**NATIONAL NUCLEAR DATA CENTER**

**Bldg. 817**

**Brookhaven National Laboratory**

**P. O. Box 5000**

**Upton, NY 11973-5000 U.S.A.**

Telephone: (631) 344-5091

FAX: (631) 344-2806

**Memo CP-C/478**

**Date:**  April 9, 2020

**To:**  Distribution

**From:**  B. Pritychenko

**Subject:**  **Towards More Human Readable EXFOR**

 Historically, the Exchange Format (EXFOR) has been created as a four major experimental nuclear reaction data compilation centers interchange format, and it became operational in 1970. The major centers had their own data formats such developed in 1964 Sigma Center Information Storage and Retrieval System (SCISRS) by the National Nuclear Data Center (NNDC) at Brookhaven National Laboratory [1]. The new interchange format was designed for flexibility and the diverse needs of the data centers, it contained all major reported quantities and enabled computerized processing of the data. It quickly gained worldwide popularity, and it was adopted in nuclear reaction data compilations. The EXFOR format was revolutionary for its time, it allowed to combine multiple publications into a single entry that consists of bibliographical and data subentries; these subentries include extensive data sets and provide complete description of the experimental methods and techniques. Its data structures consisted of keywords, codes, data tables and definitions, and each subentry is defined by ‘begin/end’ tags (BIB/ENDBIB, DATA/ENDDATA). These tags and heavy reliance on numerical data column headers (ELEMENT, EN, DATA, DATA-ERR, …) made EXFOR-formatted nuclear reaction data compilations easily readable by the early computers that relied on a FORTRAN language.

 Over its 50-year history nuclear reaction data compilations project has produced a unique collection of experimental nuclear reaction, and the development of modern computer and Web dissemination technologies helped to make it accessible to many new user groups. These user groups span over the fundamental and applied sciences and engineering. At the same time, data center’s interaction experience shows that some users are unaware of the EXFOR and LEXFOR manuals, historic rules and often have no time or intentions to read the original publications. This is the time that we are living in, and the user’s approach would not change. Therefore, we must find the new avenues for engagement with these users, and the development of computerized EXFOR formats (C4, C5) and X4+ [1] were important steps. These computerized formats extract tabulated data from the underlying EXFOR format that is human-readable for experienced users but, regretably, confusing to regular users.

 In the light of this disclosure, we may consider a few minor changes in EXFOR compilation rules:

--- First, the current rules do not require a strict order for the bibliographical subentry. While many compilations are based on research articles that have a very strict **TITLE/ AUTHOR/ AFFILIATION/ REFERENCE** sequence. This sequence is commonly accepted by everyone, so it would make sense to adopt it in EXFOR for new compilations. In fact, the Sarov EXFOR editor [2] already produces compilations in such way.

--- Second, the publication world has evolved over the years, and nowadays every article from a major publisher is uniquely linked with a digital object identifier (doi). **Unfortunately, there is no requirement in EXFOR to include doi links into compilations.** This must change.

---Third, EXFOR contains **multi-column tables that can include up to 18 columns**. Unfortunately, only six columns could be shown in a single row, and the remaining columns are wrapping to the next line as an example below:



The wrapping data lines are not a problem for computers, but it is a huge issue for human eyes, and the wrapping effect makes EXFOR tables not human readable. There is no easy solution in this case except splitting data into smaller subentries (separate data sets) when possible and effectively reduce the number of columns.

--- Fourth, element (atomic), mass and isomer numbers in data tables were absolutely essential for FORTRAN programs. It is no secret that many users straggle to covert their numerical values into chemical symbols, i.e. very few users would immediately identify Z=51 with antimony (Sb), and we should capitalize on modern computer technologies and deliver data in a user comprehensive form. In addition, **ELEMENT/MASS/ISOMER** columns could be avoided in data tables.

---Fifth, the present-day rules of EXFOR compilation for sub-field 8 (SF8) such as **AG** (times isotopic abundance and statistical weight factor), **DAM** (Divided by atomic mass of the target nucleus) or **FRC** (Fractional). These rules have been developed to describe the original data ``as they were published” with very elegant reaction strings. Unfortunately, the majority of EXFOR users is not familiar with the EXFOR compilation rules and confused by nine different sub-fields; they often ignore SF8=AG or DAM. It is highly unlikely that users would change their behavior, and it is the NRDC network problem. The possible suggestion here to accept less elegant, simplified reaction strings that would more intuitive to physicists. These more structurally complex but self-explanatory reaction strings would make EXFOR formatted data more readable in their original stand-alone form.

In conclusion, the EXFOR format was introduced fifty years ago as a data interchange format between the four major centers for computers. In subsequent years it became the default compilation format for experimental data, and the corresponding database is used worldwide by more and less sophisticated users. We should make EXFOR compilations more user friendly and self-explanatory for nuclear physicists, engineers and application developers.

**References**

## V.V. Zerkin, B. Pritychenko`` The experimental nuclear reaction data (EXFOR): Extended computer database and Web retrieval system,” [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](https://www.sciencedirect.com/science/journal/01689002) **88**, 31-43 (2018).

1. G.N. Pikulina, S.M. Taova, ``Activities of the RFNC-VNIIEF Center of Nuclear Physics Data on the Compilation of Experimental Data for the EXFOR International Library: EXFOR-EDITOR Software Package,” Physics of Atomic Nuclei **81**, 1450 (2018).

**Distribution:**

a.koning@iaea.org

abhihere@gmail.com

aloks279@gmail.com

cgc@ciae.ac.cn

dbrown@bnl.gov

draj@barc.gov.in

dvoytenkov@ippe.ru

ebata@nucl.sci.hokudai.ac.jp

fukahori.tokio@jaea.go.jp

ganesan555@gmail.com

gezg@ciae.ac.cn

iwamoto.osamu@jaea.go.jp

j.c.sublet@iaea.org

jhchang@kaeri.re.kr

jmwang@ciae.ac.cn

kaltchenko@kinr.kiev.ua

jim.gulliford@oecd.org

manuel.bossant@oecd.org

marema08@gmail.com

masaaki@nucl.sci.hokudai.ac.jp

mmarina@ippe.ru

mwherman@bnl.gov

nicolas.soppera@oecd.org

n.otsuka@iaea.org

nrdc@jcprg.org

nurzat.kenzhebaev@gmail.com

ogritzay@kinr.kiev.ua

oscar.cabellos@oecd.org

otto.schwerer@aon.at

pikulina@expd.vniief.ru

pritychenko@bnl.gov

samaev@obninsk.ru

sbabykina@yandex.ru

scyang@kaeri.re.kr

selyankina@expd.vniief.ru

sonzogni@bnl.gov

stakacs@atomki.hu

stanislav.hlavac@savba.sk

s.a.dunaeva@yandex.ru

sv.dunaeva@gmail.com

taova@expd.vniief.ru

tarkanyi@atomki.hu

vvvarlamov@gmail.com

v.zerkin@iaea.org

yolee@kaeri.re.kr

zholdybayev@inp.kz

zhuangyx@ciae.ac.cn