

EXFOR, GitHub and license issues

Arjan Koning, Shin Okumura, Ludmila Marian, Georg Schnabel

Nuclear Data Section



Role and Mission of NDS

- Distribution and dissemination of nuclear data
 - Databases and evaluated data libraries
- International network coordination
 - **NSDD** (Nuclear Structure and Decay Data Evaluators)
 - International cooperation for the evaluation of nuclear structure and decay data (ENSDF)
 - **NRDC** (Nuclear Reaction Data Centers)
 - International cooperation for collecting the nuclear reaction experimental data (EXFOR)
 - **INDEN** (International Nuclear Data Evaluation Network)
 - International collaboration to improve both the evaluation methodology as well as to produce new (updated) evaluated nuclear data files
- Nuclear data development
 - **CRP** (Coordinated Research Project)
 - Normally a period of 3-5 years that involve 3 Research Coordination Meetings (RCMs).
 - **DDP** (Data Development Project)
 - Consultant Meeting/Technical Meeting for following up of CRP projects
- Trainings and workshops

Open science

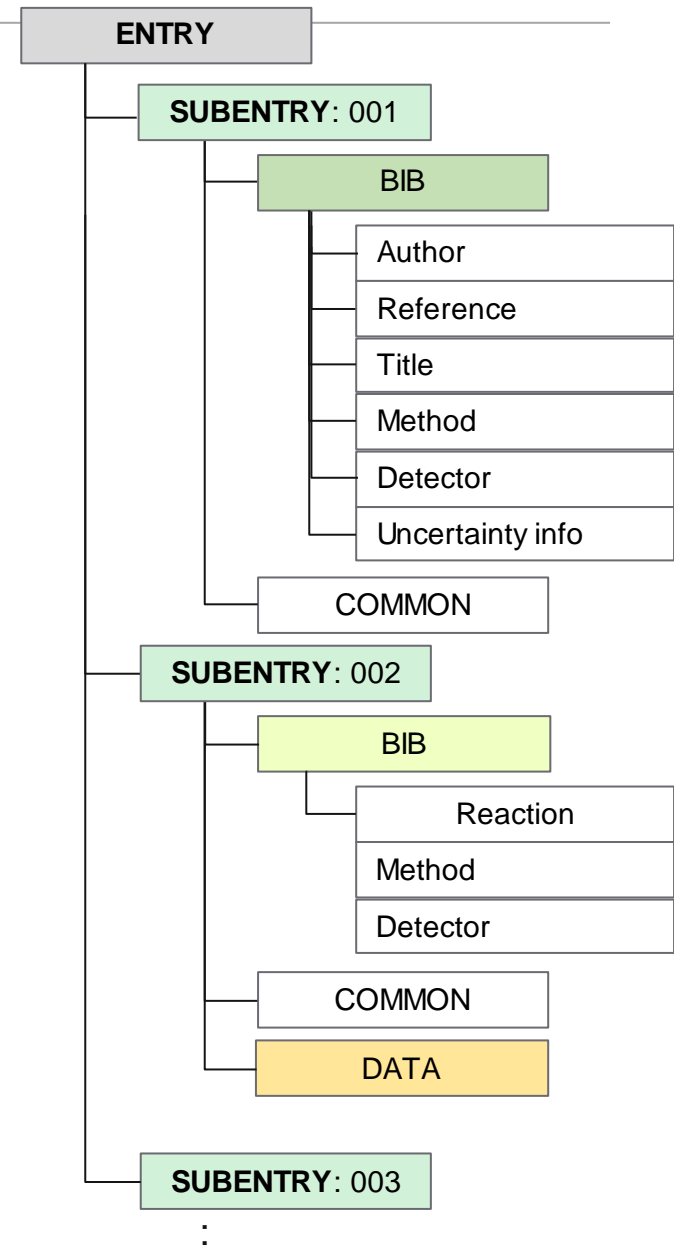
- “Open science” is defined by UNESCO as making scientific research and data accessible to all.
- “Open Science, leaving no one behind” for the Sustainable Development Goals (SDGs) in “World Science Day for Peace and Development”.
- IAEA, one of the UN Organisations, is providing access to data for peaceful use of nuclear science and technology.
 - International Nuclear Information System (INIS)
 - Experimental Nuclear Reaction Database (EXFOR) and ENDF Databases
 - (and other IAEA-NDS databases)
 - Delayed neutron, stopping power database
 - RIPL-3
 - CoNDERC, CINDA ...etc

Adhering to FAIR/Open Science principles

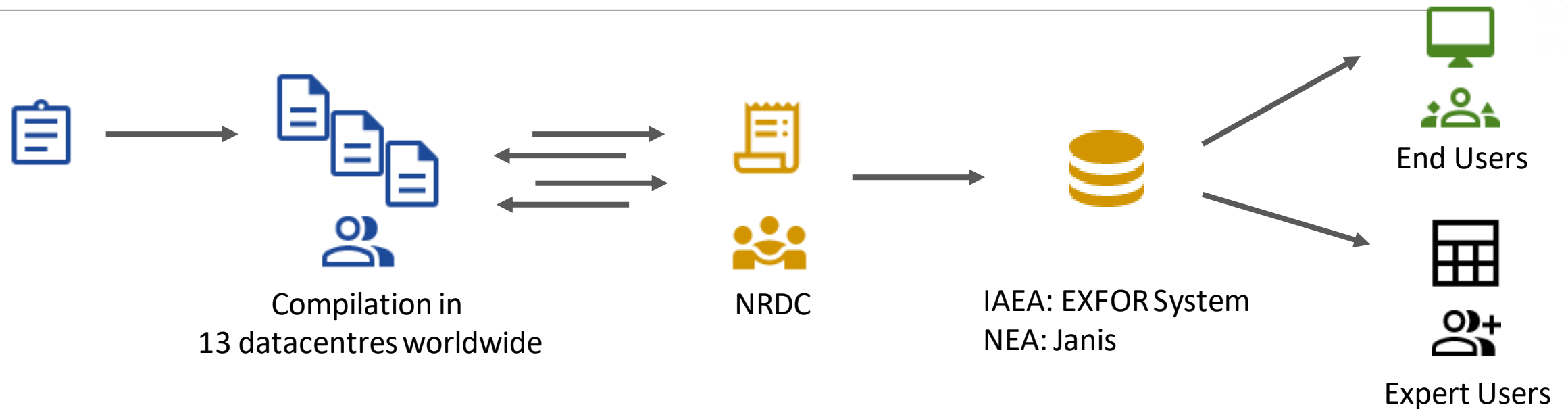
- The development of data repositories need to adhere to:
 1. Open Science best practices and
 2. Principles known as FAIR (Findable, Accessible, Interoperable, Reusable)

EXFOR is a format

- EXFOR: the exchange format designed to allow transmission of nuclear reaction data between the Nuclear Reaction Data Centers (NRDC)
- Many keywords users don't understand
 - FACILITY (CCW,1USALAS)
 - DETECTOR (THRES)
- WPEC-SG50 on accessible and curated experimental data tries to open up this knowledge to nuclear physicists: computer readability and physics interpretation of format



EXFOR compilation and transmission of data



1. Compile EXFOR data from publications in 13 datacentres
2. Data check among NRDC
 - “prelim” file is for the check by NRDC
 - “trans” file is for the transmission to the EXFOR system
3. Store (create, update, or delete) into system
4. Provide interface to end users via <https://nds.iaea.org/exfor/>
 - X4pro package, C4, or C5 format data for expert users

Use cases of C4/C5 format and EXFOR master file

Name	Author	Purpose	Data format	URL
ENEL (EXFOR Nuclear Extraction Library)	Ahmed A. Selman	Package for MATLAB	C4	https://www.mathworks.com/matlabcentral/fileexchange/110930-enel-exfor-nuclear-extraction-library
TALYS-Related Software and Databases	Arjan Koning	Developing TENDL using experimental data from EXFOR, also used in https://nds.iaea.org/dataexplorer/ “EXFORtables”	C4, C5	https://nds.iaea.org/talys/
EXFOR SQL and NucML	Pedro Jr. Vicente-Valdez	Modernizing the EXFOR Database using Google BigQuery and Python Pipeline for ML-based Nuclear Data Solutions	C4	https://pedrojrv.github.io/projects.html
exfor-couchdb-docker	Georg Schnabel	EXFOR database in JSON format using CouchDB	EXFOR master files	https://github.com/IAEA-NDS/exfor-couchdb-docker
x4i	David Brown	Parsing EXFOR: For the US evaluation (forked version: x4i3 by Anatoli Fedynitch)	EXFOR master files	https://github.com/brown170/x4i
exfor-parserpy	Georg Schnabel	Parsing EXFOR: As a part of the developments of evaluation framework of SG-50/WPEC	EXFOR master files	https://github.com/IAEA-NDS/exfor-couchdb-docker
EXFOR_Parser	Shin Okumura	Parsing EXFOR: Convert EXFOR to python dictionary, JSON, tabulated format for the visualization	EXFOR master files	

.... + several local evaluation systems

Nuclear Data (cross section and fission product yield) Plotter

<https://nds.iaea.org/dataexplorer>

The image displays three overlapping screenshots of the IAEA LIBRARIES-2022 and LIBRARIES-2021 Data Explorer web application. Each screenshot shows a different nuclear data plot:

- Top Left Screenshot:** LIBRARIES-2022 Data Explorer. Target element: Au, Target mass: 197, Reaction: n,g. Plot: Cross section [barn] vs Incident energy [eV] on a log-log scale. Title: Au197(n,g), found 88 experimental dataset(s). The plot shows a sharp resonance peak around 10 eV.
- Top Middle Screenshot:** LIBRARIES-2022 Data Explorer. Target element: Mo, Target mass: 100, Reaction: p,x. Plot: Cross section [barn] vs Incident energy [eV] on a linear scale. Title: Mo100(p,x)Tc099m, found 9 experimental datasets. The plot shows a broad resonance peak around 10^7 eV.
- Top Right Screenshot:** LIBRARIES-2021 Data Explorer. Target element: U, Target mass: 238, Reaction: n,f. Plot: Fission yields [Fission] vs Mass number on a linear scale. Title: U238(n,f) at E=MeV, found 39 Y(A) and 91 Y(Z,A) experimental dataset(s). The plot shows two distinct peaks at mass numbers 95 and 140.

Each interface includes a search bar, a 'Tips' button, and a 'Dataset List' table at the bottom. The plots are interactive, with vertical sliders for zooming in on the data.

EXFOR master in Github/IAEA-NDS

- https://github.com/IAEA-NDS/exfor_master
 - Preservation of historical EXFOR versions
 - Recovered from more than 430 backup zip files stored in NDS since 2005
 - All updates are merged into main branch
 - One can go back to any timestamp
 - Identical to the current data loaded in the EXFOR retrieval system
- https://github.com/IAEA-NDS/exfor_dictionary
 - Conversion of EXFOR dictionary to JSON format

Assignment of database license

- Software licensing terms and conditions usually include
 - Fair use of the software
 - Limitations of liability
 - Warranties and disclaimers
 - Intellectual property rights
- Clear license assignments
 - Protect developers' intellectual property based on copyright laws
 - Limit what users or other parties can do
 - Help users have with clear scopes and maintain a positive relationship with developers
- Bring nuclear databases up to par with those in other fields.

Database license:

- *Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike.*
- Creative commons
 - Initiated by American non-profit organization
 - Accepted by many countries



The Creative Commons Public Domain Dedication



CC0 (aka CC Zero) is a public dedication tool, which allows creators to give up their copyright and put their works into the worldwide public domain. CC0 allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, with no conditions.



CC BY: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use.

CC BY includes the following elements:

BY – Credit must be given to the creator

DB licensing for scientific databases

Database	License	URL
CERN, Open Data	CC0	https://opendata.cern.ch/
CERN, HEPData	CC0	https://www.hepdata.net/
US, Worldwide Protein Data Bank	CC0	http://www.wwpdb.org/
US, Review of Particle Physics	CC BY-NC 4.0 (Attribution-NonCommercial 4.0 International)	https://pdg.lbl.gov/
UK, Crystallography Open Database	CC0*	http://www.crystallography.net/cod/
US, High Throughput Experimental Materials Database (HTEM DB)	Copyright only**	https://htem.nrel.gov/
UK, The Cambridge Structural Database	Copyright only	https://www.ccdc.cam.ac.uk/

* <https://creativecommons.org/about/ccllicenses/>



BY – Credit must be given to the creator



NC – Only noncommercial uses of the work are permitted

** NREL-authored documents are sponsored by the U.S. Department of Energy under Contract DE-AC36-08GO28308. Accordingly, with respect to such documents, the U.S. Government and others acting on its behalf retain a paid-up nonexclusive, irrevocable world-wide license to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. Use of documents available from or referenced by this server may be subject to U.S. and foreign Copyright Laws.

Need to think of

- Who are "we" to decide? Question for EXFOR but any database
 - NRDC consists of 13 datacentres in the world
 - IAEA maintains the database and system
- What kind of licenses to assign?
 - Is CC0 or CC-BY 4.0 suitable for nuclear data and nuclear application communities?
 - EXFOR is an important example here, what about other databases? ENSDF? ENDF?
 - What does CSEWG/USNDP think of all this?