A modern nuclear database structure beyond the ENDF format

NEA/WPEC Sub Group 38 Report

12-13 May 2016

Dennis P. McNabb

Lawrence Livermore National Laboratory

LLNL-PRES-654634

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

We've made excellent progress toward the development of a modern database structure/representation

- Detailed requirements have been drafted and reviewed in 3 documents
  - Many thanks to David Brown, Bret Beck and Caleb Mattoon for documentation efforts
  - Basic numeric and text data
  - Particle information
  - Reaction information

  } *Essentially done, will be finalized when SG38 closes.*

- Specifications are mature enough to move forward with infrastructure development (documentation ongoing)
  - Terminology/conventions chosen
  - Hierarchy agreed to, what's required versus optional
  - Still capturing it all into a written document

  } *Will never be done! Instead needs to be under version control.*

My proposal is to close SG38 over the coming year with a summary report that references all documentation prepared by SG38, and start two new subgroups.

The two subgroups

Long-term subgroup: International standard for a general nuclear database structure

- A long-term sub-group is needed to serve as a governance body
- Format is reasonably mature, enough to work on infrastructure

Regular subgroup: Code infrastructure to support a general nuclear database structure

- Code infrastructure will enable international adoption
- There is a lot of work to do, so I think this deserves a separate focused effort
- After completion, long-term subgroup can oversee efforts

I propose to resurrect Long-Term Subgroup B on Formats and Processing, but with a new name

New Name: *International standard for a general nuclear database structure*

- SG38 has developed a nuclear data structure standard that can meet the needs of a broad set of nuclear data users and providers
- Necessary to establish an international governance body to endorse, promote, and maintain the new format as the future international standard
- WPEC was established by the Nuclear Science Committee (NSC) of the Nuclear Energy Agency (NEA) to promote exchange of information
  - This standard provides a framework for exchanging nuclear reaction data

The subgroup aims to be inclusive and consensus driven

- All Participants from the nuclear data community in good standing with an active interest in contributing to format definition and its application will be allowed an equal voice in the definition of the standard;
- Consensus among Participants will be attempted for all decisions taken by the formats community when feasible; and
- The Collaboration will promote and facilitate the transfer of technical expertise and general-purpose infrastructure in support of a successful format standard.

Governance Board (*Up to 2 official representatives from ENDF, JENDL, ROSFOND/BROND, JEFF, CENDL and the IAEA*) + WPEC Secretary

Propose David Brown as first chair

- Represents WPEC and project interests
- Maintain membership and communication lists
- Organize/lead meetings and necessary votes
- Provide for the dissemination of new versions of the standard and associated tools
  - Official representatives are essential conduits to each project community!
  - The reason for two reps: one for evaluation community, one for data processing
  - Not all projects/institutions expected to provide 2 reps
- Given that numerous decisions are being made at a technical level, governance board members expected to have authority
- WPEC typically would only approve major revisions

Tasks and Deliverables

- Approve and release the initial version;
  - Deliverable 1: Documentation of the initial version.
- Provide a productive environment to modify and extend the standard;
  - Deliverable 2: Collaborative platform to maintain and discuss the standard
- In close collaboration with the short-term subgroup on infrastructure ensure that important and useful tools for using the new standard are developed and maintained; and
- Release new versions with appropriate documentation as necessary, without overburdening stakeholders.

Sub-group 43 proposal: Leaders: Fausto Malvagi, Jeremy Conlin, Caleb Mattoon
Code infrastructure to support a general nuclear database structure

Fundamental needs (focus of tasking under this proposal):

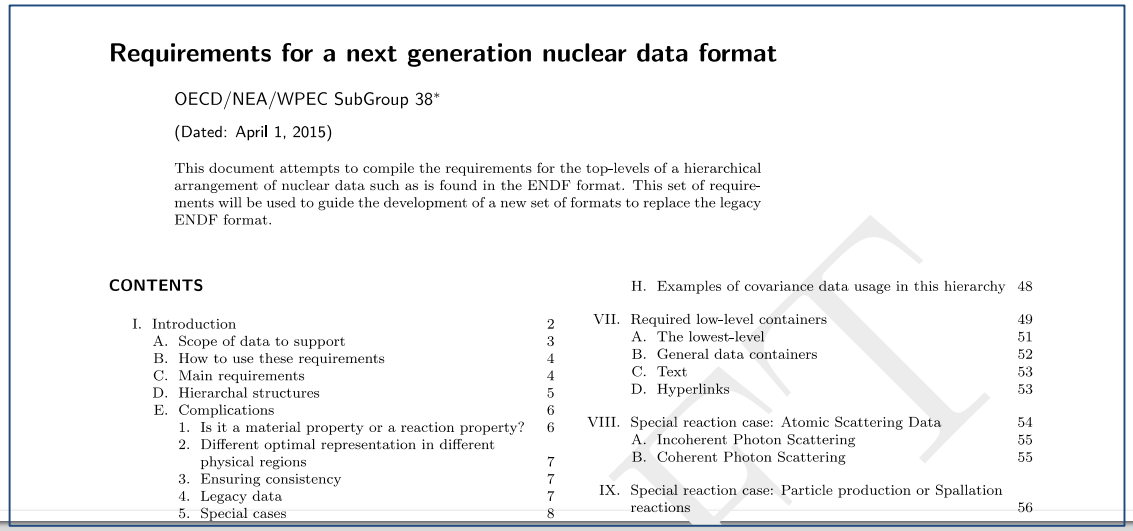
- An Application Programming Interface (API) for reading and writing data in the new structure; and
- Checking codes to help validate new evaluations and fix problems identified during validation. These include checks for proper formatting and completeness.

Support for standardization and structure adoption (discussions):

- Standards to support checking that the physics content is sensible;
  - (e.g. conservation of energy)
- Initial infrastructure for manipulating and processing nuclear data;
- Tools for generating new evaluations using the new structure;
- Visualization tools; and
- Tools to assist with uncertainty quantification (UQ) studies using the covariance estimates that are being expanded with recent releases of nuclear data libraries.

The main work product this year is 3 draft documents that capture the detailed requirements and draft specifications of the proposed structure

1. Low-level data structures
   - Requirements and specifications more driven by general considerations and computational issues
   - Perhaps applicable to a broader range of scientific data storage applications
2. Top-level reaction hierarchy
   - Requirements and specifications driven by nuclear reaction database considerations
3. Particle properties hierarchy
   - Looking to the future, particle properties not dependent on reaction mechanism should be stored in a separate evaluated database to avoid internal inconsistencies, e.g. masses, decay properties



Requirements for a next generation nuclear data format

OECD/NEA/WPEC SubGroup 38*

(Dated: April 1, 2015)

This document attempts to compile the requirements for the top-levels of a hierarchical arrangement of nuclear data such as is found in the ENDF format. This set of requirements will be used to guide the development of a new set of formats to replace the legacy ENDF format.

CONTENTS

The level of detail and iteration has been substantial

(~180 pages)

We've broken the work into seven products

1. Low-level data structures
   - Requirements and specifications more driven by general considerations and computational issues
   - Perhaps applicable to a broader range of scientific data storage applications
2. Top-level reaction hierarchy
   - Requirements and specifications driven by nuclear reaction database considerations
3. Particle properties hierarchy
   - Looking to the future, particle properties not dependent on reaction mechanism should be stored in a separate evaluated database to avoid internal inconsistencies, e.g. masses, decay properties
4. Collaborating on visualization, manipulation, and processing tools
   - Agree on a path forward for comparing and cross-validating tools working with the new format
5. API for reading and writing data in the new structure
6. Testing and quality assurance practices
7. Governance

We have captured our plans in a report

A possible path forward

- I think it is reasonable to expect that we have a useable draft format specification by next year
  - Version 0.1
- However I expect QA will take some time and the format specification will continue to evolve
- At next year's meeting, we are planning to propose a long-term Subgroup in order to establish a governance model for the format going into the future
- For the first few (5?) years, we suggest that each Data Project release two versions of their databases
  - Production version in ENDF102
  - Opportunity version in ENDFX (or whatever name we decide to adopt)

This approach allows the new format to mature (QA)
and the user communities time to become familiar with new tools

Governance Board leadership roles

- Chairperson
  - Organizes and leads meetings to discuss and develop the format and supporting infrastructure.
  - There will be at least one meeting of the collaboration per year.
- Secretary
  - Maintains membership lists (members in good-standing and members of the Governance Board)
  - Executes any formal votes, if required.
  - Maintains a website to disseminate the documentation and infrastructure for the formats standards.
- Institutional members
  - Nominate new members in good standing, and in doing so provide some evidence for the nominee's expertise and willingness to contribute. The Secretary will request from the Governance Board a vote on the acceptance of each new member. Concurrence from a simple majority shall suffice.
  - Institutions that do not participate in community discussions and meetings for a period of three years or more will be asked to step down as members.

Collaboration member roles

- Improve the format and data quality, supporting infrastructure,
- Generally encourage and support the broad use of the new standard by the nuclear data community
- Collaboration members are expected to release their contributions with a relevant open source license that enables the community to make full use of their contributions.
  - For computer codes, a public domain or open source attribution (e.g. BSD) copyright is recommended
  - For documentation or specifications, a public domain or agreed upon attribution-share alike license is recommended

Votes on format specifications and other proposals (e.g. new release)

- Structure to encourage productive, consensus-driven decisions.
- All votes are fully public.
- Members have 3 options during a vote:
  1. Agree with the proposal as it stands
  2. Abstain from voting due to a lack informed knowledge
  3. Submit an argument against the proposal and provide a sensible counter-proposal