Los Alamos NATIONAL LABORATORY

Meeting on selecting a database type and format for the SG-50 database: introduction and user needs from a smaller-scale every day user's perspective

Denise Neudecker
Thanks to D. Brown, A. Lovell, M. Paris

WPEC SG-50 9/27/2021

LA-UR-21-29288

Los Alamos
NATIONAL LABORATORY

Introduction

The main charge of this meeting is:

(A) to decide what database type and format we want to use for SG-50, and

(B) to get volunteers for developing this interface from data in the EXFOR database and format to the SG-50 counter-part.

These two are strictly tied together. We will only select a database type and format that the volunteers from point (B) feel comfortable developing!

First part: What are user needs and other requirements on the database.

5 min	D. Neudecker	Introduction
13 min	D. Neudecker	User needs from a smaller-scale every day user's perspective
12 min	A. Koning	User needs from a larger-scale every day user's perspective
5 min	A. Lewis & DN	What do we already have?
10 min	A. Hayes	Experiences dealing with hierarchical ENSDF data
45 min	Discussion	• What are further user needs? • Do we need offline options? Online enough? • Cyber-security considerations at the hosting labs? • Volunteers for developing interface from EXFOR data to SG-50 database
10 min	Break	

Second part: What database type and format should we use?

15 min	V. Zerkin	X4Lite2 and progress in EXFOR data automatic renormalization system
60 min	Discussion	• Experiences with DBs and which one? • What are the preferences of developers on databases? • Which format? • Final selection of databases.
5 min	Discussion	Next steps

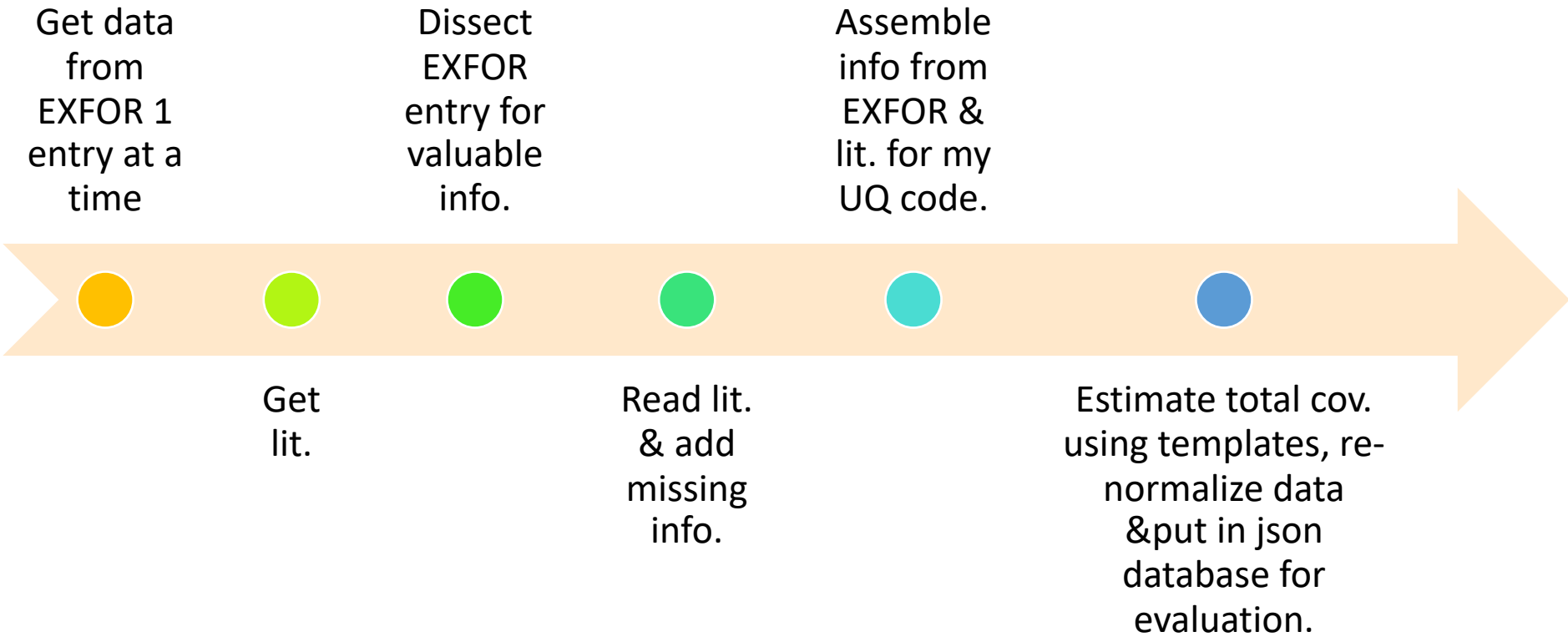
Los Alamos NATIONAL LABORATORY

User needs from a smaller-scale every day user's perspective

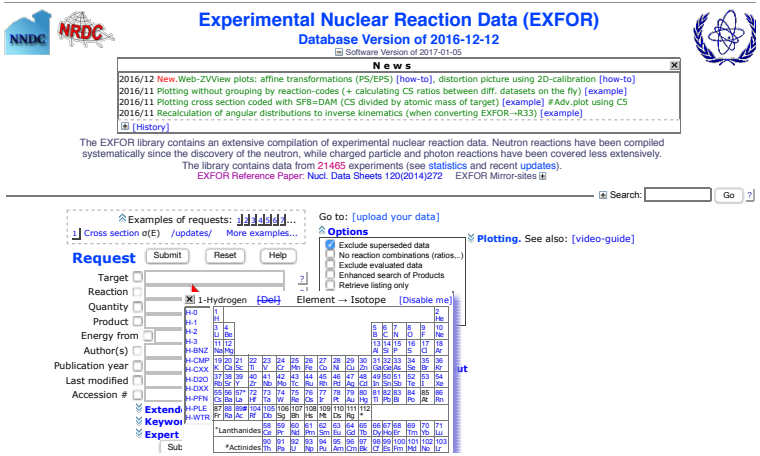
I use EXFOR pretty much every work-day for:

- Selecting the input for my nuclear data evaluations → This is a small-scale use of one entry at a time. I am going through all the details in the EXFOR entry.
- Counter-checking whether nuclear data are realistic as part of my nuclear data validation effort. → larger-scale use of all entries of one observable at a time.

As an evaluator, I do the following:



Step 1: getting the data from EXFOR



EXFOR data: http://www.nndc.bnl.gov/EXFOR/40493.002
Data retrieved from the EXFOR database version of June 23, 2020.

Search by Entry: 40493
Authors:
M.V. Savin
Ju.A. Khokhlov
V.N. Ludin

```
ENTRY        40493001     850329               20050926        0000
SUBENT       40493001     850329               20050926        0000
BIB             12           16
INSTITUTE    (4CCPKUR)
             #(4CCPKUR) Institut Atomnoi Energii I.V. Kurchatova,
Moskva, Union of Soviet Socialist Republics
REFERENCE    (R,YFI-27,4,7911)
             (C,73KIEV,4,63,7305) PRELIMINARY RESULTS
             #(R,YFI-27,4,7911) Rept: Yaderno-Fizicheskie Issledovaniya Reports,
No.27, p.4 (1979), Russia
             #+       #Title=AVERAGE NUMBER OF PROMPT NEUTRONS AT THE U-235 FISSION
BY THE NEUTRONS IN THE ENERGY INTERVAL 0.2 - 1.0 MEV
             #+       #Authors=M.V.Savin,Ju.A.Khokhlov,V.N.Ludin
             #(C,73KIEV,4,63,7305) Conf: 2.Conf.on Neutron Physics, Kiev 1973, Vol.4,
p.63 (1973), USSR
AUTHOR       (M.V.SAVIN,JU.A.KHOKHLOV,V.N.LUDIN)
TITLE         AVERAGE NUMBER OF PROMPT NEUTRONS
              AT THE U-235 FISSION BY THE NEUTRONS
              IN THE ENERGY INTERVAL 0.2 - 1.0 MEV
FACILITY     (LINAC)
             #(LINAC) Linear accelerator
```

- Enter reaction code to get data of specific isotope and observable x,
- I use x4+,
- I save every EXFOR entry that I work on locally as a pdf on my desktop to have a copy of my evaluation input. → I have seen entries vanish and change which is problematic for keeping track and documenting my evaluation input.

Step 2: getting literature

- I try to get as much literature as possible for each data set (yes, multiple per experiment),
- I love the DOI links!!!!

```
            #(1USAANL) Argonne National Laboratory, Argonne, IL, United States of
America
REFERENCE   (J,JNE,21,157,1967)
            #(J,JNE,21,157,1967) Jour: Journal of Nuclear Energy, Vol.21, p.157
(1967), UK
            #+        #URL=http://dx.doi.org/10.1016/0022-3107(67)90125-6
            #+        #DOI=10.1016/0022-3107(67)90125-6
            #+        #Title=Energy dependence of Nu bar (p) for neutron-induced
fission of U235 below 1.0 MeV.
            #+        #Authors=J.W.Meadows,J.F.Whalen
AUTHOR      (J.W.Meadows,J.F.Whalen)
TITLE       Energy dependence of Nu bar (p) for neutron-induced
            fission of U235 below 1.0 MeV.
FACILITY    (EXFOR-1USAANL)
```

Step 3: dissecting the EXFOR entry, I copy out:

- Information on measurement features (detector type, sample mass, impurities, etc.) → I use this to find common trends among measurements that have a sub-set of features available.
- the monitor reaction and old monitor values to re-normalize to the newest nuclear data.
- Info from the “correction” field as it gives me an understanding what corrections could be missing.
- Info from err-analysis and re-name unc. following a common nomenclature to identify which ones are missing or underestimated.
- The data, of course.

Step 3: dissecting the EXFOR entry, I copy out:

```
            1% in neutron spectra).
            - Multiple scattering and neutron attenuation
              corrections have not been performed.
ERR-ANALYS Uncertainty in outgoing neutron energy due to
            calibration is
                       less than 2% for En < 5MeV,
                       4%           for En < 10 MeV,
                       8%           for En > 10 MeV
           (ERR-T) Sum of counting and experimental uncertainties.
                   - Both uncertainties were propagated throughout
                     the analysis when significant.
                   - Uncertainty associated with the experiment
                     beyond counting statistics and efficiency
                     correction are rather negligible.
                   - At high energy portions of the spectra,
                     statistical uncertainties due to background
                     subtraction are the DATA              6        275        12
                     to overall error.
           (ERR-1) Uncertainty in effici
```

One example of a rather detailed ^{239}Pu PFNS EXFOR entry.

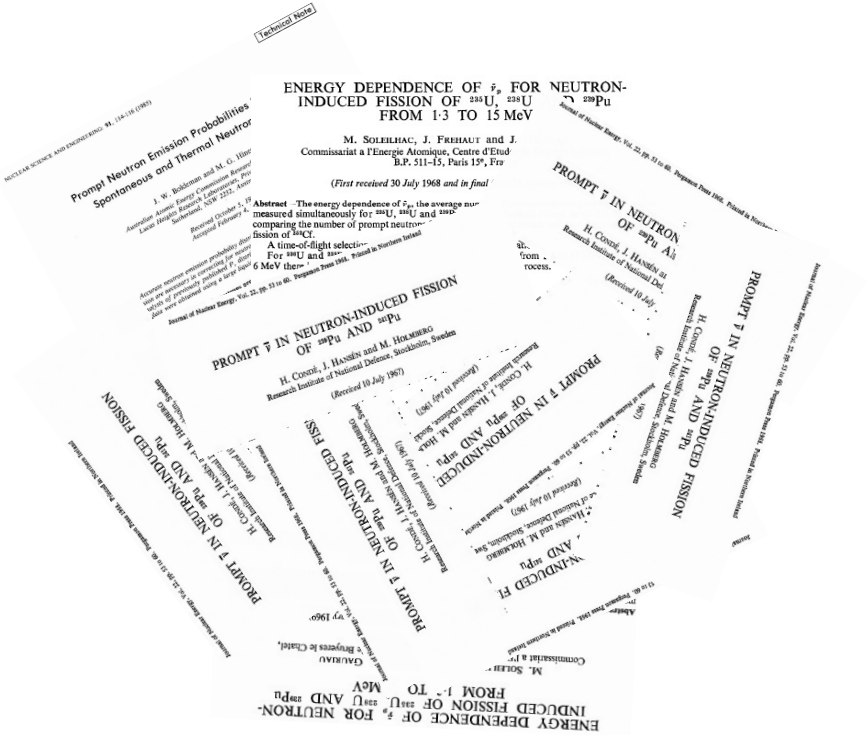
#Legend: 6 x 275 x 12 : data columns * lines * column width

#EN	Energy of incident projectile, laboratory system	MEV	MeV
#E	Energy of outgoing particle, lab. system	MEV	MeV
#DATA	Diff. fiss. neutron multiplicity d/dE(n) #+ 94-PU-239(N,F),,NU/DE,,REL	ARB-UNITS	arbitrary Units
#ERR-T	Total uncertainty (1-Sigma)	ARB-UNITS	arbitrary Units
#MISC-MIN	Lower limit of miscellaneous value	MEV	MeV
#MISC-MAX	Upper limit of miscellaneous value	MEV	MeV

#/Legend

EN	E	DATA	ERR-T	MISC-MIN	MISC-MAX
MEV	MEV	ARB-UNITS	ARB-UNITS	MEV	MEV
.5	.596	72442	5512.35	.571056	.621056

Step 4: reading literature and add missing information



I look for the following information:

- what measurement feature apply,
- what corrections are missing,
- and what uncertainties are missing.

→ I often find information that is missing and very relevant for me but it takes a lot of time to feed back, so I gave up and just put it in my own database.

Step 5 & 6: using information from literature and EXFOR to create my own evaluation (json) database.

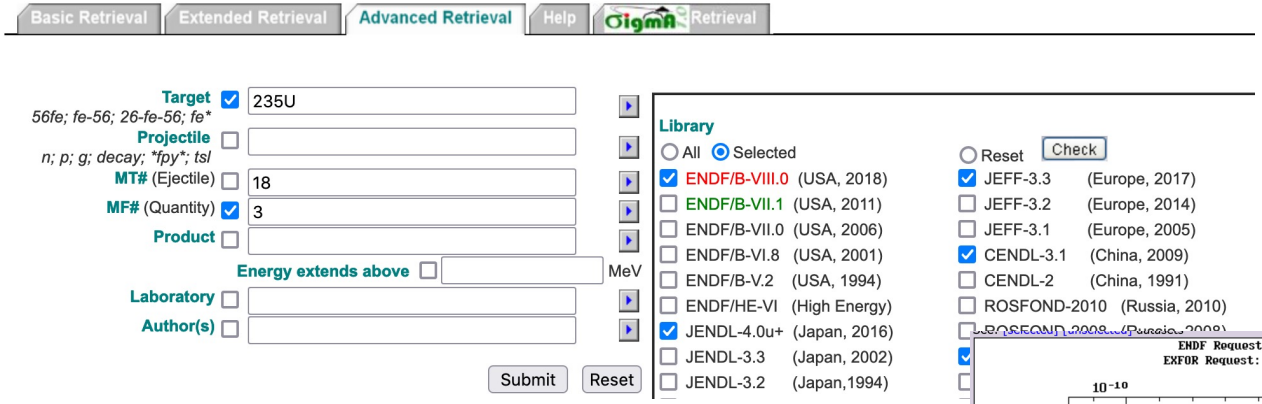
My post-processing steps include:

- Re-normalizing to newest monitor reaction,
- Weeding out outliers,
- Identifying missing uncertainties and corrections,
- Adding missing uncertainties via templates of expected uncertainties.

The final database includes:

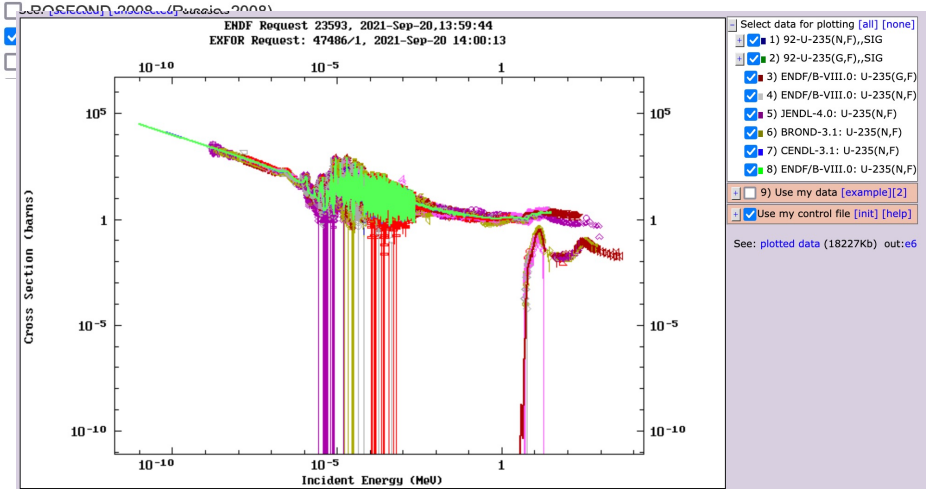
- Data: lattice (energies), data, total covariances (across data-sets),
- Information: EXFOR accession number, references, metadata features (on hardware, missing corrections, samples, etc.).

I use EXFOR for counter-checking whether nuclear data are realistic:



- I use the plotting interface with MF/MT (really nice and easy),

- Use the retrieval tool for EXFOR and plot.



What input from and capabilities of EXFOR makes my life currently easier?

- First of all, the existence of EXFOR is great!!! I see fellow scientists in other fields getting their data out of books …,
- DOIs when available,
- the clear identification of the monitor reaction with monitor values,
- listed corrections,
- all listed uncertainties,
- all listed features,
- everything that points to the dataset being an outlier,
- Plotting tools.

What would really help me if changed or added?

- Searching for specific nuclear data is often non-intuitive (PFNS, nu-bar). → sometimes things are stored in a way that is not easy to find and looking through previous evaluations has helped me pin-point data that I missed but were in EXFOR,
- I sometimes would like to type in MF, MT and get the data related to that,
- DOI/URLs for all cited papers/ report/ etc., would be great!,
- easy way to get to the features,
- common nomenclature of uncertainties,
- common nomenclature of corrections undertaken,
- highlighting which uncertainty sources are missing,
- highlighting which corrections are missing.

What would really help me if changed or added?

- Easy way to get back to EXFOR on what is missing,
- you cannot plot all relevant data (PFNS?, angular distribution?, nu-bar you need to work-around),
- the data for automatic plotting are not re-normalized → having updated monitors would be really helpful,
- weeded out outliers would be really helpful,
- knowing which data were used for previous evaluations would be perfect,
- having realistic uncertainties on data would be great,
- having an easy-to-plot download format would be nice, but maybe I did not try hard enough.

Comments from other users:

- What formats/ tools do they use for evaluations:
  - Focused evaluations: web application (pdf), X4, X4+, c4, -> will be stored for evaluation repository in many cases!
  - Special reaction codes (DA, Pol, Sig, …) frequently used
  - Viktor's web-app for re-normalizing with new monitors and getting total covariances
- For quick-plotting (cs, Eout spectra, ang. Dist.): x4i dedicated EXFOR API
- Data looked at: how complete is data set (A, energy), year published, uncertainties, DOI, etc.
- Needs: better filtering (e.g.,: pre-neutron/ independent/ cumulative FY instead of FY only), DOI/URL for reports, bibtex entry, search inelastic data by residuals, give alternative suggestion of what you can search for instead of -NO DATA FOUND-

What do we currently have:

- A requirement document for the new format,
- A few examples (nu-bar, (n,f) cs, activation, transmission, etc.) of EXFOR data transformed into the new format,
- Do we need anything else as a good starting point?

I am sure that covered only parts of user needs, so:

Discussion:

- What are other user needs?
- Do you need to use EXFOR online and offline?
- What are cyber-security needs at the labs?

First part: What are user needs and other requirements on the database.

5 min	D. Neudecker	Introduction
13 min	D. Neudecker	User needs from a smaller-scale every day user's perspective
12 min	A. Koning	User needs from a larger-scale every day user's perspective
5 min	A. Lewis & DN	What do we already have?
10 min	A. Hayes	Experiences dealing with hierarchical ENSDF data
45 min	Discussion	• What are further user needs? • Do we need offline options? Online enough? • Cyber-security considerations at the hosting labs? • Volunteers for developing interface from EXFOR data to SG-50 database
10 min	Break	

Second part: What database type and format should we use?

15 min	V. Zerkin	X4Lite2 and progress in EXFOR data automatic renormalization system
60 min	Discussion	• Experiences with DBs and which one? • What are the preferences of developers on databases? • Which format? • Final selection of databases.
5 min	Discussion	Next steps

Next steps:

- For big WPEC meeting, we would like to see:
  - Some prototype coding for translating the database,
  - Some more examples that can be used by developers,
  - Discussion on what is really difficult in the requirement document and needs to be re-done,
  - First discussion of a format (containers, datatypes, etc.).
- Next meeting (end of October) will be on layers 2 and 3 (outlier identification, identifying missing uncertainties, etc.) that we know what additional functionality we need for those layers.