Automation in the Nuclear Data Pipeline

David Brown

National Nuclear Data Center

Brookhaven National Laboratory

BROOKHAVEN NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

BROOKHAVEN SCIENCE ASSOCIATES

Steps in the ENDF library preparation

1. **ENDF/A** — store the in-development evaluations here
2. **Phase I Review** — simple physics and format checks (“unit tests”)
3. **Phase II Review** — detailed validation in critical assemblies; led by CSEWG validation committee (“integration tests”)
4. **Library release as ENDF/B**

This process has been in place since ENDF/B-I (1968).
It is good engineering practice

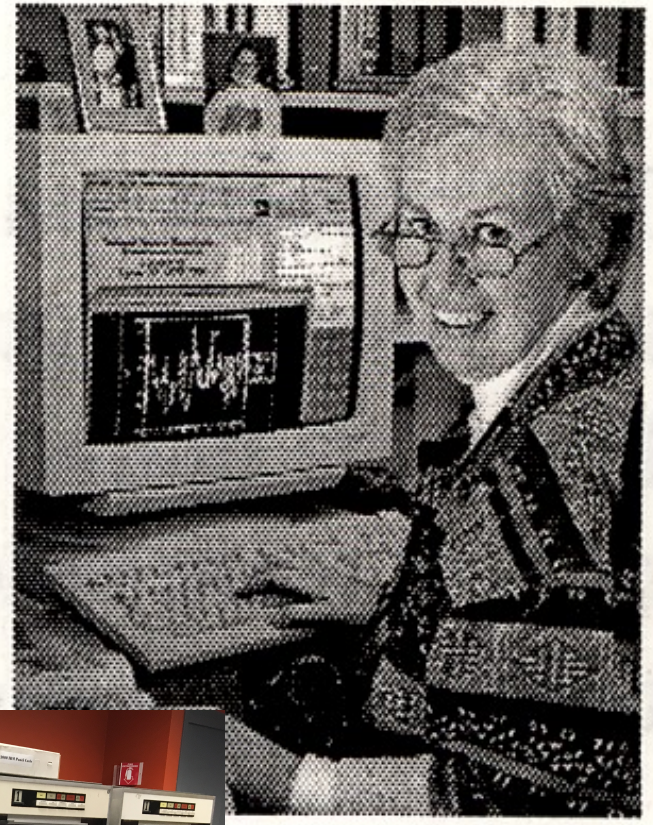
The ENDF library in days of yore...

ENDF/B-VI.8 was the last ENDF library produced using the “old ways”

Summary of Tapes in ENDF/B-VI release 8

ENDF/B-VI (and all older ENDF libraries) was released as a series of tapes containing different portions of the full library. As new revisions were released, users were to replace materials/datasets from older tapes with that in the newer tapes. Below we summarize these tapes.

TAPE #	Release	Library	# MATS	# Records	Contents
100	0	Neutron sublibrary	7		
101	0	Neutron sublibrary	16		
102	0	Neutron sublibrary	37		
103	0	Neutron sublibrary	22		
104	0	Neutron sublibrary	40		
105	0	Neutron sublibrary	68		
106	0	Neutron sublibrary	38		
107	0	Neutron sublibrary	5		
108	0	Neutron sublibrary	10		
109	0	Neutron sublibrary	14		
110	0	Neutron sublibrary	16		
111	0	Neutron sublibrary	4	replaced	
112	0	Neutron sublibrary	4	replaced	
113	0	Neutron sublibrary	6		
114	0	Neutron sublibrary	3		
115	0	Neutron sublibrary	6		
116	0	Neutron sublibrary	5		
117	0	Neutron sublibrary	4	replaced	
118	0	Thermal neutron scattering	5		
119	0	Thermal neutron scattering	4		
120	1	Neutron sublibrary	19	44 272	
121	1	Neutron sublibrary	8	14 459	
122	1	Neutron sublibrary	4	29 326	
123	1	Neutron sublibrary	4	43 411	
124	1	Neutron sublibrary	5	34 487	
125	1	Neutron FPY sublibrary	11	replaced	
126	1	Neutron FPY sublibrary	16	replaced	
127	2	Neutron sublibrary	21	41 980	

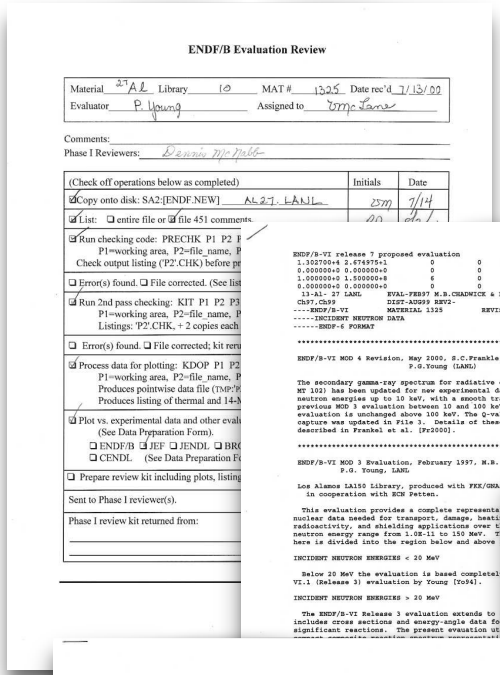


Vicki McLane
(1940-2011)

BROOKHAVEN NATIONAL LABORATORY

Phase I reviews used hybrid of e- & snail mail

Review packet (snail) mailed to reviewers


ENDF/B Evaluation Review

Material 27Al Library 10 MAT # 1325 Date rec'd 7/13/00

Evaluator P. Young Assigned to Dennis McNabb

Comments:

Phase I Reviewers: Dennis McNabb

(Check off operations below as completed)	Initials	Date
☑ Copy onto disk: SA2:[ENDF.NEW]_AL27.LANL		7/14
☑ List: ☐ entire file or ☑ file 451 comments.		
☑ Run checking code: PRECHK P1 P2 P3 P1=working area, P2=file_name, P Check output listing (P2'.CHK) before pr		
☐ Error(s) found. ☐ File corrected. (See list		
☑ Run 2nd pass checking: KIT P1 P2 P3 P1=working area, P2=file_name, P Listings: 'P2'.CHK, + 2 copies each		
☐ Error(s) found. ☐ File corrected; kit rerun		
☑ Process data for plotting: KDOP P1 P2 P1=working area, P2=file_name, P Produces pointwise data file (TMP.'P Produces listing of thermal and 14-M		
☑ Plot vs. experimental data and other evalu (See Data Preparation Form). ☐ ENDF/B ☑ JEF ☐ JENDL ☐ BRO ☐ CENDL (See Data Preparation Fo		
☐ Prepare review kit including plots, listing		
Sent to Phase I reviewer(s).		
Phase I review kit returned from:		

ENDF/B-VI release 7 proposed evaluation

ENDF/B-VI MOD 4 Revision, May 2000, S.C.Frankle, R.C.Reedy, P.G.Young (LANL)

The secondary gamma-ray spectrum for radiative capture (MF 12, MT 102) has been updated for new experimental data at incident neutron energies up to 10 keV, with a smooth transition to the previous MOD 3 evaluation between 10 and 100 keV. The MT=102 evaluation is unchanged above 100 keV. The Q-value for radiative capture was updated in File 3. Details of these changes are described in Frankel et al. [Fr2000].

ENDF/B-VI MOD 3 Evaluation, February 1997, M.B. Chadwick and P.G. Young, LANL

Los Alamos LA150 Library, produced with FKK/GHASH in cooperation with ECN Petten.

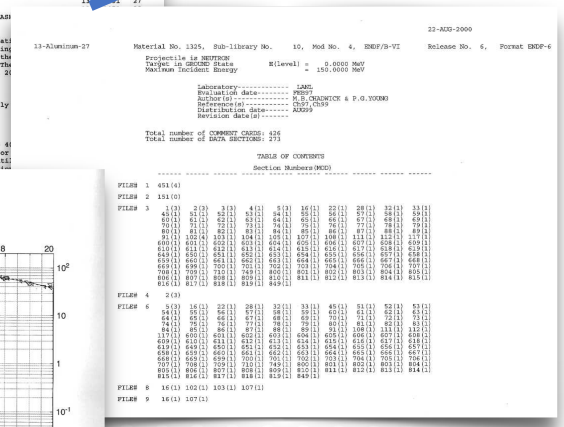
This evaluation provides a complete representation of nuclear data needed for transport, damage, heating, radioactivity, and shielding applications over the neutron energy range from 1.0X-11 to 150 MeV. The here is divided into the region below and above 2

INCIDENT NEUTRON ENERGIES < 20 MeV

Below 20 MeV the evaluation is based completely on VI.1 (Release 3) evaluation by Young [Yo94].

INCIDENT NEUTRON ENERGIES > 20 MeV

The ENDF/B-VI Release 3 evaluation extends to 4 includes cross sections and energy-angle data for significant reactions. The present evaluation uti


13-Aluminum-27 Material No. 1325, Sub-library No. 10, Mod No. 4, ENDF/B-VI Release No. 6, Format ENDF-6

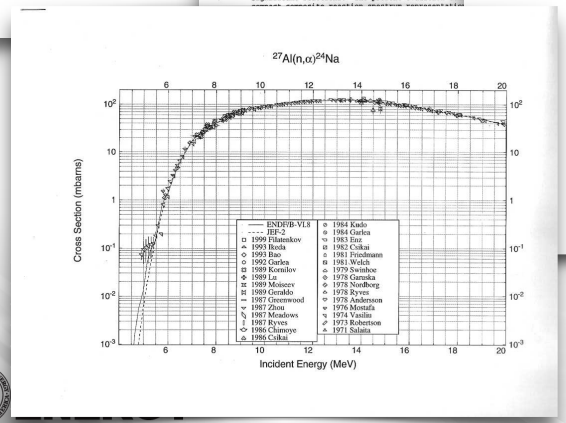
Projectile is NEUTRON
Target in GROUND State E(level) = 0.0000 MeV
Maximum Incident Energy = 150.0000 MeV

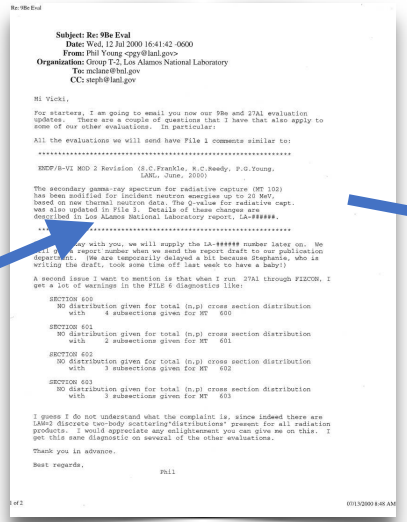
Laboratory---------- LANL
Evaluation date----- PEB97
Author(s)----------- M.B.CHADWICK & P.G.YOUNG
References(s)------- Ch97,Ch99
Distribution date--- AUG99
Revision date(s)---

Total number of COMMENT CARDS: 426
Total number of DATA SECTIONS: 573

TABLES OF CONTENTS

Section Numbers(MOD)




Re: 9Be Eval

Subject: Re: 9Be Eval
Date: Wed, 12 Jul 2000 16:41:42 -0600
From: Phil Young <pgy@lanl.gov>
Organization: Group T-2, Los Alamos National Laboratory
To: mclane@bnl.gov
CC: steph@lanl.gov

Hi Vicki,

For starters, I am going to email you now our 9Be and 27Al evaluation updates. There are a couple of questions that I have that also apply to some of our other evaluations. In particular:

All the evaluations we will send have File 1 comments similar to:

ENDF/B-VI MOD 2 Revision (S.C.Frankle, R.C.Reedy, P.G.Young, LANL, June, 2000)

The secondary gamma-ray spectrum for radiative capture (MT 102) has been modified for incident neutron energies up to 20 MeV, based on new thermal neutron data. The Q-value for radiative capt. was also updated in File 3. Details of these changes are described in Los Alamos National Laboratory report, LA-######.

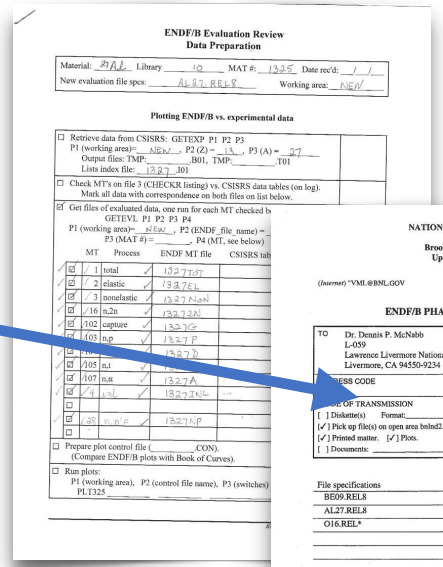
I guess I do not understand what the complaint is, since indeed there are LAB-2 discrete two-body scattering"distributions" present for all radiation products. I would appreciate any enlightenment you can give me on this. I get this same diagnostic on several of the other evaluations.

Thank you in advance.

Best regards,

Phil

Response to "referee report" e-mailed


ENDF/B Evaluation Review
Data Preparation

Material: 27Al Library 10 MAT #: 1325 Date rec'd:
New evaluation file spec: AL27.REL8 Working area: NEW

Plotting ENDF/B vs. experimental data

Retrieve data from CSISRS: GETEXP P1 P2 P3
P1 (working area)= NEW , P2 (Z) = 13 , P3 (A) = 27
Output files: TMP: .B01, TMP: .T01
Lists index file: 1327 .I01

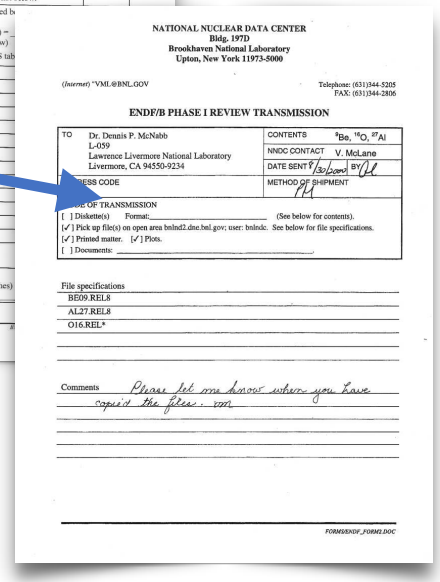
Check MT's on file 3 (CHECKR listing) vs. CSISRS data tables (on log).
Mark all data with correspondence on both files on list below.

Get files of evaluated data, one run for each MT checked b
GETEVL P1 P2 P3 P4
P1 (working area)= NEW , P2 (ENDF_file_name) =
P3 (MAT #) = , P4 (MT, see below)

	MT	Process	ENDF MT file	CSISRS tab
☑	1	total	1327MT	
☑	2	elastic	1327EL	
☑	3	nonelastic	1327NON	
☑	16	n,2n	1327N2N	
☑	102	capture	1327G	
☑	103	n,p	1327P	
☐			1327D	
☑	105	n,t		
☑	107	n,a	1327A	
☑	1	UNL	1327UNL	
☐				
☑	28	n,np	1327NP	
☐				

Prepare plot control file (.CON).
(Compare ENDF/B plots with Book of Curves).

Run plots:
P1 (working area), P2 (control file name), P3 (switches)
P1,1325


NATIONAL NUCLEAR DATA CENTER
Bldg. 197D
Brookhaven National Laboratory
Upton, New York 11973-5000

(Internet) "VML@BNL.GOV

Telephone: (631)344-5205
FAX: (631)344-2806

ENDF/B PHASE I REVIEW TRANSMISSION

TO: Dr. Dennis P. McNabb, L-059, Lawrence Livermore National Laboratory, Livermore, CA 94550-9234

CONTENTS: ^{9}Be, ^{16}O, ^{27}Al
NNDC CONTACT: V. McLane
DATE SENT: 8/30/2000 BY:
METHOD OF SHIPMENT:

TYPE OF TRANSMISSION
[ ] Diskette(s) Format: (See below for contents).
[✓] Pick up file(s) on open area bnlnd2.dnc.bnl.gov; user: bnlndc. See below for file specifications.
[✓] Printed matter. [✓] Plots.
[ ] Documents:

File specifications
BE09.REL8
AL27.REL8
O16.REL*

Comments: Please let me know when you have copied the files. vm

FORM/ENDF_FORM1.DOC

Final acceptance report, snail mail

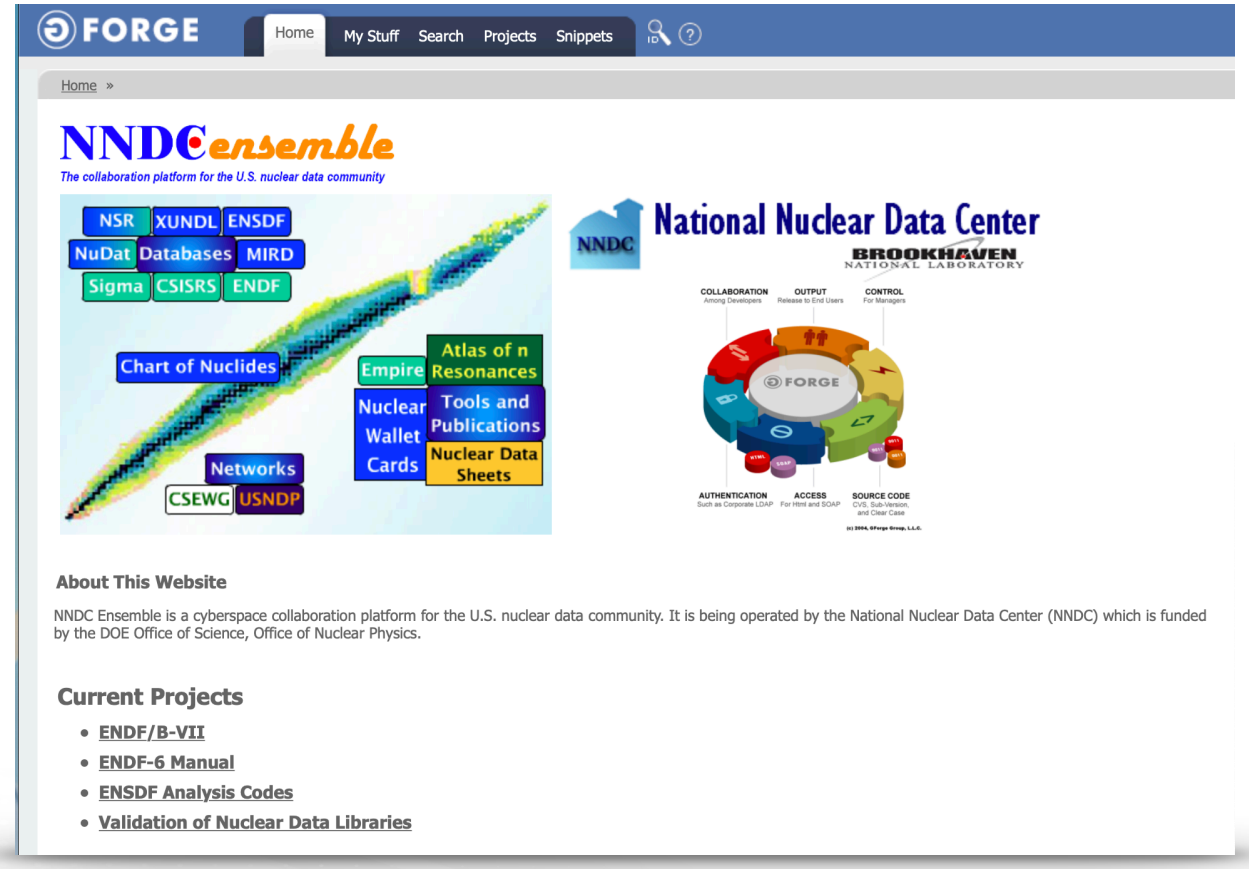
n+^{27}Al, P. Young (LANL) July 2000
Phase I review: D. McNabb

BROOKHAVEN NATIONAL LABORATORY

And then things began to change...

In 2009, Michal Herman introduced a revision control system (svn) to the ENDF process

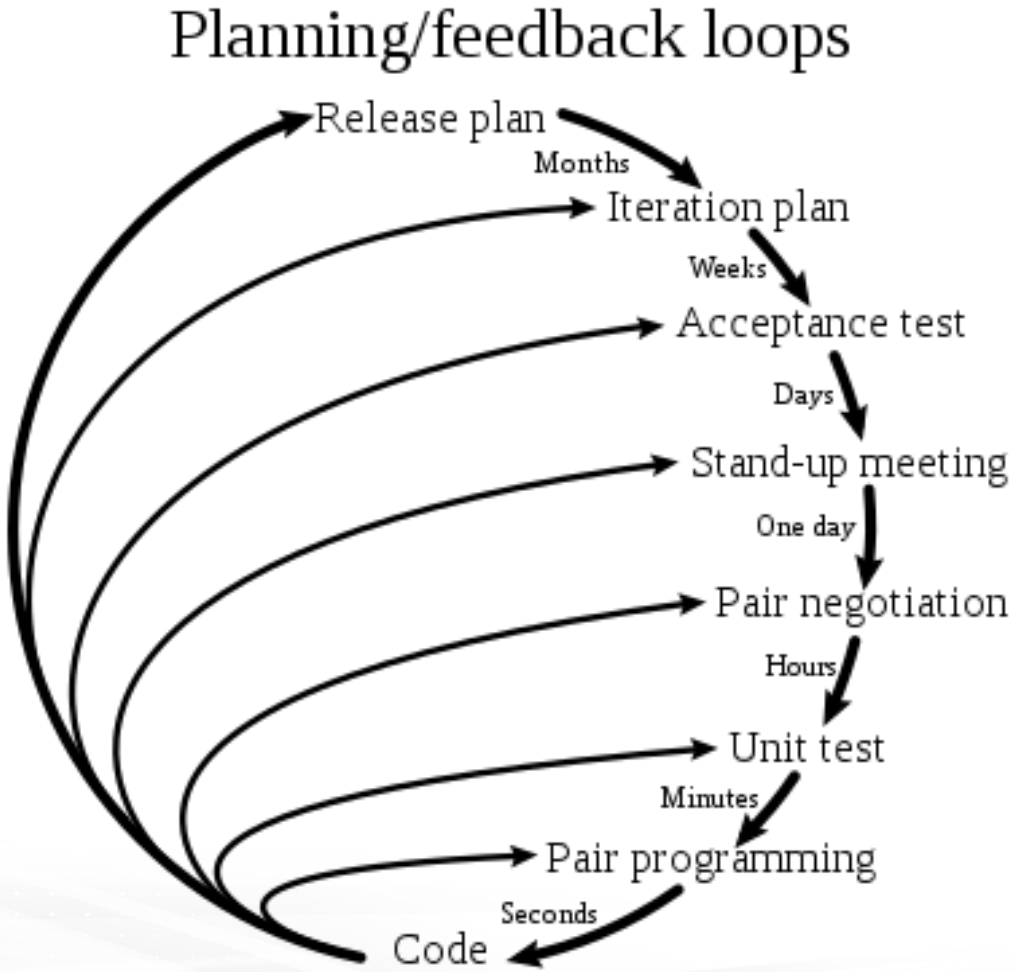
- No longer humans manage the tapes and the changesets
- Revisions no longer stored in (paper!) meeting minutes, now are logged with the changeset
- ***First tagged release: ENDF/B-VII.0 in 2009 (3 years after the release!)***
- ENDF/A is now just a branch



https://ndclx4.bnl.gov/gf/

In the early 2000's, everything was eXtreme!

- ***Extreme Programming (XP)***
- Run tests on every commit to an RCS
- Enables rapid code development
- ***Bugs detected early, as they are created, rather than being left as "landmines" to be discovered later "the hard way"***
- Key part of Agile Programming
- ***Test running automated with a continuous integration system***



M. Fowler and J. Highsmith, "The agile manifesto," *Software Development*, vol. 9, no. 8, pp. 28–35, 2001.

Birth of ADVANCE

- In September 2011, a new NNDC hire (me) was tasked with preparing the ENDF/B-VII.1 release
- The release was due in December
- No Phase I testing had be done (yet)
- Ramon Arcilla and I stood up the first ENDF continuous integration system
- ***ENDF/B-VII.1 released on 22 December, 2011***

ENDF/B-VII.1/neutrons/Ta

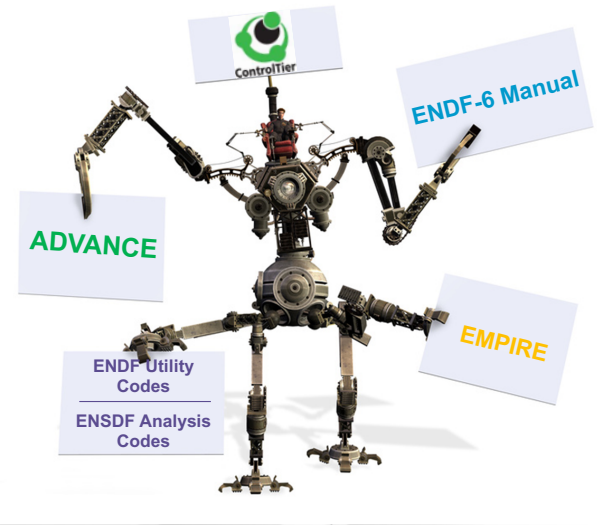
status	isotope	abundance	svn log	ENDF (orig. file)	STANEF (log)	STANEF (output diff)	CHECKR (log)	FIZCON (log)	PSYCHE (log)	fudge (warnings log)	reactionSuite (gnd file)	covarianceSuite (gnd file)	NJOY (log)	AMPX (log)
	^{180}Ta	0.012 +/- 0.002 %				Files identical			CORE	No file		No file	No file	No file
✓	^{181}Ta	99.988 +/- 0.002 %				Files identical			(28 URR dens. (a))	(1 skip total, 1 redun URR (a))		No file	No file	No file
✓	^{182}Ta	0.0 %				Files identical			(1 iffy Q)	(2 xtra interp, 1 skip total, 1 redun URR (a))		No file	No file	No file

What do the messages mean?

1. **fudge** info "redun URR (a)": Redundant data in URR
2. **fudge** info "xtra interp": Multple interpolation regions in a file where there didn't need to be that many.
3. **fudge** unimplimented "skip total": Fudge currently ignores the total cross section.
4. **psyche** unknown "URR dens. (a)": Level density in URR not in agreement with PSYCHE's, possibly misguided, expectations
5. **psyche** unknown "iffy Q": Non-threshold reaction with Q value differing from PSYCHE's, possibly misguided, expectations

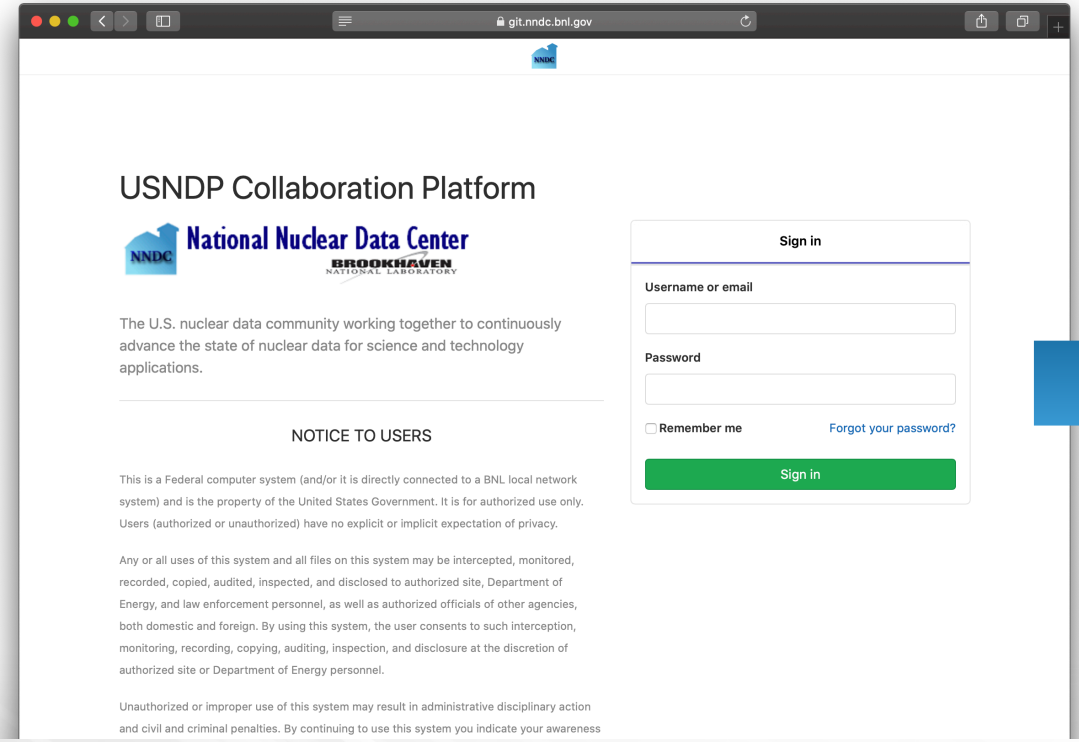
Page generated at 2011-11-08 10:57:48.143028
BNL and/or NNDC boilerplate here

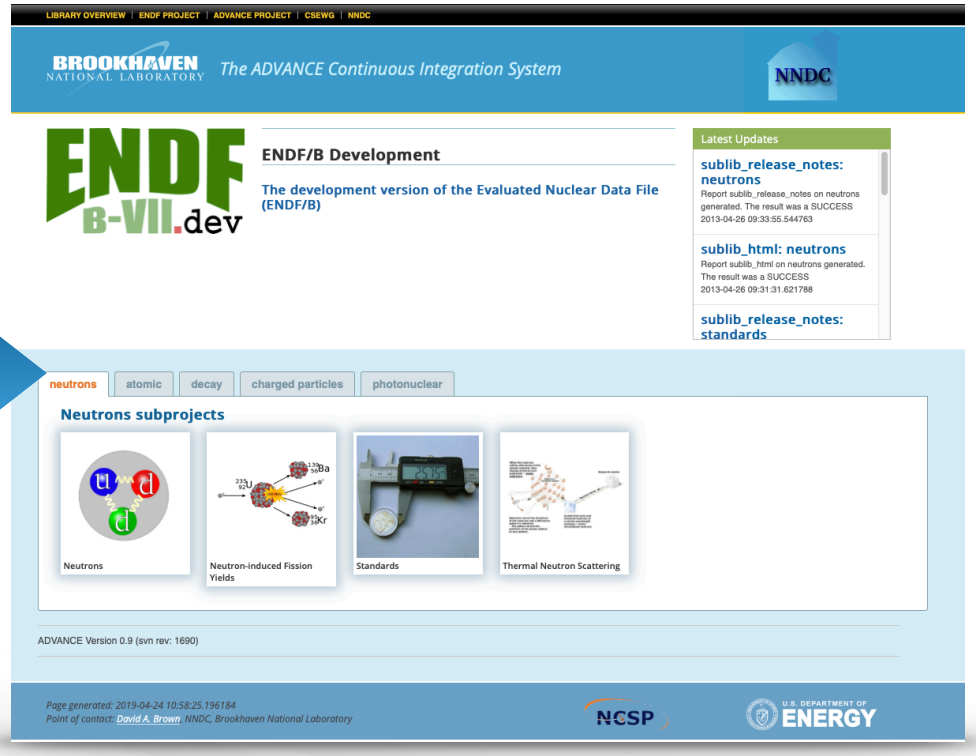
First build report, circa 2011



ADVANCE today

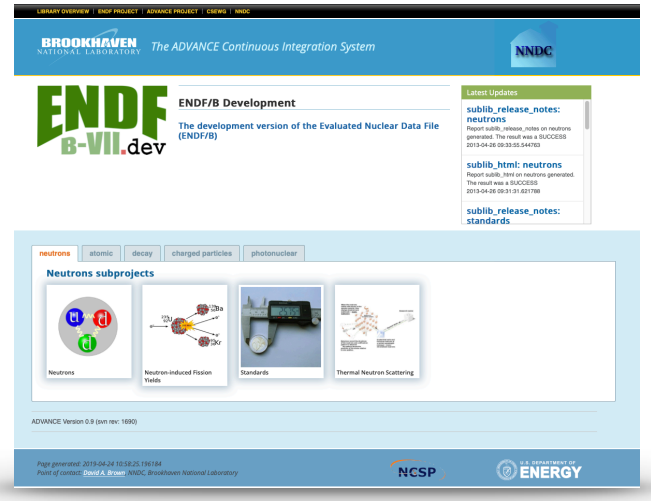
We are transitioned from GForge to gitlab.nndc.bnl.gov, but ADVANCE still running

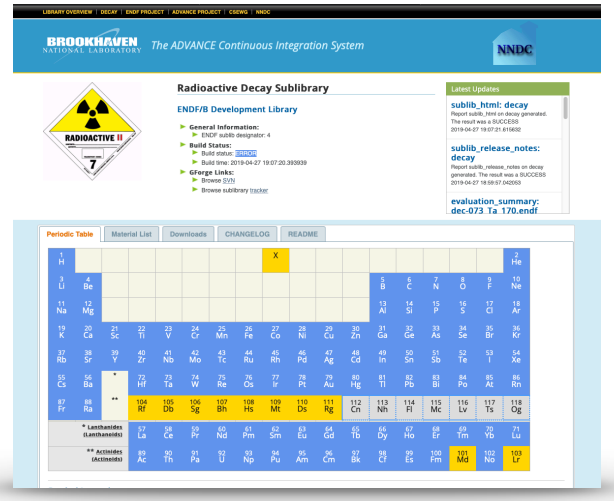


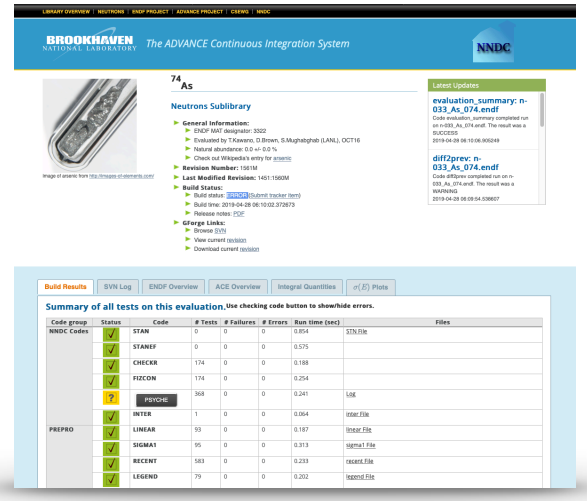


Build reports published to

https://www.nndc.bnl.gov/endf/b7.dev/qa







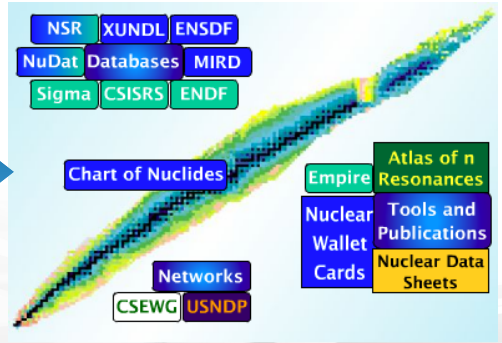
git.nndc.bnl.gov

Changes

advance2.nndc.bnl.gov

Status

www.nndc.bnl.gov



Phase I testing automated for nearly 10 years

Can we automate reading build reports? What other tests can we automate?

Code	Test	pre-VII	Now
	File summary complete & correct	🤔	
STAN, STANEF, CHECKR, fudge	**ENDF format compliance**	🖥	🖥
FIZCON, fudge	**Mathematical correctness** (e.g. probabilities valid, covariances positive)	🖥	🖥
FIZCON, PSYCHE, fudge	**Physical correctness** (e.g. Q, thresholds, energy deposition/KERMA)	🖥	🖥
INTER, fudge (inter.py)	**Compute & check integral metrics** (e.g. RI, thermal cross sections, MACS)	🖥	🖥
fudge	**Completeness** (all outgoing particles, including gammas)	🤔	🖥
ADVANCE	**Comparisons to microscopic experimental data** (EXFOR)	🤔	🖥
	Assessment of application suitability (e.g. usable for fast reactors or spaceflight)	🤔	
	Reasonable (e.g. covariances, angular distributions)	🤔	
NEW fudge (grokres.py)	**Resonance quality** (missing resonances? widths realistic?)		🖥
PREPRO, fudge, NJOY (not SCALE yet)	**Can process for user codes**		🖥
	Is state of the art? Is best we can do?	🤔	

Treating Data Like Software: A Case for Production Quality Data

Jennifer M. Schopf
Woods Hole Oceanographic Institution
Woods Hole, MA 02543
(Currently at the National Science Foundation, GEO/OAD)
jschopf@whoi.edu

ABSTRACT

In this short paper, we describe the production data approach to data curation. We argue that by treating data in a similar fashion to how we build production software, that data will be more readily accessible and available for broad re-use. We should be treating data as an ongoing process. This includes considering third-party contributions; planning for cyclical releases; bug fixes, tracking, and versioning; and issuing licensing and citation information with each release.

Categories and Subject Descriptors

E.5.3 [**Data**]: Files - *Organization, Structure*; E.4.3 [**Data**]: Coding and information theory - *Formal models of communication*; H.1 [Information Systems] - Models and principles

General Terms

Management, Documentation, Design, Verification

Keywords

available digitally, the ability to find and access data is increasingly difficult.

In order to address the need for better data preservation and access, we propose that data sets should be managed in a similar fashion to how we maintain production quality software. These *production data sets* are not simply published once, but go through a cyclical process of development, verification, deployment, support, analysis, and then development again. Attention is given to ensuring the data is understandable, useful, and updated over time, the same way software products need updating over time, even if the core functionality does not change.

This short paper gives a brief definition of what is meant by data in this context. It then addresses at a high level standard factors that are part of the development of (academic) production software, and describes how similar processes can be applied to enable data sets to have extended lifecycles and improved usability. A key premise is that if this approach can be integrated into common practice, it will result in a higher level of preservation and

Jennifer M. Schopf. 2012. Treating data like software: a case for production quality data. In Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries (JCDL '12). Association for Computing Machinery, New York, NY, USA, 153–156. DOI:https://doi.org/10.1145/2232817.2232846

Automation in the future?

Our workflow is already a Bayesian network!

Without aggressive optimization, we cannot bring full power of machine learning to bear

A Bayesian network

We need an automated system to keep up with future flux of data

This is more than just automation taken to the extreme – this is a scalable distributed open source machine learning framework hosted on virtual machines

Free the physicists to do physics!

Modernizing USNDP databases is key for vision

Nuclear Science References	Structure databases (XUNDL & ENSDF)	Reaction database (EXFOR)
1960's vintage infrastructure => proposed modernization	1960's vintage infrastructure => modernization in progress	1960's vintage infrastructure => modernization complex due to international collaboration
• Needs Natural Language Processing • Has extensive training data	• Needs Natural Language Processing • Needs table/figure extraction • Has extensive training data	• Needs Natural Language Processing • Needs table/figure extraction • Has extensive training data
Bibliographic/meta data only	"Homogeneous data" "few" classes of measurements/data	"Heterogeneous data" "many" classes of measurements/data
Potential journal collaboration	Active journal collaboration with Phys. Rev and EPJ	Potential journal collaboration

We are starting the process this FY

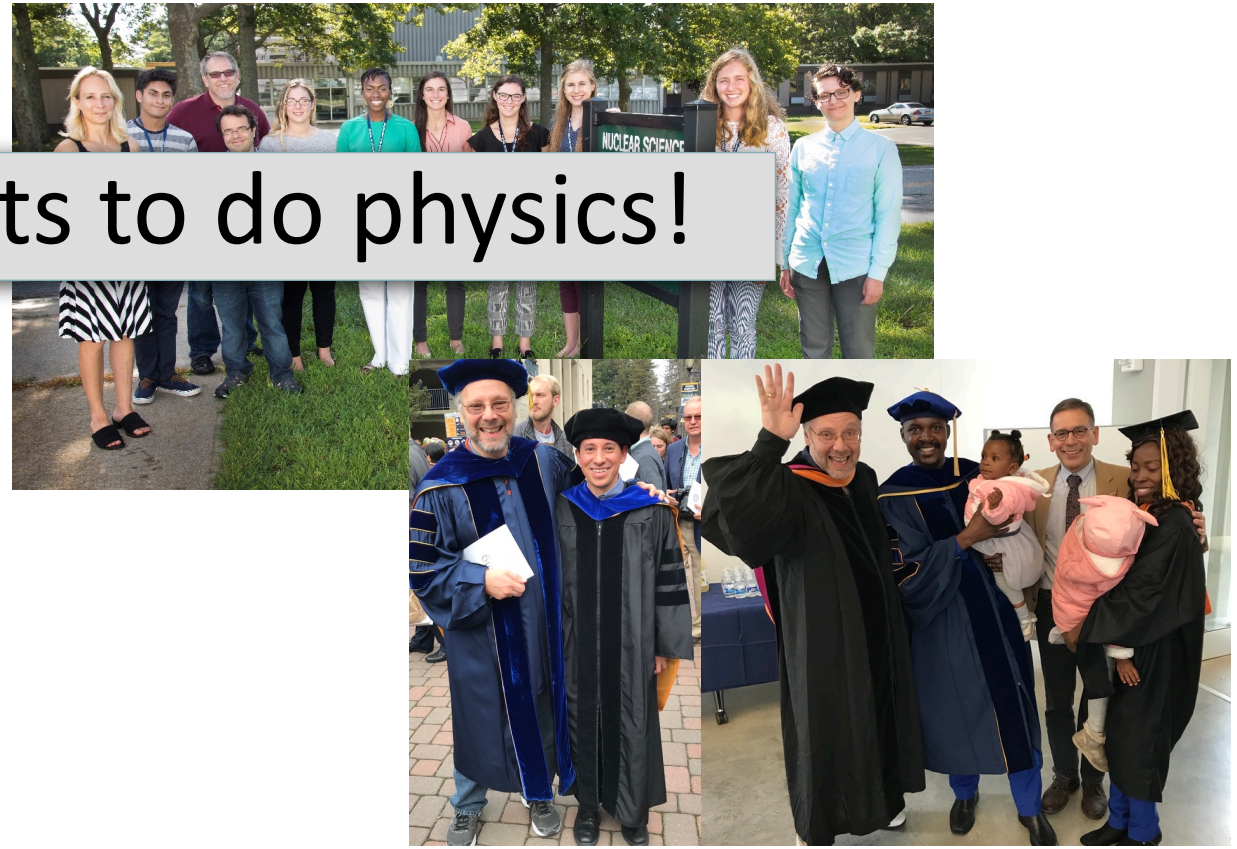
But, if we automate, where do the people go?

- Perform new (but targeted) experiments
- Improve models
- Improve hardware & software
- Deal with “tough cases”
- ***New physics***
- ***New applications***

But, if we automate, where do the people go?

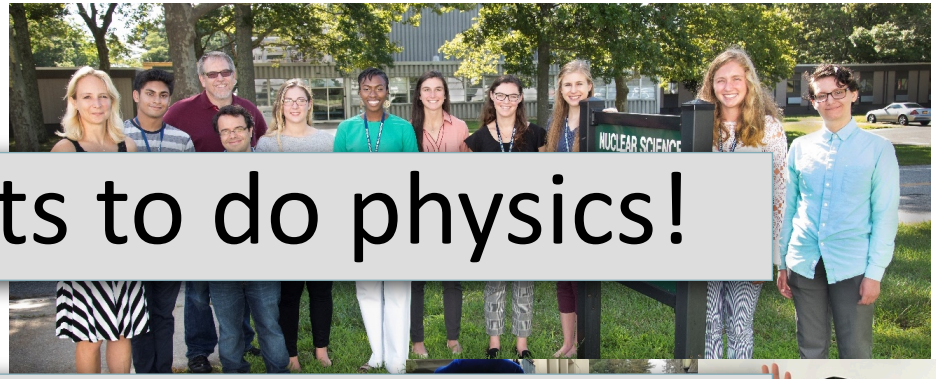
- Perform new (but targeted) experiments
- Improve
- Improve hardware & software
- Deal with “tough cases”
- ***New physics***
- ***New applications***

Free the physicists to do physics!



But, if we automate, where do the people go?

- Perform new (but targeted) experiments
- Improve
- Improve hardware & softwar
- Deal wi
- ***New ph***
- ***New ap***



Free the physicists to do physics!

But don't forget to support the "boring infrastructure stuff" that needs to get done to support the physicists!