# Compilation Databases

*Boris Pritychenko*

*National Nuclear Data Center, BNL, Upton, NY 11973*

**BROOKHAVEN**
NATIONAL LABORATORY

*a passion for discovery*

U.S. DEPARTMENT OF **ENERGY** | Office of Science

# Compilation Databases

- There are three major and two minor compilation databases at NNDC: NSR, EXFOR, XUNDL and B(E2)↑, ββ-decay.

- We will concentrate on NSR and EXFOR databases:

  - Nuclear Science References (NSR): all low- and intermediate-energy references for a broad use, not just nuclear structure and decay as before 90ies.

  - Experimental Nuclear Reaction Data (EXFOR): all low- and intermediate-energy reaction data sets for neutron-, charged- and photo-induced reactions, not just neutron-induced as before 80ies.

- The compilation scope and quality controls for NSR and EXFOR database have evolved over the many years of operation. These facts plus lack of advanced computer tools in the past are responsible for missing references and data sets.

# Compilation Operations

- Compilations represent a middle layer between research operations and evaluated data.

- Compilation databases are fundamental for nuclear data evaluation activities.

- Compilations of experimental results are done once while evaluations are revisited on 10-12 years basis. QA requirements are strict.

- Compilation activities rely on significant contributions of contractors.
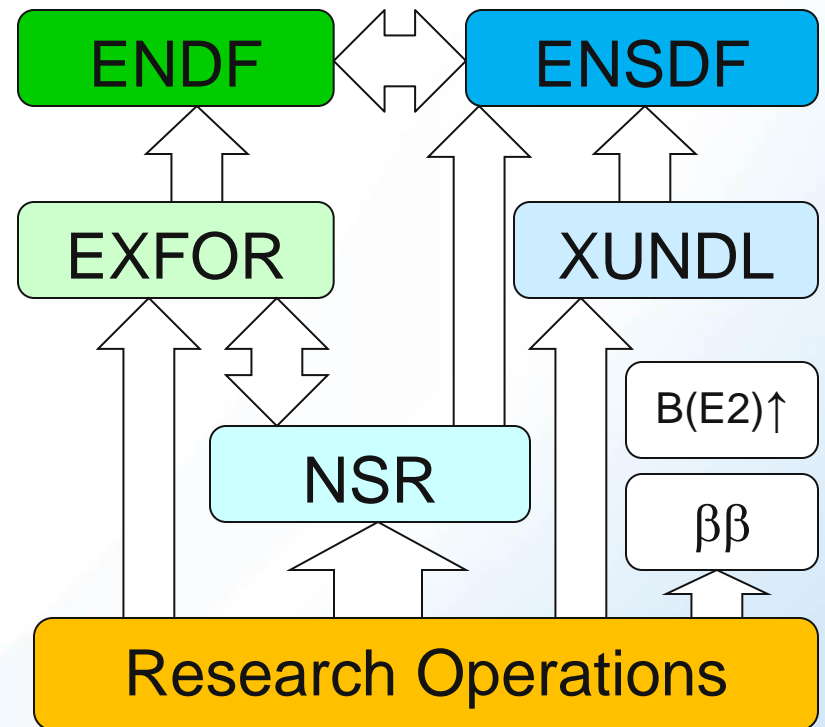
# NSR Compilations in FY2016

- NSR team: 1.5 NNDC, 2 contractors and 1 IAEA collaborator.

- Our goal is provide the coverage for current publications; however, due to many historical and technical reasons a substantial number of articles was missed in the past and we are we are proactively recovering these references.

- Our major requirement is speed, prompt creation of entries for ENSDF.

- NSR Quality Assurance: Manager + Users + Evaluators + Compilers inputs. We do not have a bug database, we just fix bugs immediately.

- Direct communication with Phys. Rev. C: ~15% of authors submit keywords to NSR.

- NSR database is updated 2-3 times a week (Most frequently-updated nuclear database).

# FY 2016 NSR Statistics

- NSR References:
  - 3263 new article entries, total NSR: 222684 (10 times bigger than EXFOR)
  - 859 modified (bug fixes) article entries
  - 1856 keyworded article abstracts
- NSR Dictionary updates:
  - 1750 new authors, total NSR: 96685
  - 7 new journals, total NSR: 516
  - 225 new reactions, total NSR: 7904
  - 408 new nuclides, total NSR: 6415
- NSR Database updates: 117 in FY 2016
- NSR Web retrievals: 442175

# Oak Ridge Library for NSR Users

- NSR like ENSDF has been started at Oak Ridge National laboratory and transferred to Brookhaven in 1980.

- Oak Ridge library had many unique and valuable publications that of importance for ENSDF evaluations.

- In May of 2016, NSR technical staff member Joann Totans travelled to Oak Ridge and recovered unique publications: 35 boxes of reports, private communications and theses that have been published prior to 1980. Oak Ridge employee C. Nasaraja has helped Joann.

- These publications have been shipped to NNDC, and J. Totans is presently processing them as time permits.

- We are planning to add these publications to a common PDF database. V. Zerkin (IAEA) is working with J. Totans.

# EXFOR Effort in the U.S. and Canada

- NNDC is a member of Nuclear Reaction Data Centers (NRDC) network. We are the largest EXFOR contributor that is responsible for compilation of nuclear reaction data sets from the U.S. and Canada (Area #1):

  - Area 1: USA/Canada – 37.16%,
  - Area 2: Europe – 34.79%,
  - Area 3: Asia/Africa/Australia/LA – 13.35%,
  - Area 4: Former SU/RF – 14.70%.

- Authors asked to submit original data sets and approve compilations for EXFOR (New articles compilation by S. Hlavac).

- It is hopeless to ask authors about 20-30 year old articles. We digitize and use Optical Character Recognition technologies at NNDC.

- We work in close relations with researchers: CSEWG, DNP, JINA, personal contacts.

- EXFOR Team: 0.5 NNDC,  2 contractors, and 1 IAEA collaborator.

# EXFOR Compilations in FY2016

- New EXFOR entries (experiments): 126, corrected entries: 122
- A substantial number of charged particle and photonuclear experiments are still missing in EXFOR, some existing compilations are incomplete due to many reasons. We continue to address these issues.
- Data recovery mission at Oak Ridge lab in March of 2015. Many thanks to M. Dunn and K. Guber. All recovered data sets have been compiled and submitted to EXFOR database.
- NNDC compilations challenge the status quo with EXFOR editor: number of subentries and compilation size
- We are involved in a strong nuclear data dissemination effort. EXFOR Web application was upgraded in collaboration with the IAEA.
- We would like to update the database at least once a month??? As soon as the IAEA would release an update.

# Overlap Between Nuclear Reaction and Structure???

- Publication analysis: Overlap in between ENSDF and EXFOR?

- NSR keynumbers allow to link both databases. ~66% and 100% of EXFOR and ENSDF publications, respectively, have NSR keynumbers.

- The ~10-15 % of  EXFOR content in ENSDF and ~25-30% of ENSDF content in EXFOR database reflect 22K vs. 55K references in EXFOR and ENSDF, respectively.

- Accurate accounting for incomplete EXFOR and NSR overlap,  decay data, duplicate and missing publications shows ~15% overlap between the two databases.
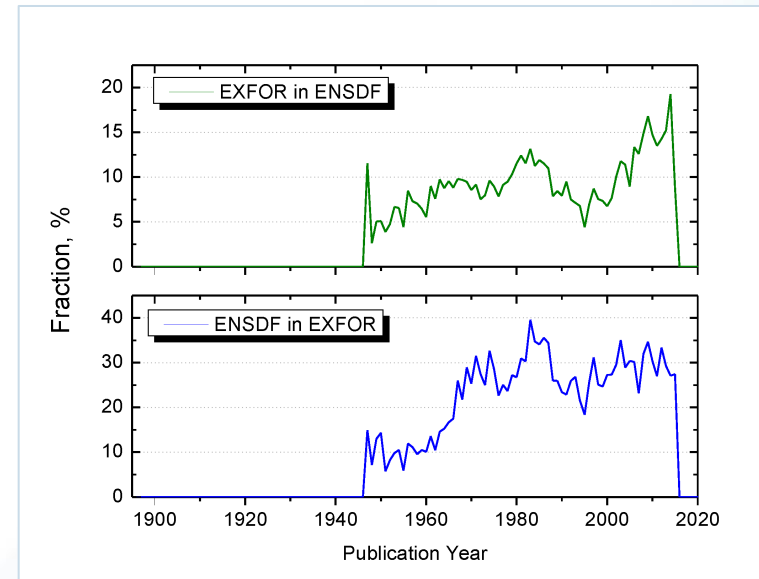
# Role of Compilation Activities

- The main function of compilations is to support nuclear data evaluation and research activities.

- Nuclear data compilers are often do not have a sufficient expertise in compilation topics. Original authors' compilations are the best.

- Compilers are expected to reproduce the original papers. Correction of results should be left to the evaluators, who have a better expertise.

- Quality of compilations should be good, however occasional errors are not as critical as in nuclear data evaluations or standards…

- These errors can be fixed after the interaction with users. NSR model: We should concentrate on user interactions and fixes rather than on hypothetical high-quality compilations that do not exist.

- NSR model: Compilations have to be performed in timely fashion. Evaluators and researchers cannot wait for months until a particular paper will be compiled.

- Database manager should not be a bottleneck in compilation process.

# NNDC Model of Operation

- ENSDF/ENDF evaluators demand high speed with good quality:
  - Evaluators cannot wait for NSR keynumbers, so we produce keynumbers for PRC, NPA and EPJ as soon as journal issues are complete; we add keywords later. Smaller journals are compiled immediately at NNDC.
  - EMPIRE developers need the latest and the most complete EXFOR.
- NSR or EXFOR compilers and contractors are treated by NNDC with respect, no one is interested in increasing bureaucracy and lost productivity.
- Database management is flexible: **T**here are more **ways** than one to **skin a cat**.
- Brookhaven National Laboratory is a research institution. Database manager should have his/her own scientific interests. It would positively affect the compilation activities.
- User relations and support are of paramount importance.

# Conclusions

- Compilation activities are essential for nuclear data evaluations and fundamental research.

- Compilation of the current nuclear science references is going well, NNDC is recovering Oak Ridge library references.

- Compilation of the current experimental nuclear reaction data sets is going well, NNDC is proactively recovering missing and unobtainable data sets.

- Collaboration with the IAEA and contractors is essential for compilation activities. It is a very cost-effective way of doing business.

- We will continue to work on E-library for nuclear structure and reaction papers by adding new PDF references from the Oak Ridge library.

- We will continue to work on the new approaches and technical tools for recovery of missing and unobtainable data sets and publications.

**BROOKHAVEN**
NATIONAL LABORATORY

# NRDC 2013, 2016

- I'm not a typical project manager because I'm involved in compilation, data dissemination and research activities.

- My first (rookie) year as an EXFOR compiler, first insider experience with compilation and 2 questions:

  - EXFOR rookie comment: Observing discussions/arguments between Naohiko and Otto. How difficult EXFOR should be, if two most qualified people are still discussing the compilation?? Sometimes, I think that this is the most difficult database at NNDC, just look at the dictionaries.

  - Finally, famous Quantity vs. Quality plot from U.S. business education: Quantity x Quality = constant

- In 2016 I still have these questions.





Quality or Quantity?

BROOKHAVEN
NATIONAL LABORATORY
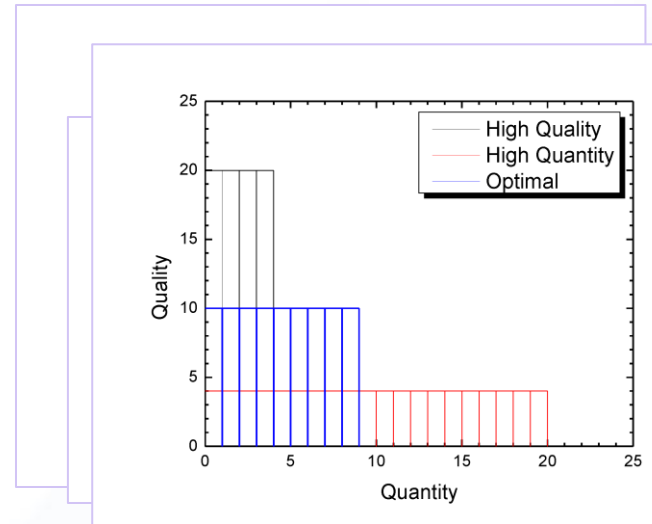
# EXFOR Compilation Suggestions

- Often we experience a complete lack of collegial relations; center's comments and valuable suggestions are ignored by the database management. Many decisions are solely made at the IAEA without any consultations, the excuse that previously a very few people replied.

- Gate keeping, EXFOR is a nuclear reaction database, not structure and decay: $\beta$-delayed neutrons to satisfy the CRP needs.

- Overreliance on obsolete EXFOR documents and 40-year old NRDC decisions: raw data for TOF cross sections, insistence on converting $S(\alpha,\beta)$ into cross sections, … we may lose our current users???

- Strong needs for a more realistic and fast quality control system, any perfect 20-30 year old compilations would fail modern checking.

- We need to increase productivity and develop new products, the NRDC output is low in the last 5-7 years.

- We need more frequent database updates: July 1 – September 12 ??? In FY 2015 we had 12 updates, while in FY 2016 only 9 updates.