Summary
Further Development of EXFOR

Consultant's meeting 6 - 9 March 2012,
IAEA Headquarters, Vienna, Austria

version 09 March, end of Meeting

EXFOR was originally (1969) created to be an “exchange format” to share data between different nuclear data centers, with different information systems and missions. By 2005, all data centers had agreed to merge all of their EXFOR libraries into one central master library and project coordination was housed in the Nuclear Data Section at the IAEA. In the mean time, the EXFOR Library has evolved into a mature and comprehensive database of knowledge transfer. Indeed, the EXFOR Library is the “Mother of All Libraries” from which nearly all evaluated nuclear application libraries (particle transport, dosimetry, etc) are derived.

Over the years, many web services have been developed to improve dissemination and data management. Currently EXFOR comprises the Format, the Information Systems (external websites, with ever-expanding retrieval and plotting capabilities, and backend data management), and the Library itself. When we discuss the future of EXFOR, we are discussing the future of all of these components of the EXFOR system and even the management of EXFOR.

The conception of EXFOR lies far in the past and despite the great foresight of its inventors and developers, new demands and impressive software developments in the past decade call for a careful evaluation of the status of EXFOR. Now is an auspicious time to do this because of the accumulated experience gained by the LLNL/BNL collaboration in the redesign and migration the ENDF format to the modern hierarchical GND format. In addition, there are many data projects around the world which would like to expand EXFOR’s capabilities beyond its original design goals to accommodate new kinds of data such as the hypernuclear data currently housed in Hokkaido University’s NRDF database.

We note that the EXFOR system serves many users and many communities. Our core users consist of scientists wishing to archive their data, theoreticians seeking to benchmark their models, experimentalists who seek to compare datasets, compilers who archive the data in EXFOR and evaluators who use the EXFOR data as input for their work. That said, the NRDC has limited resources and cannot compile all forms of data for all users. The compilations are guided by user needs and the extent to which the scope of archived data and format of EXFOR can be adapted to meet those needs.

The goal of this consultants meeting, was to first discuss the EXFOR Format and as a consequence the connection with the EXFOR System and Library. Along the way, we have collected users experience (especially problems) using the EXFOR Format, the EXFOR Information Systems and the EXFOR Library itself and then to present possible plans of action to address these problems. These are collected in the talk summaries and the actual presentations.

From these presentations and our collective discussions, it is our consensus strategy to keep EXFOR exchange format essentially untouched (for now). But, we have gained valuable experience from other format projects (XSAMS, ENDF/GND, ROOT, NRDF, EXFOR/XML). Several projects have proposed and even experimented with XML as an exchange mechanism for hierarchical data. This seems to be a very profitable approach and worthy of future work.

Below, we present the areas of general consensus from members of this meetings and proposals for future action items.

General consensus on extension of the EXFOR Library:

1. *Need to add incident source spectrum for average data*. For example, Maxwellian Average Cross Sections are useful for validating evaluated data, yet the corresponding spectrum is not compiled in the EXFOR Library. An effort to address this need is in progress.

2. *Need to store sufficient experimental data to allow a meaningful reevaluation of experimental data*. The template for the compilation of transmission factors provided by S. Kopecky et al. in the case of cadmium transmission is a good example of what is needed for the reanalysis of data.

3. *Need to collect experimental response and resolution functions.* The resolution is important for allowing an evaluator to match to the experimental data in EXFOR. In the cases of the Lead Slowing Down Spectrometer and measurements of double differential data and time of flight data the resolution function is quite complex. Another good example was shown by K. Kato's presentation in this meeting.

4. *Need to collect experimental covariance data so evaluator can encode all correlations in their evaluations.* We would like to solicit this covariance information from experimentalists. Of course, not all of the information needed to construct a proper covariance is available. So, we need a standard format for representing covariance data and supporting software.

5. *Need to expand EXFOR Library with all of the associated documents for a dataset including author communication, PDFs, the experimentalist's data reduction codes, etc.* This information may not be generally available outside of the EXFOR compilers due to copyright problems. Nevertheless as much as possible should be generally accessible.

General consensus on the EXFOR Format:

1. *The EXFOR Format is old, but it is not necessarily "broken" .* The logical structure of the data hierarchy is well formed. The specific formatting of the data within the hierarchy is not very flexible. However, continued extensions are possible even if

these extensions might be more convenient in another format. Currently, there is no crucial issue forcing a major change of format.

2. *An editor (for compilers) and web interface (for users) insulate people who interact with the EXFOR Library from the EXFOR Format.* This will allow us to modify the underlying structure/format/language of the data without interfering with compilers or users (so long as the software tools are maintained).

3. *There is dissatisfaction with the restraints imposed by the EXFOR REACTION string for complex reactions.* It does not easily enable coincidence data or intermediate particles (e.g. in breakup or sequential reactions). However, no consensus on how to address these shortcomings could be obtained. Work will continue to develop acceptable improvements to the REACTION coding. As an alternative we can compile this data use the existing MSC flag (the EXFOR miscellaneous flag). A poorer alternative would be to not compile this data but keep the links to the original information (e.g. through NSR).

4. *There is dissatisfaction regarding the EXFOR dictionaries*. Although dictionaries are essential for the EXFOR system, there are many dictionaries, they are often cryptic and the size of the dictionaries is growing. In addition, many are obsolete and there is a need for revision and possible simplification. Finally, they are hidden from the user. While most users are not interested in the contents of the dictionaries, they should be generally available.

5. *Explore alternative future formats for EXFOR.* There are many kinds of data that do not neatly "fit" in EXFOR. We will investigate options for alternative format, especially because projects like EXFOR/XML, GND and ROOT offer possible paths forward.

Meeting Recommendations

Although there are many items which we arrived at a consensus regarding either the format or the library itself, we either did not agree on approaches to deal with those items or we are already dealing with them. Below we list proposals for future actions where general consensus was achieved.

Continue development of EXFOR/XML. Use EXFOR/XML as a standard output format since it can be easily transformed into simple to understand HTML and is isomorphic with the legacy EXFOR Format. Improve it based on user feedback and seek to harmonize it with GND. Prepare for the task of porting checking codes and other codes to a new format.

Improve existing EXFOR codes. Develop an editor for EXFOR/XML and improve checking codes and digitization codes. Improve the user experience by changing the default output format of the EXFOR retrieval system.

Establish an "EXFOR knowledge base". This knowledge base would to integrate all the different elements of documentation in one framework, including EXFOR Formats Manual, LEXFOR Manual, EXFOR Basics Manual, Dictionary Manual, the EXFOR dictionaries (in full including explanations), Protocol, Network document and maybe even CP Memos. This knowledge base should aid the compilers but should also be available general users.

Set up an "EXFOR wiki". The wiki would supplement and clarify the EXFOR manual.

New kinds of data. Hypernuclear reaction data are already compiled by Hokkaido U. in the NRDF database and has a large scientific interest. Some high energy p-Nucleus collision data was already compiled by Brookhaven National Laboratory, complete with needed EXFOR Format extensions, and is needed for proton radiography applications. Muon-induced reaction data are not compiled by anyone but muons comprise the dominant component of cosmic rays at sea level and therefore this data is important for understanding backgrounds in a variety of basic science and practical applications. Currently high energy data (E > 1 GeV/A) compilation is voluntary, but we recommend adding already compiled data in EXFOR.

MSC Code in REACTION field. The current REACTION formalism cannot express some classes of experimental quantities (e.g., coincidence measurements, sequential reactions). Though we recognize some of them are useful to validate theoretical models, we recommend keeping such data with an approximate REACTION code with MSC in SF8 unless the quantities are repeatedly reported and there is an obvious need to describe the REACTION with more accuracy.

Governance. For such an important database it is essential that a proper set of rules and goals are laid down that are agreed by the international community and in particular the Nuclear Reaction Data Centres. A strategy document is required that establishes which user groups the library should serve, which services should be maintained and developed. Purpose, scope, quality, validation, interfaces and user-orientation are important keywords to be addressed in such a document. The EXFOR Protocol document deals with many of these issue, but is not highly advertised.

User feedback. Feedback from users is now a weak point and a method for a systematic approach to establishing users interest and needs should be established. Through an EXFOR user group that periodically reviews the services that are offered providing recommendations for improvements and new directions. Such a user group could meet at the upcoming ND2013 meeting where many EXFOR users will be in attendance.