WPEC Subgroup Proposal

Arjan Koning
NRG, Petten, The Netherlands

Title

Improvement of accessibility and quality of the EXFOR database

Justification for a subgroup

The fast increase in computer power and the automation of nuclear model codes, cross section plotting software, and checking and processing codes have made the easy accessibility of experimental data more important than it was in the past. In fact, the advance in the aforementioned computational evaluation tools entails that the retrieval of experimental data is now becoming a delaying factor in contemporary data evaluation. A subgroup is proposed that establishes EXFOR as an easy accessible and correct database, with as most important spin-off a computational database that contains the entire EXFOR in tabular format. This enables

- A huge increase in efficiency for the production of evaluated data files.
- Easy and extensive validation of nuclear model codes.
- Continuous feedback from users to the Data Centers now that all data are accessible.

Bringing the EXFOR database to this level can only be accomplished through collaboration of the EXFOR providers, i.e. the maintenance teams at the Data Centers (IAEA-NDS, BNL, NEA, JAEA, CNDC, Obninsk) and EXFOR users such as nuclear physicists and nuclear data evaluators. Both communities are already associated with various WPEC projects, making WPEC the most appropriate platform for this proposal. To constrain the ambitions, the current proposal does *not* address the completeness of EXFOR (i.e. in terms of compiled works), nor a drastic change of the official EXFOR format.
EXFOR (in the USA: CSISRS) is by far the most important and complete experimental nuclear reaction database. Over the past 40 years, experimental data have been added to the database by various compilers who used different (and allowed!) formatting rules to store the data. The result is a database that contains, at least for neutrons, the numerical data of almost the entire history of reaction measurements. However, the main question is whether and how the user has access to all these data. For the study of a few detailed reactions, helpful web interfaces exist, but a genuine step forward in the production of data libraries requires that all nuclear data that exists in EXFOR can be retrieved in an unambiguous manner, and this is lacking at the moment. Moreover, EXFOR is known to contain various errors which, admittedly, in the past have not been reported back sufficiently by the user community.
There are many reasons why the quality of EXFOR is important:

1. Nuclear model codes. Evaluated data libraries are relying more and more on nuclear model calculations which can provide data for energies, reaction channels and nuclides for which no experimental data exist. An easy comparison with all existing experimental data would be extremely helpful for validating these model codes.

2. The evaluation of individual isotopes. To increase the efficiency, it is important that an evaluator has easy access to *all* available data, and that these data are correctly represented in the database.

3. The global understanding of covariance data as a very important ingredient for advanced reactor studies. It is not only required *that* the data are good, but also to know *how* good they are. Once this is properly assessed for key nuclear reactions, on the basis of *all existing* experimental and theoretical information, proposals for new measurements can much better be justified, as their impact on technological applications will be clearer. An important condition is that the *existing* data in EXFOR are retrievable, complete and reliable.

4. Good experimental work deserves to be correctly represented and easy accessible, leading e.g. to an amount of bibliographic references that does the work justice.

These issues alone require that EXFOR is not only a large and complete database, but also a properly validated database (similar to the Evaluated Nuclear Data Files).
Finally, an unambiguously defined good-quality experimental data base fits into a larger and more modern nuclear data framework, in which the entire process of data evaluation and validation is better automated and performed under a more strict quality assurance scheme. Eventually, this could make the evaluation of a data file reproducible, instead of an incremental ad-hoc process.

Subgroup coordinator

—

Subgroup monitor

JEFF: A.J. Koning

List of subgroup participants

- ENDF: Rochman, Herman (BNL), Brown (LLNL)

- JENDL: Fukahori (JAEA)

- JEFF: Plompen (IRMM), Kopecky (JUKO), Forrest (UKAEA), Konobeyeev (FZK), Talou (CEA), Henriksson (NEA)

- BROND: ?

- CENDL: CIAE, Nankai Univ.

- IAEA: Zerkin, Schwerer

Definition of the project and proposed activities

The subgroup addresses the accessibility, quality assessment, and formatting issues of the EXFOR database, from the point of view of a user. The subgroup does *not* propose a new EXFOR format, but instead strives to harmonize the current format, i.e. to store the current contents in a more unambiguous way, ensuring completeness of retrievals by users. The quality assessment, validation, and correction of the EXFOR database can be done on 5 levels. In increasing order of effort these are:

1. Correction of the most obvious errors:

   A. Physical errors, e.g. if barns are given where millibarns should be given. (or MeV instead of eV, etc.). These errors are often directly visible and readily emerge upon trying to process the entire EXFOR library.

   B. Format errors. There are errors in the reaction identification, and in the codes used at the Data Centers to handle the numerical data, e.g. X4toC4. This prevents proper use of the reaction data, and worse, reaction data may appear not to exist.

   C Format harmonization. Some nuclear reactions are stored under two, or even more, different reaction identifiers, again leading to incomplete retrievals.

   Responsible: Data Centers. Input: Users.

2. Completion of the translation of the entire EXFOR database into tabular format. For this [1B] and [1C] needs to be solved.
   Responsible: Data Centers. Test: Users.

3. Correction of more difficult to find, but still obvious dimensional errors, e.g. millibarns instead of barns, through comparison with nuclear model codes or other measurements of the same reaction.
   Responsible: Data Centers. Input: Users.

4. Quality flagging of the data through comparison with other measurements of the same reaction and nuclear model codes.
   Responsible: Users. Input: Users.

5. A review team consisting of experienced experimentalists, who judge each experiment (or at least an important subset of the database) in detail, regarding measurement method used, quality of data analysis, reputation on the authors and lab, etc. and then assign a quality flag.
   Responsible: Users. Input: Users.

Relevance to evaluated data files

If every data evaluator has the same easy access to the entire EXFOR database, there is a decreasing possibility that essential input is overlooked in the evaluation process. Well-automated systems for data file evaluations now exist. A notable example of the power of computer technology is WP21, where automated numerical and graphical comparisons helped to select the best evaluation for a large number of fission products in a relatively easy manner. This means that we are approaching the stage, and some of us have already reached this stage, where retrieval and selection of experimental data is the main delaying factor for a data evaluator.
Successful implementation of the scheme proposed here

- **minimizes the delay with which experimental results are used or adopted in data files.**
- **makes the process of evaluating data files much more efficient.**

Deliverables

- Data Centers: Unified EXFOR database in computational format. Annual releases, in increasing quality and completeness.
- Users: Quality Database: an EXFOR related database with quality flagging.

Time-schedule and milestones

- t_0 + 6 months: Collection and comparison of all software that processes EXFOR (X4toC4, NEA's code, Cullen's code, JANIS), and outline of ideas to merge this.
- t_0 + 6 months: Correction of EXFOR with first lists of errors (Forrest/Kopecky, Koning, etc.)
- t_0 + 12 months: First version of the full computational library. Target: All cross sections.
- t_0 + 12 months: Procedures + format of Quality Database.
- t_0 + 24 months: Second version of the full computational library. Target: All cross sections + angular distributions + differential energy spectra + double-differential energy-angle spectra.

- t_0 + 24 months: First version of the Quality Database.

- t_0 + 36 months: Third version of the full computational library. Target: All quantities of the second version + resonance parameters + ratio's + everything else.

- t_0 + 36 months: Second version of the Quality Database.

Possible long-term milestones:

- t_0 + 48, 60, ... months: Next version of the full computational library.

- t_0 + 48, 60, ... months: Next version the Quality Database.