

Beyond the ENDF format: A modern nuclear database structure

NEA/WPEC Sub Group 38
29-30 November 2012

Dennis P. McNabb

 Lawrence Livermore
National Laboratory

LLNL-PRES-XXXXXX

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC



Now is the time to work on replacing ENDF format

- Indications that we are reaching limits of ENDF format
 - Covariances
 - Needs for correlated data, eg fission multiplicity-energy-angle
- Generational shift is underway
 - Admittedly this an anecdotal statement, perhaps this is always true
- Nuclear physics is going through a renaissance
 - High performance computing is allowing for major theoretical steps
 - Fundamental science community is focused on:
 - An application of nuclear data: nucleosynthesis
 - Neutrinos
 - QGP
 - Nucleonic structure

Developing a new nuclear data structure will transfer expertise to a new generation, provide new capabilities, enhance data sharing

- Next generation is more comfortable with and interested in modern concepts
 - XML, HDF5, MySQL, Python, Java, JSON, C++
 - Leverage vast, well-tested infrastructure
 - Automated format controls
 - Routines for code access libraries automatically generated
 - Large user base from governments to industry
- Remove artificial limits imposed by legacy formats
 - Extensible without sacrificing backwards compatibility
 - Data precision essentially unlimited
- Link disparate databases to each other
 - Reactions, level structure, mass tables can all be cross referenced

There is a cost to change,
but modern programming and database practices have real benefits

Today we take the first steps towards developing and rolling out a new format/structure that will probably take nearly a decade to adopt

- Features, needs, and desires for a new format (now-June 2014)
 - Scope and vision, list of desired features (Today's workshop)
 - Project plan / Functional requirements (June 2013 workshop)
 - Submit a requirements document to WPEC for feedback after the June meeting
- Develop a format/structure to meet requirements (Nov. 2013-June 2017)
 - Iterate, iterate, iterate
- Develop tool sets to make new format more useful (June 2014-???)
 - QA, visualization, search, etc.
- Evaluation projects test (June 2015-June 2017)
 - CSWEG plans to release next ENDF/B8 in new and current formats
- Rollout by evaluation projects for widespread adoption (by June 2020?)

Today we are focusing on scope and vision:
What is it we want the new format to accomplish and improve?

We have attempted to organize the day to cover all aspects of how a new format could improve our community

- Welcome & recent history with formats
 - How we got here and why we want to do this.
- Purpose of the new data structure
 - Scope, high-level goals and philosophies, what should we NOT do?
- Nuclear Data Systems Overview
 - How can we improve the interactions between experiment/theory, evaluation, application codes, and regulators? Enhance workflow?
- Benefits and requirements for evaluation & processing
 - What features do we want to improve data evaluation and data processing?
- Format and Organization perspective
 - What features do we want to make maintaining, storing, access, sharing easier?
- Requirements for basic data structures
 - What types of data do we want to be able to store and retrieve?

There is obviously a lot of overlap between these areas.
Goal is to capture the needs and desires of all stakeholders.

By the end of tomorrow, we should have the pieces for a draft document regarding the scope and vision for a new format

- Writing this down is very important!
 - Not every stakeholder can be here today
 - We will have a website or wiki to enhance discussion
 - We need broad consensus and impact to succeed
 - Document is a way to solicit further input from the community
 - Focuses the development effort on what is important
- Each session should provide me at least 1 page of input that captures discussion and issues before leaving Friday
 - I suggest chairs take notes, work with speakers to develop text on Friday afternoon
- In the end, ~10 features that we want from a new format
 - We can assign features to the various sessions on Friday afternoon

Let's not worry about computer science issues --
Focus on the features/requirements that improve usability and data sharing

What's a feature? And other guidance for writers...

- Our thoughts will be dis-organized during the discussion
- Writers should try to coalesce all our thoughts for each section of the document.
- Each feature should include:
 - Name
 - Short description of the functionality desired (2-3 sentences or bullets)
 - An example of how feature will be used
- I will keep a running list of proposed features. Main effort on Friday afternoon is to coalesce into ~10 for document.
- Don't worry about perfection
 - Bret, Caleb and I will edit into something coherent and send back out to SG38 mailing list to get your edits and feedback

Proposed next steps

- January 2013 – Distribute draft Vision and Scope document amongst SG38
- February-March 2013 – Distribute draft more broadly for open feedback
 - ND2013 meeting to discuss draft, discuss agenda for next workshop?
- June WPEC mtg – Hold workshop to develop project plan and more detailed functional requirements
 - Goal is to develop a view of the work that needs to be done and get volunteers for the various components

Summary points from Andrej Trkov

- The time is right to look for a modern replacement, with due consideration for the following:
 - Development of data processing capabilities, starting from the data in the new format.
 - Backward compatibility through robust translation codes between the new and the old format until the majority of processing tools have been adequately validated.
 - Standardisation of the format features on the international level to maintain the possibility of easy data comparison and exchange.
- Further developments in the data processing tools should reflect these trends, focusing on the following:
 - Further verification and validation of covariance processing methods.
 - Development of a common tool for generating a global covariance matrix of nuclear data, including all available cross-reaction and cross-material correlations.
 - Consider if we can move away from histogram covariance representation into a piecewise linear domain.
 - Pursue the development of a common tool for statistical sampling of the cross sections and other parameters and generation of “perturbed” ENDF files.

Some thoughts on a different representation of cross sections and covariances

- TAB1 equivalent record would contain three values for each point: energy, cross section and the relative uncertainty.
- Uncertainties would be represented by the correlation matrix relating to the pointwise values (and not group average).
- The potential advantages of this representation include:
 - The range of values in the correlation matrix is much smaller, since the diagonal is always one.
 - It might be easier to identify and eliminate insignificant correlations.
 - The matrix might exhibit better numerical stability in the calculations.
 - The uncertainties would not have discontinuities.