Progress in international collaboration on EXFOR library

N. Otuka, B. Pritchenko, M. Fleming, Y. Jin, G. Pikulina, R. Suzuki, V. Devi, M. Mikhailiukova, S. Okumura, N. Soppera, T. Tada, S. Takács, S. Taova, V.V. Varlamov, J.M. Wang, S.C. Yang, and V. Zerkin

1 Nuclear Data Section (NDS), International Atomic Energy Agency, A-1400 Wien, Austria
2 National Nuclear Data Center (NNDC), Brookhaven National Laboratory, Upton, NY 11973-5000, USA
3 OECD Nuclear Energy Agency Data Bank (NEA DB), F-92100 Boulogne-Billancourt, France
4 China Nuclear Data Centre (CNDC), China Institute of Atomic Energy, Beijing 102413, China
5 Centre of Nuclear Physics Data (CNPD), All-Russian Research Institute of Experimental Physics, 607188 Sarov, Russia
6 Department of Medical Physics, Hokkaido University Hospital, Sapporo 060-8648, Japan
7 Department of Applied Sciences, Institute of Engineering and Technology Bhaddal, Ropar 140108, Punjab, India
8 Nuclear Data Centre (CJD), Institute for Physics and Power Engineering, 249033 Obninsk, Russia
9 Nuclear Reaction Data Centre (JCRG), Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan
10 Laboratory of Cyclotron Applications, Institute for Nuclear Research (ATOMKI), H-4026 Debrecen, Hungary
11 Centre for Photonuclear Experiments Data (CDFE), Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, 119991 Moscow, Russia
12 Korea Nuclear Data Center (KNDC), Korea Atomic Energy Research Institute, Daejeon 305-600, Republic of Korea

Abstract. The EXFOR library has served as the unique repository of experimental cross section and other nuclear reaction data for 50 years. The Nuclear Reaction Data Centres (NRDC) have compiled data sets from more than 22000 experimental works for the EXFOR library. Our collaboration and effort on improvement of EXFOR coverage are described in this paper, as well as tools for digitization of numerical data from graph images developed by us for EXFOR compilation.

1 Introduction

The EXFOR library [1] has served as the unique repository of experimental nuclear reaction data for many decades. Data exchange of the neutron-induced reaction data (neutron data) in the EXFOR format was started in July 1970 [2], following the agreement among the Four Neutron Centres in 1969 [3]. Currently, the EXFOR library contains neutron data, charged-particle induced reaction data, charged-particle data, and photon induced reaction data (photonuclear data) measured in more than 22,000 experimental works. It is maintained and developed by the 13 data centres belonging to the International Network of Nuclear Reaction Data Centres (NRDC) [4], under the coordination of the IAEA Nuclear Data Section.

This article summarizes the status and progress in our collaboration since our report to the last Nuclear Data conference (ND2016) [5]. We also introduce digitization tools developed by data centres for extraction of experimental data plotted on graph images.

2 Status of collaboration

The NRDC Protocol [6] defines the work shared among us such as compilation scope, compilation responsibility, procedures of dictionary and manual updates, data transmission, and code development. The EXFOR scope has remained unchanged: Data for neutron and light-ion ($A \leq 12$) beams below 1 GeV belong to the category of compulsory compilation. The highest priority is given to compilation of recently published data in this category. The NDS regularly scans about 40 journals to identify articles with data suitable for EXFOR and to monitor progress in their compilation. All new and old articles are listed on the Article Allocation List [7] until they are compiled.

Table 1 summarizes the compilation responsibility of each data centre, which ensures that all data in the category of compulsory compilation are compiled by us.

The 13 centres are complemented by two compilation groups in Kazakhstan and Mongolia. The Institute of Nuclear Physics (Almaty, Kazakhstan) compiles data measured in Central Asian countries in collaboration with Institute of Nuclear Physics (Tashkent, Uzbekistan), and the National University of Mongolia (Ulaanbaatar, Mongolia) compiles heavy-ion ($A > 12$) induced reaction data measured in West European countries, which are not covered by any centre. Data for heavy-ion ($A > 12$) and photon beams belong to the category of voluntary compilation, although some data centres routinely compile these data. For example, JCRG and NNDC are systematically compiling data measured at powerful radioactive ion beam facilities to archive the results of the frontier of basic nu-
Table 1. Nuclear Reaction Data Centres (NRDC). ND: Neutron induced reaction data; CPND: Charged-particle induced nuclear reaction data; PhND: Photonuclear data

| Centre     | Location                                      | Scope of compilation                  |
|------------|-----------------------------------------------|---------------------------------------|
| ATOMKI     | Institute for Nuclear Research, Debrecen, Hungary | CPND measured by ATOMKI               |
| CDFE       | Lomonosov Moscow State University, Moscow, Russia | Photonic data                         |
| CJD        | Institute of Physics and Power Engineering, Obninsk, Russia | Neutron data from former USSR*         |
| CNDC       | China Institute of Atomic Energy, Beijing, China | ND and CPND from China                |
| CNPD       | All Russian Scientific Research Institute of Experimental Physics, Sarov, Russia | CPND from former USSR*                |
| JAEA/NDC   | Japan Atomic Energy Agency, Ibaraki, Japan     | (Dissemination of the JENDL library)* |
| JCPRG      | Hokkaido University, Sapporo, Japan           | CPND and PhND from Japan              |
| KNDC       | Korea Atomic Energy Research Institute, Daejeon, Republic of Korea | ND, CPND and PhND from Korea           |
| NDPCI      | Bhabha Atomic Research Centre, Mumbai, India  | ND, CPND and PhND from India           |
| NDS        | International Atomic Energy Agency, Vienna, Austria | ND and CPND from “rest of the world”‡ |
| NEA DB     | OECD Nuclear Energy Agency, Boulogne-Billancourt, France | ND and CPND‡ from Data Bank Countries  |
| NNDC       | Brookhaven National Laboratory, Upton, NY, USA | ND, CPND and PhND from USA and Canada  |
| UkrNDC     | Institute for Nuclear Research, Kyiv, Ukraine | ND, CPND and PhND from Ukraine         |

*Except for Ukraine. †Compilation of ND from Japan under planning. ‡Except for CPND from Japan and heavy-ion (A > 12) induced reaction data.

3.1 EXFOR-NSR comparison

The Nuclear Science Reference (NSR) [11] is a bibliography of nuclear physics articles, indexed according to content, which spans more than 100 years of research. Its coverage is wider than EXFOR, and one can expect to find more experimental nuclear reaction works in NSR than EXFOR. Both NSR and EXFOR entries have been im-

3 Retroactive compilation

The NDS started regular journal scanning for experimental data of reactions induced by all projectiles (neutron, light- and heavy-ions, photon) in 2004. Before this, CINDA [10] served as the main list of articles for EXFOR compilation. However, CINDA covered only neutron data and there was no mechanism to ensure completeness of charged-particle and photonuclear data. Additionally, some neutron data (e.g., fission product yields) were not prioritised by all centres in past decades. To improve the EXFOR coverage, the NRDC is performing retroactive compilation in parallel with compilation of newly published data. Three examples of our retroactive compilation projects are introduced in the following subsections.

3.2 Incident neutron data / EXFOR / Pu239 (F)

Data 23912.05, 1996. K. Nambu [8] published DATA
Figure 2. Cumulative number of EXFOR entries, where one entry is one experimental work.

Figure 3. Cumulative number of EXFOR entries by compiling centre.

Table 2. Experimental nuclear reaction articles compiled in NSR and missing in EXFOR. “for EXFOR” means articles for addition in EXFOR. “EXFOR entries” gives the number of existing EXFOR entries.

| Projectile | Articles in NSR | EXFOR entries |
|------------|----------------|---------------|
| Neutron    | 142            | 10558         |
| Proton     | 371            | 4523          |
| Alpha      | 1179           | 1797          |

Table 3. Cumulative number of EXFOR entries.

3.1 EXFOR-NSR comparison

Cumulative number of EXFOR entries, where one entry is one experimental work.

Figure 4. Newly compiled experimental spectra of neutrons emitted following β-decay of $^{87}$Br [16, 17].

3.2 Delayed neutron multiplicities and spectra

Delayed neutron data of a specific precursor are decay properties of the precursor nuclide and do not depend on how the precursor was formed. Due to this reason, the EXFOR scope of delayed neutron quantities was limited to the total and group-wise $\bar{\nu}_d$, delayed fission neutron spectrum of a given neutron group, and delayed neutron emission probability ($P_\nu$), while the spectrum of neutrons emitted by a specific precursor was excluded. However, the IAEA Coordinated Research Project on β-delayed neutron emission (2013-2018) [13] realized that the ENSDF format cannot accommodate continuous spectra and requested the NRDC to extend the EXFOR scope. To satisfy their needs, the NDS checked the experimental works cited in review articles of $P_\nu$ by Rudstam [14] and delayed neutron spectra by Kratz [15], and found 24 and 10 articles for EXFOR compilation, respectively. Figure 4 shows spectra of neutron emitted following β-decay of $^{87}$Br newly compiled from two articles [16, 17].

3.3 Fission product yields

The Four Neutron Centres discussed compilation of fission product yields in 1975, and concluded that they did not have sufficient resources to devote to the compilation of this data [18], while the US and UK pioneers of the fission product yield evaluation developed their own compilation in 1970s [19–22]. In order to establish a common experimental database, EXFOR compilation of fission product yields was actively discussed in two IAEA meetings [23, 24] and the IAEA Coordinated Research Project on “Compilation and evaluation of fission yield nuclear data” (1991-1996) [25]. Since then, fission product yields have been compiled more regularly, and literature scanning by the NDS (2004-) ensures that EXFOR is mostly complete for fission product yields published in or after 2004. On the other hand, it was not clear what fraction of the data compiled and used by the US and UK evaluators is covered in EXFOR. CCDN (now NEA DB) compared Crouch’s 1977 compilation [21] with EXFOR and CINDA in 1980 [26], and the NNDC performed conversion of Meek & Rider file [19, 20] to EXFOR in 1980s and
Table 3. Fission product yield articles missing in EXFOR. The numbers from the NNDC assessment (EXFOR v.s. NSR) are for creation of, or addition to, EXFOR entries while the numbers from the NDS assessment (EXFOR v.s. ENDF/UKFY) are exclusively for creation of new EXFOR entries. The areas 1-4 correspond to USA+Canada, Western Europe + Japan, "rest of the world" and former USSR countries, respectively.

| Exp. site | NNDC [27] | NDS [29] | ENDF | UKFY |
|----------|-----------|----------|------|------|
|          | n,f       | spon     | γ,f  |       |
| Area 1   | 35        | 49       | 8    | 92   |
| Area 2   | 124       | 28       | 32   | 184  |
| Area 3   | 26        | 16       | 7    | 49   |
| Area 4   | 49        | 21       | 41   | 111  |
| Total    | 234       | 114      | 88   | 436  |

1990s in collaboration with NDS and CNDC following a request of the Studsvik meeting [23]. However, a number of experimental fission product yield articles potentially missing in EXFOR were extracted from NSR. All data with identical values in an EXFOR entry were removed. All remaining datasets were compared with all EXFOR data within a 10-year window to identify datasets that were linked (e.g. preliminary and final results). NNDC assessed these articles individually and prepared a list of articles for EXFOR compilation [28], with data for new compilations or addition to existing entries.

The NDS approach [29] was comparison of EXFOR with the citation lists of the ENDF and UKFY fission product yield evaluation summaries [30, 31]. Based on this assessment, the NDS prepared a list of articles for creation of new EXFOR entries [32].

The result of these assessments was discussed by the fission product yield experimentalists and evaluators in a recent IAEA meeting [33], and our effort on compilation was strongly welcomed by the experts. Table 3 summarizes the numbers of fission product yield articles missing in EXFOR identified by NNDC and NDS.

4 Digitization tools

Collection of original data from experimentalists is the primary role of the NRDC. Experimentalists are asked by journal editors to limit the number of figures and data tables, reducing the published data to a small fraction of the valuable output from their work. EXFOR is designed to be a complete data library and such unpublished data are routinely included through direct engagement with experimentalists. Unfortunately, a considerable fraction of the numerical data are not available from the experimentalists, even if they have been published, and in many cases it is necessary to use whatever published materials are available. This is a very typical situation when compiling differential cross sections of charged-particle reactions, which were published when the NRDC was not uniformly prioritising charged-particle data compilation. As a consequence, about 40% of EXFOR entries include data read (digitized) from graph images [34], and the NRDC still routinely performs figure digitization (even for new articles, where all attempts to contact authors are unsuccessful).

In order to assure the accuracy in our digitization, the IAEA NDS organized a benchmark of digitization performance in 2005 [35]. All centres were requested to digitize angular differential cross sections published in 2000, and five data centres (CAJaD, CDFE, CNPD, JCPRG, NNDC) submitted their digitization results in a blind intercomparison with the numerical data received from the experimentalist. This was followed with a workshop organized by the NDS to discuss good practice for good digitization, which drafted recommendations for EXFOR compilers [34]. These benchmark tests show that data digitized from symbols on a linear scale are often equal to the original data when the quality of the graph image is high, while the accuracy cannot be better than 1% when symbols are on a logarithmic scale.

Three data centres (CNDC, CNPD, JCPRG) are developing and maintaining their own digitization tools (GDGraph, ImpGraph and GSYS), which are freely available on the internet. The other NRDC data centres make regular use of these tools and contribute to their improvement by providing suggestions to the developers. These three digitization tools are briefly introduced in the following subsections.

4.1 GDGraph

GDGraph [36] has been developed by the CNDC to meet the requirements from evaluators, experimentalists and compilers in China. The first version of GDGraph released in 2000 was written in VC++. The program was re-written in Perl and its second version was released in 2006. Since then, Versions 3.0, 4.0, 5.0 and 5.1 were released in 2011, 2012, 2013 and 2016, respectively. A friendly interface has been implemented by use of WxWidgets as the graphical user interface toolkit.

GDGraph has the following features:

- All information for digitization of a graph image can be saved as a "project file", and users can resume or check the digitization work later by loading it.
- Clipboards can be used when loading a graph image or generating a numerical data.
- The graph image can be rotated by setting any angle.
- Various options are available for sizes, shapes and colors of the marker (Figure 5).
- Cursor keys can be used when they are more convenient than pointing devices (e.g., mouse).
- Combination of the magnifying glass function and cursor keys supports digitization with high accuracy (Figure 6).
Users can reuse the former digitizing data or compare with other data easily by using an import function.

![Figure 5. Digitization from solid circle symbols on GDGraph. The size of markers is adjusted to the size of the symbols printed on the graph image.](image5)

![Figure 6. Fine adjustment by combining the magnifying glass and cursor keys on GDGraph.](image6)

4.2 InpGraph

InpGraph has been developed by CNPD as a part of the EXFOR-Editor software package [37]. Its first “official” release was in 2001, and the version with the modernized interface was released in 2014. It is designed in order that the user can easily achieve a variety of goals with the aid of a comprehensive graphical interface (Figure 7). InpGraph implements various mathematical models for extraction of data from a low quality image (e.g., a graph image where the x-axis and y-axis are not orthogonal to each other).

Two types of digitizing errors estimated by InpGraph (“systematic error” and “quantization error”) quantify the digitization accuracy so that compilers may input these data into their EXFOR compilations. The systematic error is estimated from the deviation of the digitized values from the true values on ticks on an axis. For example, the systematic error in the digitized x-values $\delta_x$ is estimated by digitizing n-ticks on the x-axis and calculating

$$\delta_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x_0)^2}{n-1}},$$  \hspace{1cm} (1)$$

where $x_i$ is the value of the $i$-th tick on the x-axis reported by InpGraph and $x_0$ is the corresponding true value. The quantization error corresponds to the half-width of the pixel. The systematic error becomes larger than the quantization error when the quality of the graph image is poor.

The spreadsheet (DataTable) mode (Figure 8) allows EXFOR compilers to generate digitized data in the EXFOR format. Within DataTable mode, InpGraph has implemented the following functions:

- numerical data input and editing;
- numeric data precision setting;
- manipulation with table rows and columns;
- various calculations with arithmetic or build-in formulae;
- data line sorting;
- digitized data plotting for visual inspection;
- digitized data export to and import from various files (text, MSWord, MSExcel).

4.3 GSYS

GSYS (Graph Siuichi Yomitori System) [38] is a platform-independent Java application tool, which was originally developed by JCPRG for compilation for the Nuclear Reaction Data File (NRDF) [39]. Various new functions have been added since release of its first version in 2005 [40], and the current version (Version 2.4.7, released in 2014) is used not only for NRDF compilation but also for EXFOR compilation.

GSYS offers “automatic axis recognition”, “automatic point recognition” and a “feedback function” for better digitization accuracy.

**Automatic axis recognition:** GSYS recognizes the position of the axis and ticks on the graph image when the user drags the mouse pointer to the area including the axis. The accuracy of digitization strongly depends on the accuracy
The DataTable window of InpGraph, showing the main data table manipulation functions.

Automatic point recognition: GSYS recognizes the centre of a symbol printed on the graph image when the user clicks somewhere near the symbol (Figure 9). With this function, the user can catch each symbol more efficiently, without sacrificing accuracy.

Feedback function: This function allows graphical comparison of a graph image with a numerical data set. By using this function, any data can be compared with the graph image on GSYS. This is a useful data validation procedure of digitization as long as the digitization result does not depend on various models (assumptions) for correction of any distortions.

The feedback function also can be used to identify a source article when the source of the numerical data set is not clear. An recent example is identification of a source article of the $^{244}$Cm(n,tot) cross sections received by NDS in 1976 and compiled in the EXFOR-VIEN file V0006.002 [41]. We found by using GSYS that this EXFOR data set reproduces a data set plotted on an article published by the same author in 1978 [42] except for the energy above 1 MeV (Figure 10), and we concluded the 1978 journal publication can be used as a reference of the EXFOR data set. This example demonstrates how digitization tools can be used not only for extraction of data on a graph image, but also for various analyses.

5 Summary

The status of NRDC collaboration and our recent effort for improvement of the EXFOR coverage were presented. Our goal is to make all experimental nuclear reaction data accessible, and the NRDC collaborates to increase the completeness and quality of the EXFOR library. Two EXFOR completeness assessments were discussed, including a comparison between EXFOR and the NSR, and a joint NNDC/IAEA effort to identify missing fission product yield data referenced in well-known compilations. These have revealed many nuclear reaction experimental works that are still missing in EXFOR. Digitization tools developed by NRDC members and routinely used for EXFOR compilation were introduced. These are continually improved, with highlighted features described in this paper, and have been used to provide approximately 40% of the EXFOR entries with numerical data. All of these tools are freely available online.

We are most grateful to the experimentalists who engage with compilers to provide complete data for EXFOR and increase its value for the user community. We also would like to express our thanks to V. Zerkin (NDS) and N. Soppera (NEA DB) for their maintenance and development of EXFOR tools.

References

[1] N. Otuka et al., Nucl. Data Sheets 120, 272 (2014)
[2] N.E. Holden, Report BNL-75288-2005-IR, Brookhaven National Laboratory (2005)
[3] A. Lorenz (ed.), Report INDC(NDU)-16, IAEA (1969)
[4] N. Otuka (ed.), Report INDC(NDS)-401 Rev.6, IAEA (2014); http://www-nds.iaea.org/ndr/
[5] V. Semkova et al., EPJ Web Conf. 146, 07003 (2017)
[6] N. Otuka (ed.), Report IAEA-NDS-215 Rev.2018/06, IAEA (2018)
[7] NRDC Article Allocation List: http://www-nds.iaea.org/ndr/aloc/
[8] N. Soppera et al., EPJ Web Conf. 146, 07006 (2017)
[9] V. Zerkin, NRDC working paper WP2005-32 (2005)
[10] OECD Nuclear Energy Agency, CINDA Archive 2006: The Comprehensive Index of Nuclear Reaction Data Archive 1935–2006 (OECD Nuclear Energy Agency, Paris, 2007)
[11] B. Pritychenko et al., Nucl. Instrum. Meth. A 640, 213 (2011)
[12] V. Zerkin, NRDC working paper WP2010-22 (2010); http://www-nds.iaea.org/cinda/
[13] P. Dimitriou et al., these proceedings
[14] C. Rudstam, Report INDC(NDS)-107, p.69, IAEA (1979)
[15] K.-L. Kratz, Report INDC(NDS)-107, p.103, IAEA (1979)
[16] K.-L. Kratz et al., Nucl. Phys. A317, 335 (1979)
[17] R.C. Greenwood et al., Nucl. Sci. Eng. 126, 324 (1997)
[18] G. Thompson (ed.), INDC(NDS)-68, IAEA (1975)
[19] M.E. Meek and B.F. Rider, Report NEDO-12154 (1972), NEDO-12154-1 (1974) and NEDO-12154-2E (1978), General Electric Co.
[20] B.F. Rider, Report NEDO-12154-3B (1980) and NEDO-12154-3C (1981), General Electric Co.
[21] E.A.C. Crouch, At. Data Nucl. Data Tables 19, 417 (1977)
[22] M.F. James et al., Prog. Nucl. Energy 26, 1 (1991); Report AEA-TRS-1019 (1991), AEA Technology Winfrith
[23] M. Lammer (ed.), Report INDC(NDS)-208, IAEA (1988)
[24] M. Lammer (ed.), Report INDC(NDS)-261, IAEA (1991)
[25] International Atomic Energy Agency, Report IAEA-TECDOC-1168, IAEA (2000)
[26] P.D. Johnston, NRDC Memo 4C-2/115 (1980)
[27] B. Pritychenko, O. Schwerer, J. Totans, V. Zerkin, these proceedings
[28] O. Schwerer and B. Pritychenko, NRDC working paper WP2019-19 (2019)
[29] T. Fukuda, S. Okumura, N. Otuka, these proceedings
[30] T.R. England and B.F. Rider, Report LA-UR-94-3106, Los Alamos National Laboratory (1994)
[31] R.W. Mills, Fission product yield evaluation, PhD Thesis submitted to University of Birmingham (1995)
[32] S. Okumura, NRDC working paper WP2019-20 (2019). See also Appendix A of Ref. [33]
[33] M. Fleming, T. Kawano and N. Otuka (eds.), Report INDC(NDS)-793, IAEA (2019)
[34] N. Otuka and V. Semkova (eds.), Report INDC(NDS)-629, IAEA (2013)
[35] N. Otuka and S.Dunaeva, Report INDC(NDS)-629, p.63, IAEA (2013)
[36] Jin Yongli and Chen Guochang, Report IAEA-NDS-216 (Rev.2016/10), IAEA (2016); http://www.nuclear.csdb.cn/gdgraph/
[37] G.N. Pikulina and S.M. Taova, Phys. Atom. Nucl. 81, 1450 (2018)
[38] R. Suzuki, Report INDC(NDS)-629, p.19, IAEA (2013); https://www.jcprg.org/gsys/
[39] M. Togasi and H. Tanaka, J. Info. Sci. 4, 213 (1982)
[40] K. Arai et al., NRDF Annual Report No.18, p.78, Japan Charged-Particle Nuclear Reaction Data Group (2005)
[41] K. Okamoto, O. Schwerer and H.D. Lemmel, Report IAEA-NDS-34 (Rev.2), IAEA (1984)
[42] H. Fernández Gianotti, Nucl. Sci. Eng. 65, 164 (1978)