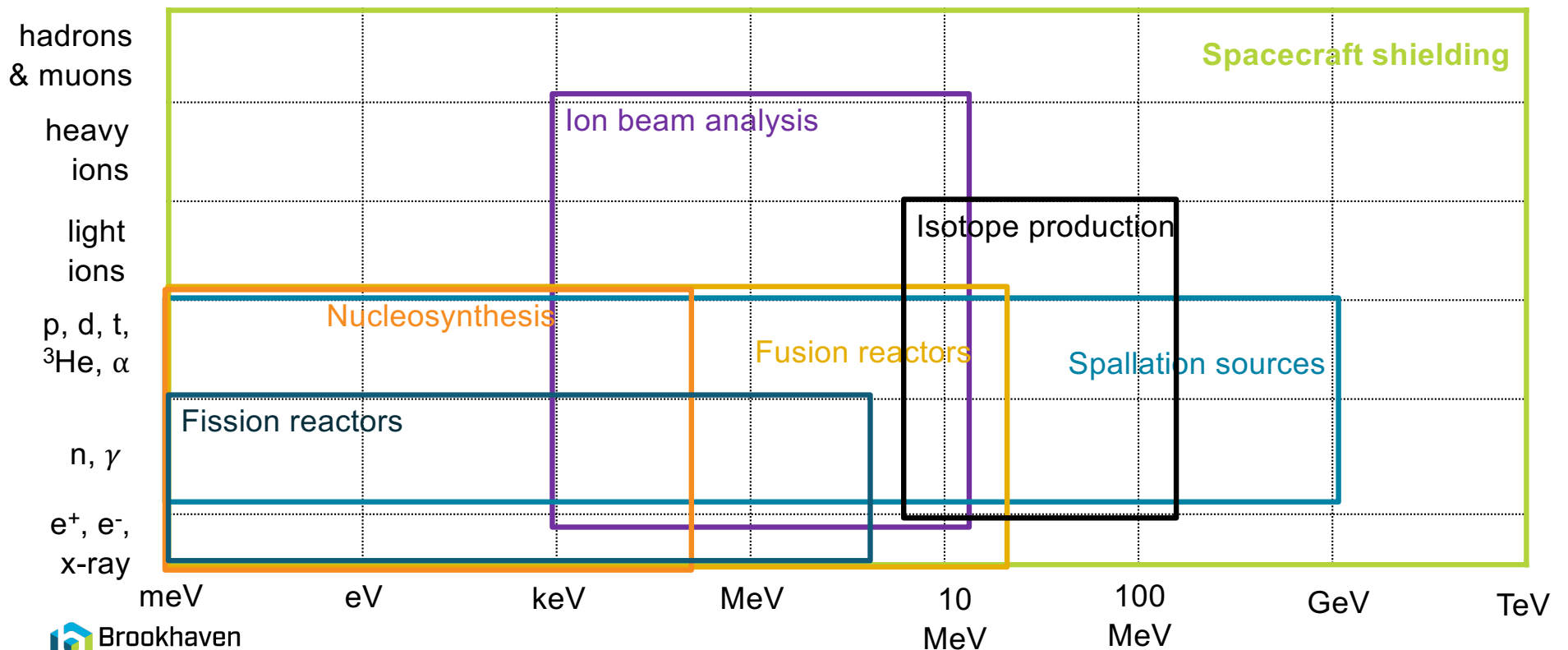


Public Reusable Research (PuRe) Data Resources aim to make data publicly available in order to advance scientific or technical knowledge

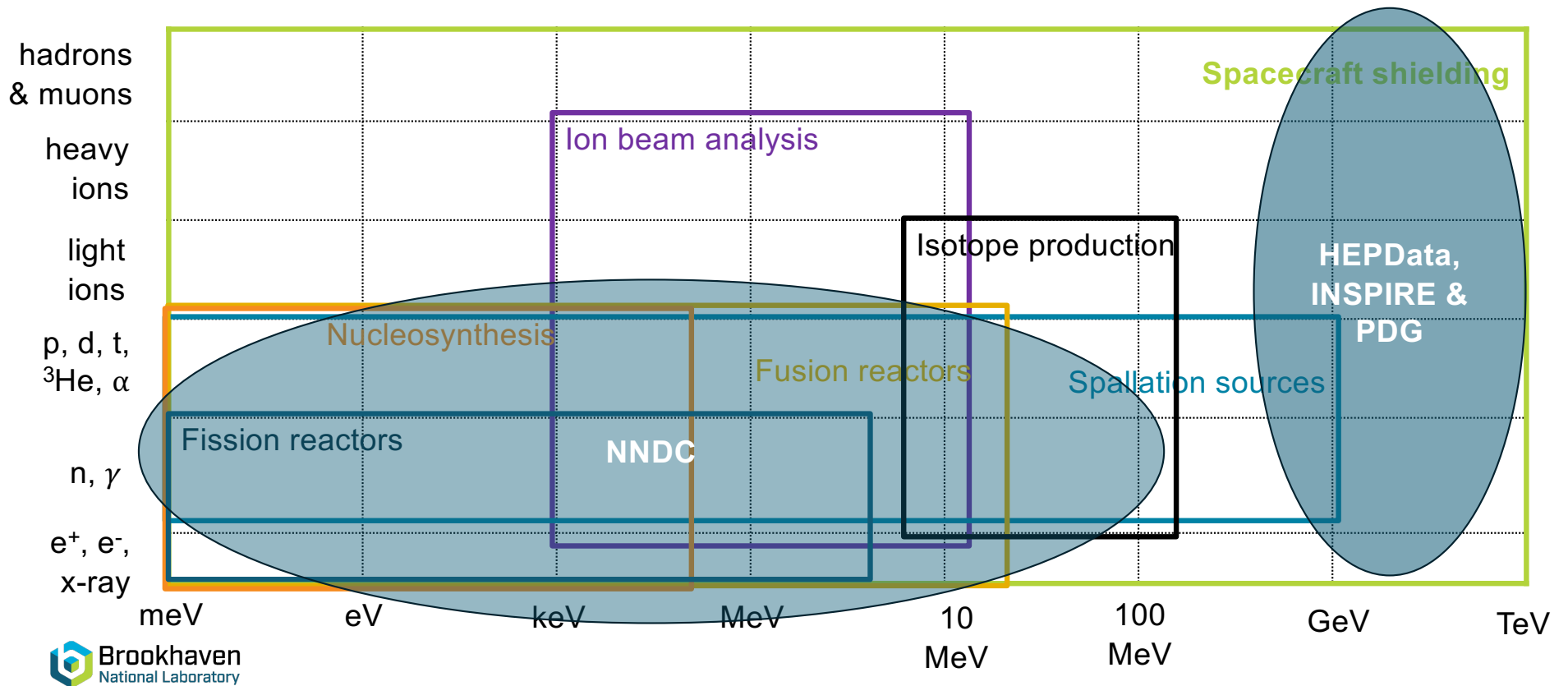
<https://science.osti.gov/Initiatives/PuRe-Data>














# Landscape of things that hit nuclei



# Landscape of things that hit nuclei



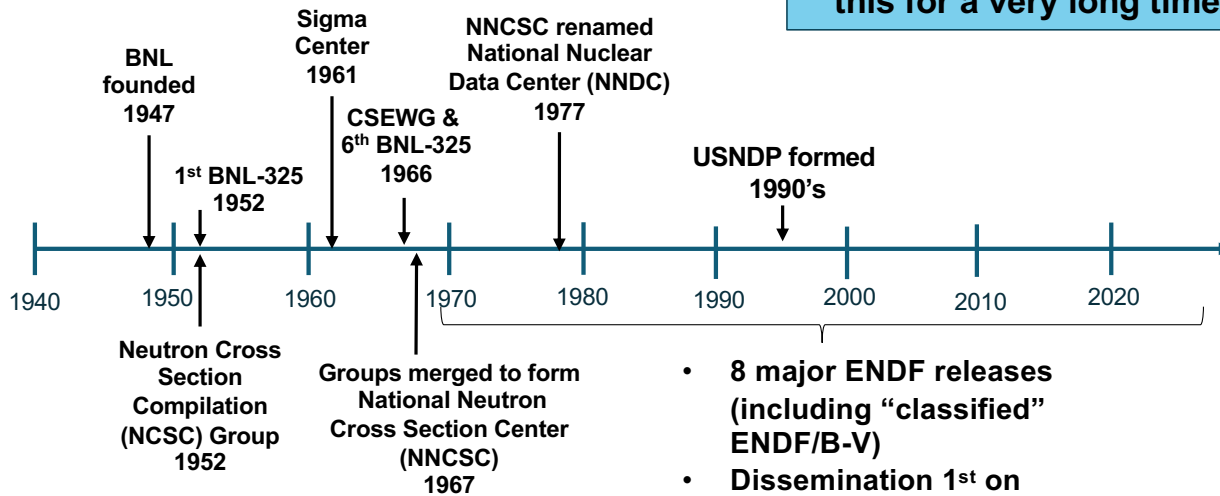
# Coverage of long-term data projects

	Bibliographic	Experiment compilation	Particle properties	Evaluated reaction	Reaction theory
Low Energy (<20 MeV/A)	NSR 	EXFOR 	ENSDF 	ENDF 	NNDC & ND community
Low-Medium (20-150 MeV/A)	NSR 	EXFOR* 		ENDF* 	LE community
High-Medium (150 MeV/A-1 GeV/A)	INSPIRE 	?**	PDG 	?	Jlab & ME community
High Energy	INSPIRE 	HEPData 		?	BNL & HE community

# NNDC details

# The NNDC is the lead unit of the US Nuclear Data Program

The NNDC has been doing this for a very long time!



<https://www.nndc.bnl.gov/about/>  
<https://www.nndc.bnl.gov/history/>

<https://www.nndc.bnl.gov/>

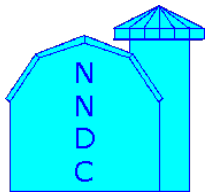
- 8 major ENDF releases (including “classified” ENDF/B-V)
- Dissemination 1<sup>st</sup> on ARPANET & later internet
- BNL-325 evolves into *Atlas of Neutron Resonance*
- 202 volumes of *Nuclear Data Sheets* published
- Many many barn logo changes



## Branding 101!

- ❑ The Greek letter sigma,  $\sigma$ , is used as a symbol for nuclear cross sections, which measure the probability that a neutron interacts with a given nucleus.
- ❑ The group founded by Donald Hughes was known as the **Sigma Center!**
- ❑ The 'barn' unit, **1 barn=10<sup>-24</sup> cm<sup>2</sup>**, is used for nuclear cross sections.
- ❑ The **barn** has been used as a symbol of the group from the very beginning!

Element	Isotope (%, T <sub>1/2</sub> )	Reaction Cross Sections (2200 m/s)	
		$\sigma_{\text{abs}}$	$\sigma_{\text{act}}$
<sup>1</sup> H	H <sup>1</sup> (~100)	330 ± 3 mb	
	H <sup>2</sup> (0.015)	0.46 ± 0.10 mb	12.4y 0.57 ± 0.01 mb



# Funding profile

## NNDC:

- ~ \$8M total
- ~ \$4.2M USNDP (DOE NP)
- ~ \$3.8M various NNSA (mainly DNN R&D, NCSP)
- Occasional funding from other programs (FES, NE, etc.)

# Resources

- 10 Scientific staff
  - 3 ENDF
  - 1 NSR/EXFOR,
  - 1 XUNDL,
  - 5 ENSDF
- 3 Professional staff (web development & cluster, gitlab support)
- 1 Administrator

- Gitlab instance
  - [git.nndc.bnl.gov](https://git.nndc.bnl.gov)
  - Kubernetes cluster for CI/CD
- Modest cluster
  - Can run Export Controlled software
  - 28 cores, 384GB memory, 14 TB disk, \$75k
  - Want to add GPUs
- website



# Data management plan

## Robust cloud backup since Sep. 2021

One of a kind at BNL

Backup of up to 15 TB of data from on-premise, mission-critical servers to AWS GovCloud

Continuous replication

During a disaster:

1. Server backup restored from AWS GovCloud to create production-ready instance.
2. Accessible within 2 hours.

Annual cost: ~\$30k fully burdened



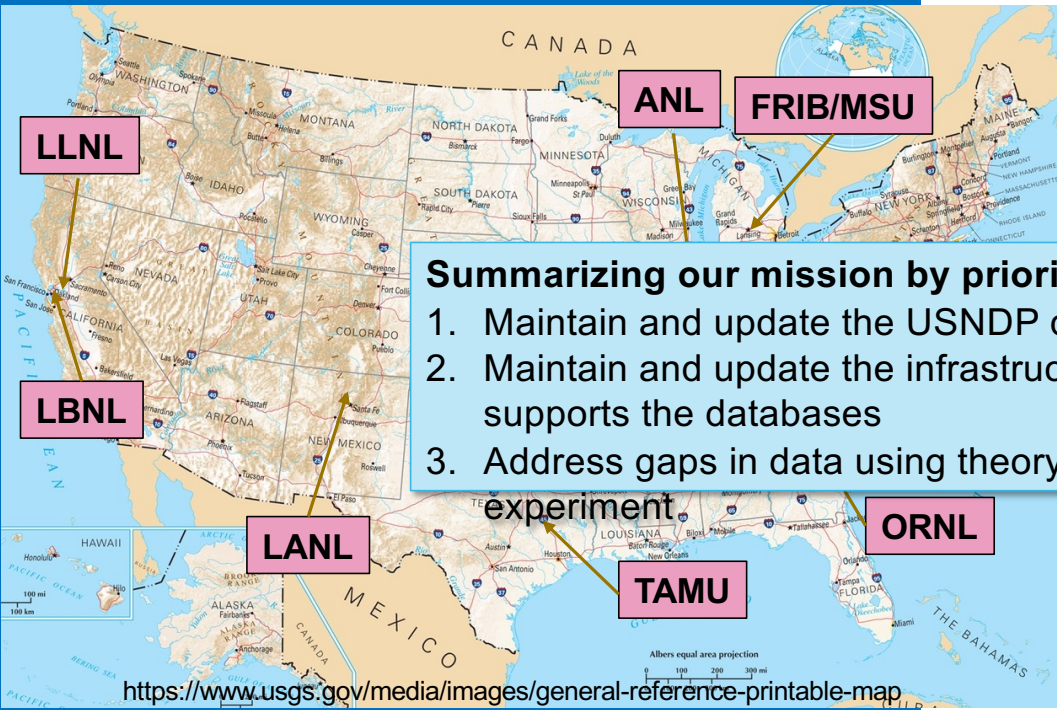
# USNDP

# US Nuclear Data Program

## Mission

The mission of the United States Nuclear Data Program (USNDP) is to provide consistent, accurate, **authoritative** nuclear data **workers in pure and applied nuclear science and** engineering. This is accomplished primarily through the compilation, evaluation, dissemination, and archiving of extensive nuclear datasets. USNDP also addresses gaps in the data, through targeted experimental studies and the use of theoretical models.

[www.nndc.bnl.gov/usndp](http://www.nndc.bnl.gov/usndp)



- Summarizing our mission by priority:**
1. Maintain and update the USNDP databases
  2. Maintain and update the infrastructure that supports the databases
  3. Address gaps in data using theory and experiment

The NNDC hosts most nuclear data needed for applications

### **Nuclear Science References (NSR)**

Nuclear physics articles indexed according to content

### **EXFOR**

Compiled nuclear reaction data

### **XUNDL**

Compiled nuclear structure and decay data

### **ENSDF**

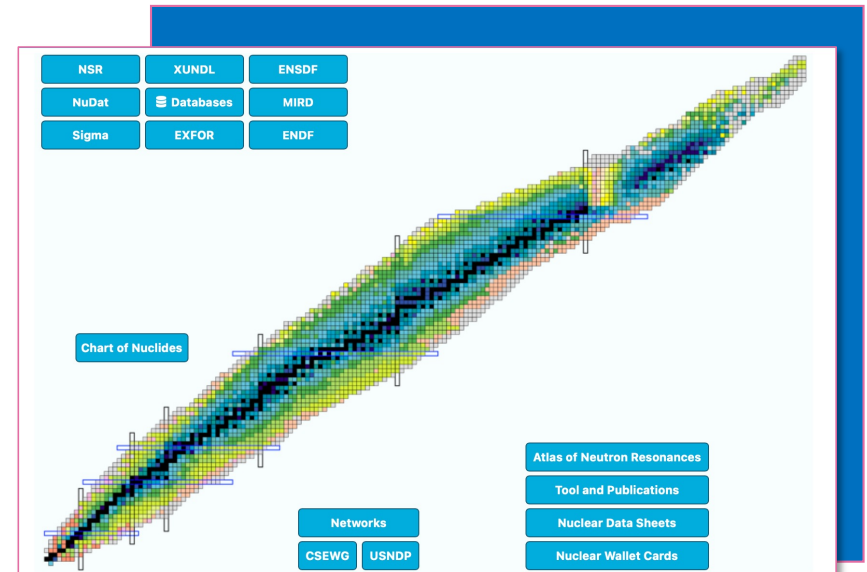
Recommended nuclear structure and decay data

### **ENDF**

Recommended particle transport and decay data, with a strong emphasis on neutron-induced reaction data

### **Nuclear Data Sheets**

Journal devoted to the publication of nuclear data articles



[www.nndc.bnl.gov](http://www.nndc.bnl.gov)

**Nuclear data science capability  
to support the development of  
new reactor concepts.**

# Each library is more than a zipfile

- Each library is essential for someone
- Governance framework for each data product
- Simple ASCII (JSON, XML, etc.) data formats for maximal portability with well defined schema
- Evolved (evolving?) workflow
- A community/village/ecosystem
  - To fund (users tell the funders)
  - To do the work
  - To use (give us a reason!)
- Partners

**To me, this is the  
most important part**

# Our Partners

- **Nuclear Data Working Group/Nuclear Data Interagency Working Group**

- NDWG holds annual WANDA workshop (with whitepaper)
- NDIAWG coordinates funding based in part on WANDA outcomes

- **Nuclear Energy Agency (OECD)**

- **Working Party on Evaluation Cooperation (WPEC)** – 3 year projects improving evaluations
- **Working Party on Nuclear Criticality Safety (WPNCSS)** – makes ICSBEP handbook
- Other working parties

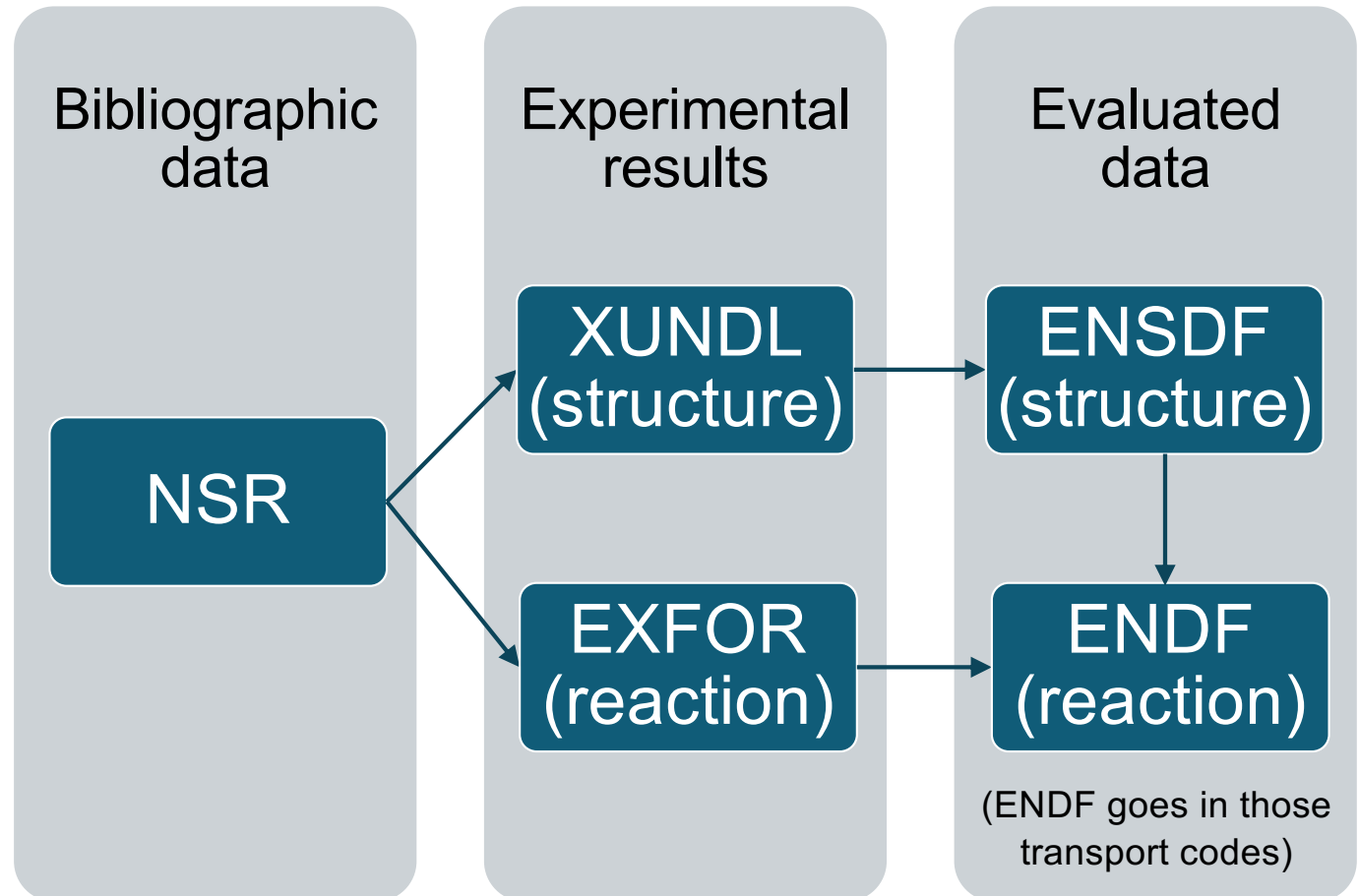
- **IAEA Nuclear Data Section**

- **Nuclear Reaction Data Centres (NRDC)** network - EXFOR
- **Nuclear Structure and Decay Data (NSDD)** network – ENSDF, XUNDL
- **International Nuclear Data Evaluation Network (INDEN)** – subset of high quality ENDF
- Various smaller 3-5 year projects



NRDC  
centers

The NNDC hosts 5 core libraries. Each plays a role supporting the ultimate mission.



# Data Preservation: EXFOR



## Nuclear Reaction Data Compilations in USA & Worldwide

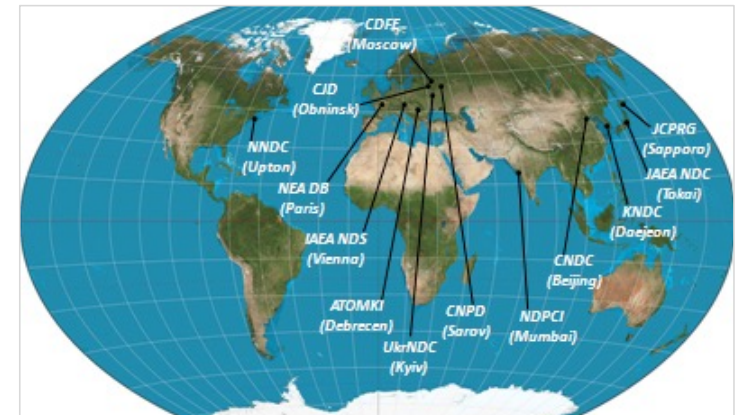
- Experimental neutron reaction data compilations have been pioneered at the Metallurgical Laboratory, University of Chicago and Los Alamos National Laboratory in 1945-1947.
- Brookhaven National Laboratory hired many *Manhattan Project* alumni when it was founded in 1947, and the lab got involved in nuclear data.
- Donald J. Hughes (1915-1960) was behind the BNL-170 (1952); it is a precursor of BNL-325 (Atlas of Neutron Resonances).
- SCISRS (Sigma Center Information and Retrieval System) at BNL (1964) was a precursor of EXFOR.
- Other data centers were created in Paris, France (NEA-Databank), Vienna, Austria (NDS-IAEA), and Obninsk, USSR (IPPE) in 1963-1964.
- Around 1970 four neutron data centers agreed on the data interchange format (EXFOR). The four centers could store data locally in its formats. The Nuclear Data Centres Reaction (NRDC) network was founded in 1979 under the auspices of the IAEA.

## 75 Years of Experimental Nuclear Reaction Data Compilations

# EXFOR - Experimental Nuclear Reaction Data

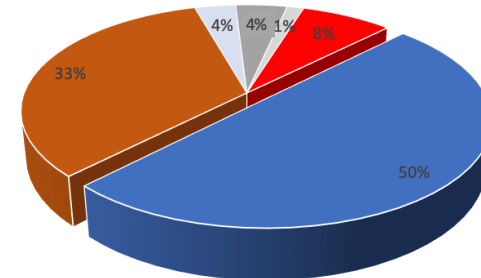
The largest experimental nuclear reaction database ([www.nndc.bnl.gov/exfor](http://www.nndc.bnl.gov/exfor), or [www-nds.iaea.org/exfor](http://www-nds.iaea.org/exfor))

- 25,226 experiments (multiple publications are grouped into one) as of February 4, 2025.
- Two largest contributors: NNDC & NEA-Databank.
- Cost of one experiment >\$1M => Database >\$25B
- EXFOR is coordinated by the IAEA's Nuclear Reaction Data Centres (NRDC) internationally.
- *EXFOR's philosophy is to compile data as published (in consultation with authors) unless obvious errors are found.*



25,226 Experiments in EXFOR

EXFOR Data Point Contributions Worldwide



NDC 1 ■ NEA-DB 2 ■ Japan 2 ■ NDS 3 ■ NDS Area 3 ■ Russian Federation

# Data archeology & Office dumps

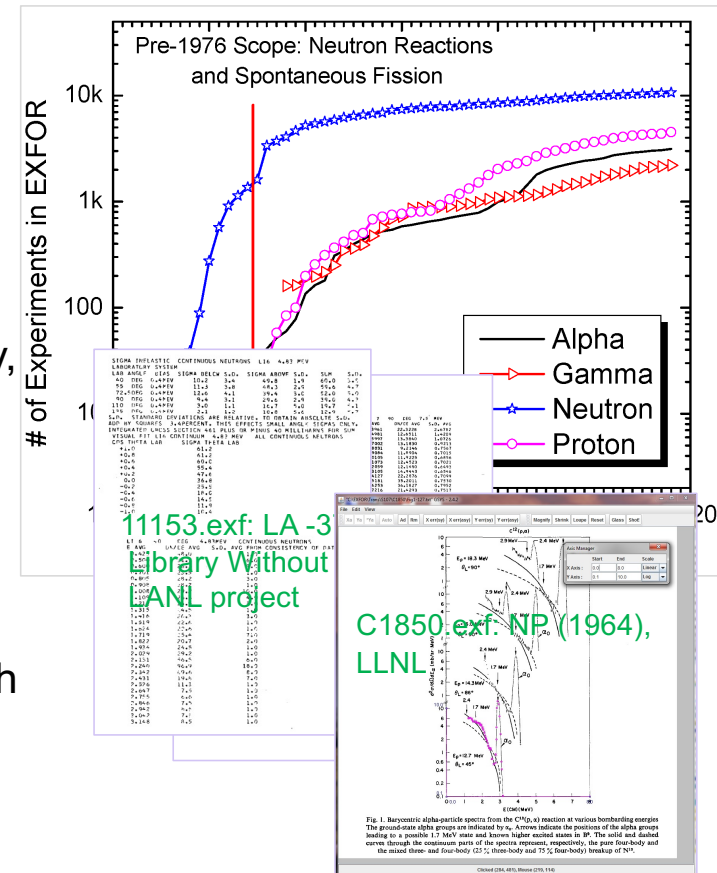
- This has happened
- This will continue to happen
- Arguably it is a failure of succession planning
- When it happens,
  - Saving the data takes time and resources
  - The data may be damaged
  - The data must be triaged
- But is clearly a tremendous opportunity



Mick Moxon, dec.  
May 12, 2020

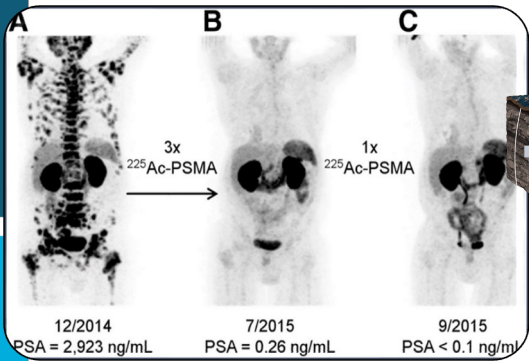
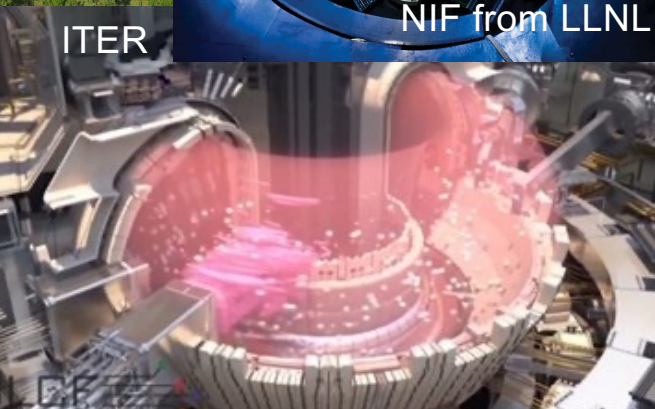
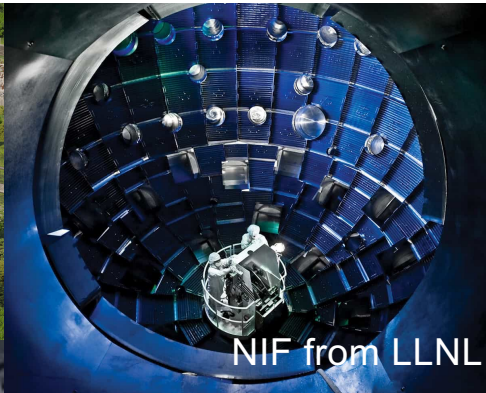
# Missing Legacy Data

- For historical and technical reasons, not all nuclear physics papers and data sets were compiled, or compilers did not obtain data for compiled entries (different compilation scope, no Internet and Email, poor computers and tools).
- EXFOR scope until 1976: only neutron reactions and spontaneous fission data (Plots for ND 2004).
- We work in EXFOR on data recovery (e.g. nuclear archeology, see the presentation of N. Otuka, and V. Dimitriou).
- Data recovery includes optical character recognition and data point digitization (Area #1: 11153, C1850). Both technologies work well with high-quality images and become challenging otherwise.
- Finally, EXFOR team members are willing to travel to research labs and universities and collect the original data to provide complete coverage of experimental nuclear reaction data.

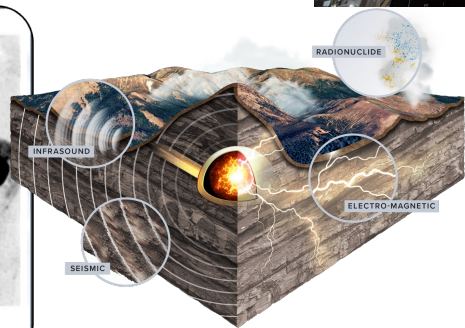


# Data Distillation: ENDF

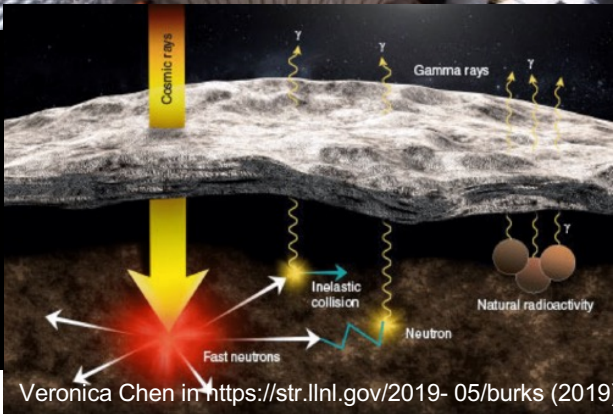
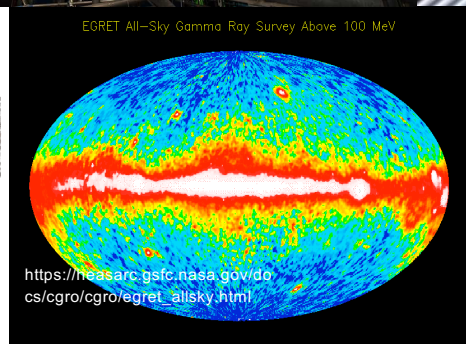
# Many different applications have nuclear data interests



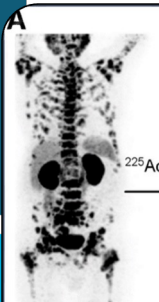
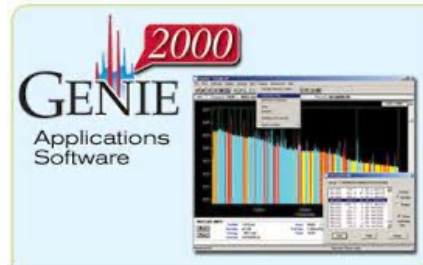
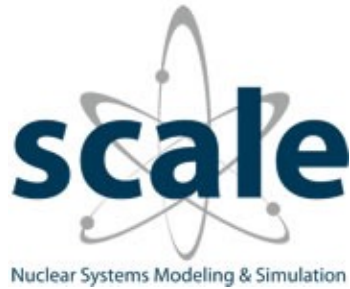
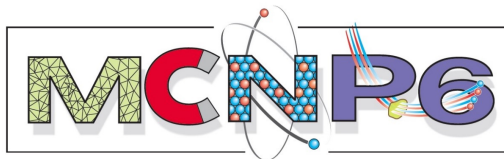
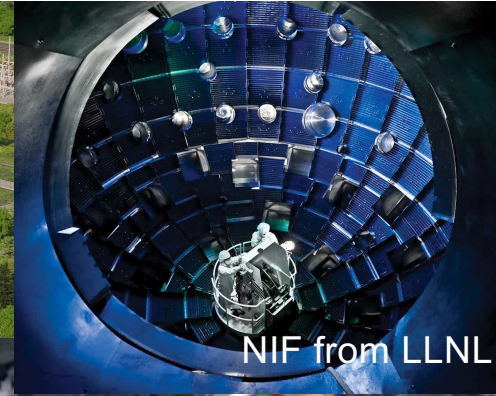
Ethan Balkin's WANDA 2024 talk



LYNM Program Overview, Chris Bradley LA-UR-21-24-967



# Many different applications have nuclear data



12/2014 PSA = 2,923 ng/mL  
 7/2015 PSA = 0.26 ng/mL  
 9/2015 PSA < 0.1 ng/mL

Ethan Balkin's WANDA 2024 talk

LYNM Program Overview,  
 Chris Bradley LA-UR-21-24-967

[https://heasarc.gsfc.nasa.gov/docs/cgro/cgro/egret\\_aisky.html](https://heasarc.gsfc.nasa.gov/docs/cgro/cgro/egret_aisky.html)



# The Cross Section Evaluation Working Group produces ENDF/B library



- **Formed 1966 & Chaired by BNL**
- **Currently ~200 members of the collaboration from 32 institutions**
  - US programs, industry and international partners
  - If you see something in the library, at some point a sponsor somewhere wanted it
- **All steps of nuclear data pipeline coordinated through CSEWG**



CSEWG meeting, BNL, Nov. 2024  
(I couldn't find a newer picture this morning)

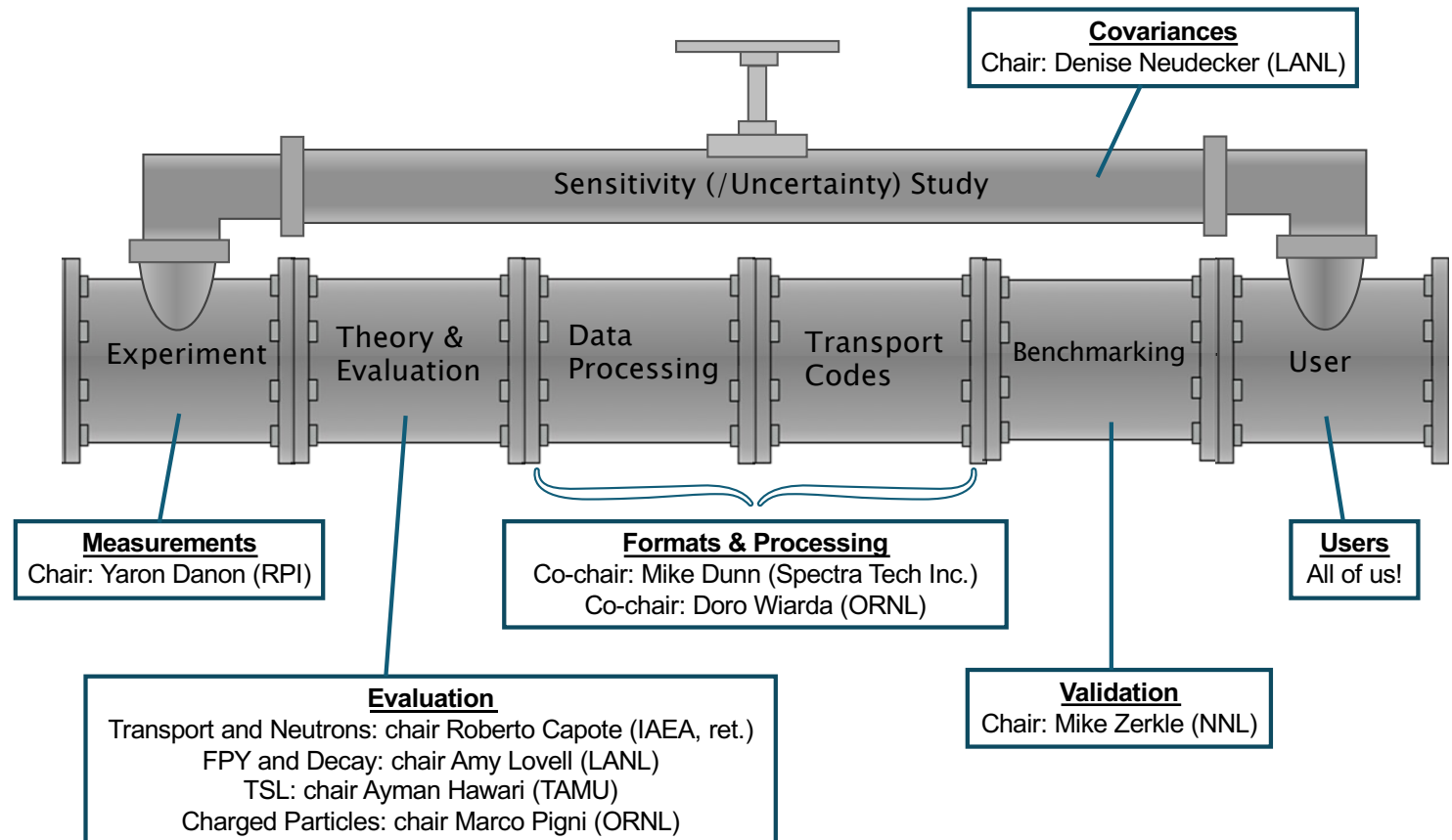
# CSEWG Executive Committee

## Chair:

David Brown (BNL)  
[dbrown@bnl.gov](mailto:dbrown@bnl.gov)

## Library Manager:

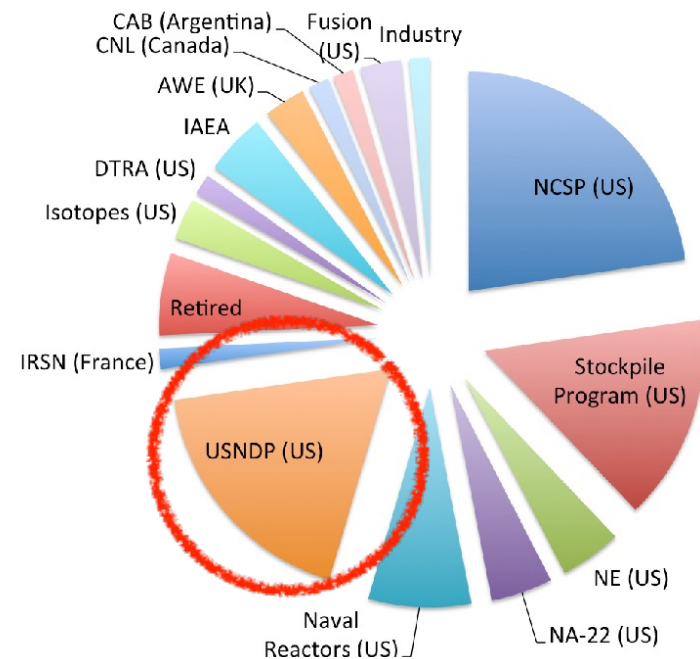
Gustavo Nobre (BNL)  
[gnobre@bnl.gov](mailto:gnobre@bnl.gov)



# CSEWG is a long-standing collaboration between the data users who are also the biggest content providers

By funding specific experimental/evaluation efforts, each program has its needs addressed with the backing of the community.

Fraction of evaluations provided for ENDF/B-VIII.0



## By sharing through CSEWG many sponsors reap the benefit of collaboration

Program	Structure & Decay Evaluation	Reaction Measurement	Reaction Theory	Compilation	Reaction Evaluation	QA (V&V, IE)	Infrastructure (gitlab, etc.)
DTRA		✓					
International (IAEA,NEA,...)		✓	✓	✓	✓	✓	✓
NA-22		✓	✓		✓		
Naval Reactors		✓			✓	✓	
NCSP		✓	✓		✓	✓	✓
Nuclear Energy						✓	
Other (IP, Fusion, ...)		✓	✓		✓	✓	
Defense Programs		✓	✓		✓	✓	✓
USNDP	✓	✓	✓	✓	✓	✓	✓

# Future proofing: Embracing Open Data

# Embracing FAIR, FARR, TRUST, and Open-Source principles

- Not surprisingly, we must follow (heck, we must drive) best practices
- Data licenses & DOIs
- Open source all codes that make sense
- Version tracking of all software & data
- Standing up REST-API's for all data products

## Our Challenges:

- Legacy debt (software & formats)
- Institutional inertia

## LENP Challenges:

- Big facilities (FRIB, Atlas) should be able to support their own users
- Small facilities (ARUNA labs) are too small!

# Future proofing: Preserving access

**Information Wants To Be Free. Information also wants to be expensive. ...That tension will not go away.**

*Brand, Stewart (1987), The Media Lab: Inventing the Future at MIT, Viking Penguin, p. 202, ISBN 0-14-009701-5.*

# ENDF/B-V was RESTRICTED for a time

**ENDF/B-V** The release of ENDF/B-V was originally scheduled for mid-1977. The general purpose library was released at the end of 1978 and the special purpose libraries shortly thereafter. For the first time, the Department of Energy placed restrictions on the distribution of an ENDF/B library. With maturity, it took more time to make improvements to the data which would warrant the release of a new version of ENDF/B. It would never again be released on a two-year cycle.

⋮

Outside the US, ENDF coexisted with local libraries and formats. The foremost of these were the Aldermaston library in the United Kingdom, the KEDAK library in Karlsruhe Germany and the BROND library in the Soviet Union. In the early 1980's Japan adopted the ENDF format for its JENDL library. Western Europe had yet to adopt a common format or library. This all changed with the decision of the US to restrict the distribution of ENDF/B-V to the United States and AECL Chalk River in Canada. From very early times, AECL had taken an active part in CSEWG and provided important data for the library. The decision by the United States to restrict ENDF/V was based on a dispute over the free exchange of data from integral experiments. ENDF/B-V evaluations could be released to a non-CSEWG organization only on a material-by-material basis with DOE approval. Processed nuclear data had the same restrictions. The NNDC had to lock the master tapes in a large safe, which now sits empty in the NNDC library. Very few such requests were granted. The world would not have access to the latest evaluations from CSEWG. The impact was greatest on Western Europe that had no unified library and greatly diminished resources. Under the sponsorship of the Nuclear Energy Agency, their member states decided to develop common evaluated nuclear data file to be called JEF for Joint Evaluated File, which used the ENDF format

C. Dunford, "A CSEWG Retrospective,  
35<sup>th</sup> Anniversary Cross Section Evaluation Working Group" (2001) p. 34-35.

# DOE's Restrictions on ENDF/B-V Splintered the data community

- Led to splintering of the global data project and creation of both JEFF and JENDL data projects
- Never led to the release of the critical assembly data that triggered the spat
- DOE is only now rejoining the NEA Data Bank

Politics should not cloud wisdom



# **Future proofing: Keeping the community vibrant**

# Succession planning

- The ND community is built of several long-running (60+ years) collaborations who deliver data to users
- Every step in the pipeline requires domain expertise
- Need to ensure expertise preserved (& grows to meet new needs!)
- Rule of Two (Sith rule) is not enough
- Mentoring does not begin just before retirement, takes YEARS
- Junior staff need leadership roles



HYMAN H. GOLDSMITH

**H**YMAN H. GOLDSMITH, the initiator of the *Bulletin of the Atomic Scientists* and its co-editor, died on August 7, 1962, at South Weyham, Vermont, as the result of an accident. The last four years of his life had been concerned, with a relentless and impatient devotion, in the cause which the *Bulletin* has served. In the spring and summer of 1945 he had taken a passionate interest in the long, anxious discussions of American atomic energy policy within the Metallurgy Project at the University of Chicago.

After Hiroshima, when the bomb's existence became public knowledge, Goldsmith was one of a relatively large group of atomic-energy scientists who thought it their first duty to guide American policy into the channels of realistic wisdom and enlightened humanity. These were days of public meetings and conferences with civic and political leaders, and scientists who spoke with diplomatic calm and lucidity were required. Goldsmith did not play a great public role in these activities, but he was a conscience-driven incendiary whose explosive scores for indifference and passivity were made acceptable and effective within the circles of scientists themselves by his obviously and profoundly distressed concern for the common good.

All his life he was a disrespector of routines and conventional expectations and restraints. When the situation called for the *Bulletin* to be created, he acted simply on his conviction of that necessity. To say that the *Bulletin* was founded on a whimster would be to describe it as overreached at birth. It lived for many months from hand to mouth, supported by the Atomic Scientists of Chicago, debts, and Goldsmith's conviction.

He chose to make his contribution to the growth of a sense of moral responsibility and political insight among American scientists through a non-protecting activity on behalf of the *Bulletin of the Atomic Scientists*. He loved the *Bulletin* as a Jackson parker. He believed it to be the best and most important journal in the world. He had the interests of the *Bulletin* on his mind wherever his manifold professional duties took him, and everywhere else as well. His vast acquaintanceship, not only among scientists, but

occasions when political frustrations, financial stringency, and the pressing demands of research and teaching have slackened the will to continue.

Goldsmith, by his own example and his angry unwillingness to understand that any reason, however good, could be good enough to justify pessimism about the *Bulletin's* future, or reduction of one's efforts on its behalf, has carried the *Bulletin* through all these crises. He was always on the run, catching a train or a plane, or hurrying to a scientific meeting; but with a tactful boisterous impatience at the door, he kept arguing, prodding, hortatory suggestions for financing, plans for new articles, or improvements, in format. No exhaustion was too deep and no night too long for him to stop discussing, with quick understanding and sharp judgment, the fundamental problems of science and politics, or developing plans for the future of the *Bulletin*.

As co-editor he never wrote an editorial or article for the *Bulletin*, but he was a scrupulously vigilant and often acutely sharp critic of whatever the other members of the Board wrote. The gadfly's stings came where they were needed. His years of random literary and historical reading at the New York Public Library had given him a feeling for the quality of style and ordinariness of presentation which made him a most valuable influence, and his mark, in one way or another, was on everything which went into the *Bulletin*.

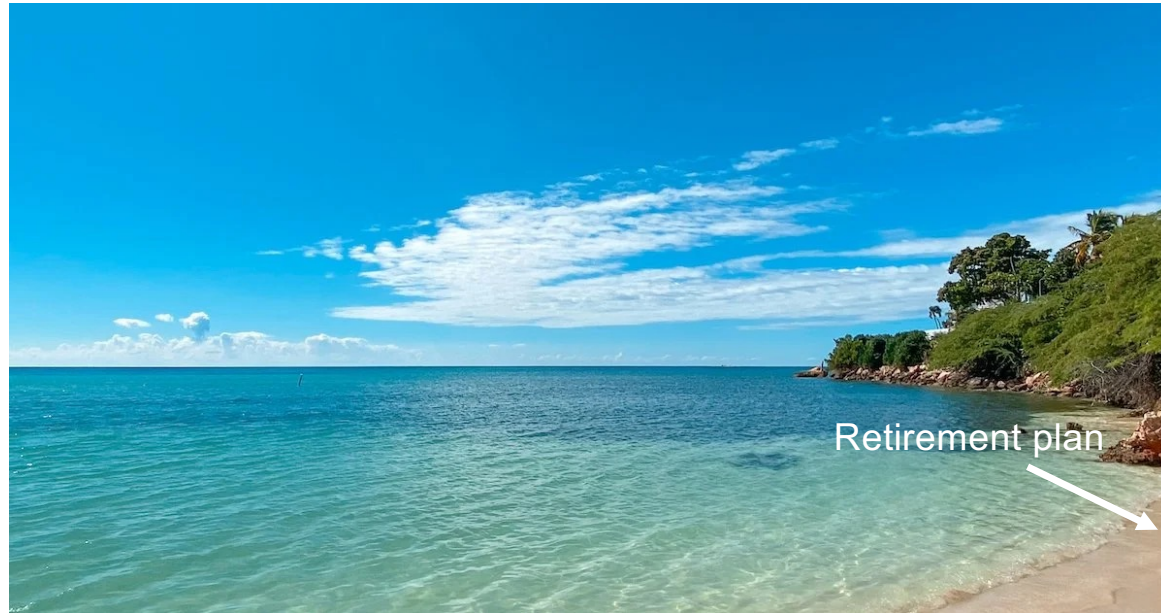
Remarkably, the man whose temper so often ran away with him in conversation, who had little tolerance for those who disagreed with him and little indulgence for human weaknesses, was an inheritor of moderation and balance both in style and in opinion when it came to the contents of the *Bulletin*. He was anxious to see all points of view presented fairly and adequately. He abhorred emotional appeals, public, sentimentalism, and pious recitations. His eye was unerring in discovering catchwords, exaggerations, or empty phraseology. He had respect only for terse, logical presentation, and fair, well-balanced judgment.

We do not know whether Hy Goldsmith was a happy man when he died, and we do not know whether he thought that his life was being filled with accomplishment. Despite



# Future proofing: the data itself

# Where do you see yourself in 30 years?



<https://www.discoverpuertorico.com/article/secluded-beaches-puerto-rico-to-get-away-crowds>

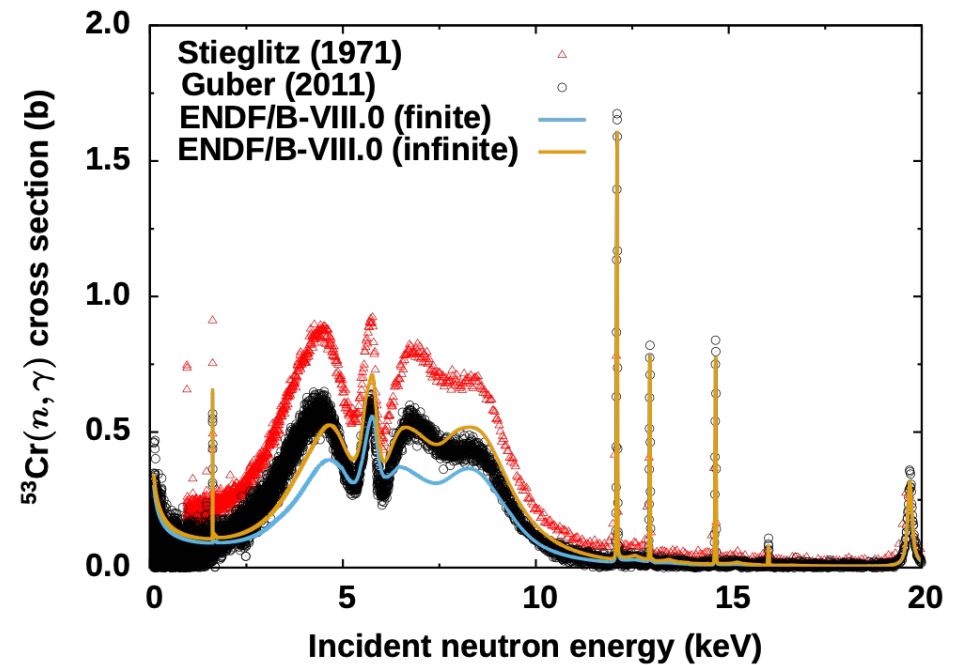
## **You might retire, but your data will still be working!**

We need a sustainable system to ensure our data will be useful far in the future

- Dumping your data in a text file is just the start
- Someone has to keep the servers alive
- Computers and operating systems change with astonishing frequency
- Also, you picked a non-standard format
- Most importantly, who's going to pay for it?
- You and 100's of researchers all have the same problem

# An anecdote

- Cr alloyed in stainless steel, so it is EVERYWHERE
- NNDC led re-evaluation of stable Cr isotopes for ENDF/B library
- Clusters of resonances around 5-20 keV in most isotopes
- Outsized impact on neutron shielding

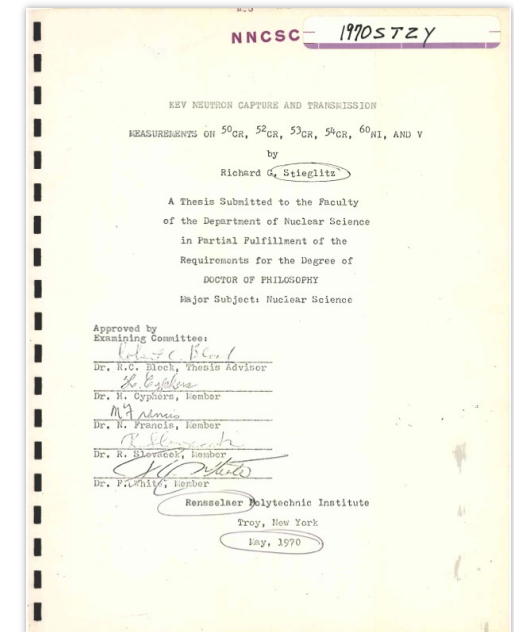
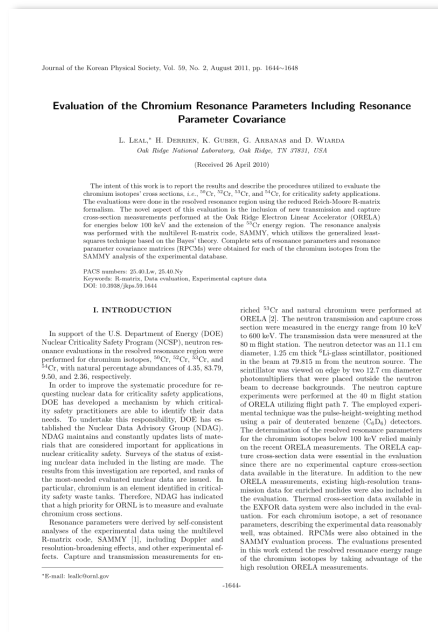


## Who is right?

- ORELA (Guber (2011)) using modern analysis (fitted with full R-matrix, using well studied resolution function), digital DAQ experienced team from ORNL and properly reported uncertainties?
- RPI student Ph.D. project (Stieglitz (1971)), with no error bars?

# Answer: the Ph.D. student was right

- Hinged on understanding multiple scattering correction in both experimental analyses
- Extensive GEANT and MCNP simulations
- Required deep dives into original literature
- Luckily original experimenters were still alive (especially Guber, who was co-author of evaluation!)

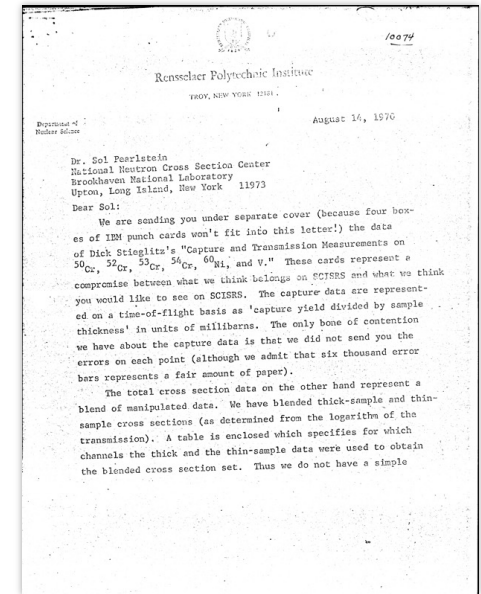


We had to "rerun" both experiments on our cluster

# About the missing uncertainties

- RPI had them. They had full covariances.
- The mean values of the data was shipped to BNL on punched cards (in four boxes)
- Sending the covariance would have meant the death of numerous trees

**Although ARPANET was one year old at the time, there was no realistic way to predict what tools would be available in 50 years**



# You don't know what the future user will need



- **You made a delicious cake, and I asked for the recipe**
  - You'd share it, right?
  - I'll need the recipe to either bake my own
  - I'll need the recipe if I need modify it (dietary restrictions...)
  - I'd also appreciate a "heads up" for any gotchas (do not substitute baking soda for baking powder)

- **We're asking for the same thing for your research**

- *Open Data means sharing the process, not just the data*
- **The Recipe:** Here's a detailed, quantified, well-documented description of what was done
- **What I'll need to modify it:** portable scripts that does the file assembly/analysis/whatever
- **Avoiding gotchas:** the container/virtual environment with all the evaluation and assembly codes, with corresponding inputs, so the user can re-run the research as you intended!

# Acknowledgements

This work was supported by the Nuclear Criticality Safety Program, funded and managed by the National Nuclear Security Administration for the U.S. Department of Energy. Additionally, work at Brookhaven National Laboratory was sponsored by the Office of Nuclear Physics, Office of Science of the U.S. Department of Energy under Contract No. DE-SC0012704 with Brookhaven Science Associates, LLC.



U.S. DEPARTMENT  
*of* **ENERGY** | Office of  
Science

