Mandate: **WPEC Subgroup proposal on a "Code infrastructure to support a general nuclear database structure"**

Justification: Code infrastructure to work with the new international general nuclear database structure will enable international adoption

ENDF-6 has had a long and fruitful history as the preferred format for storing and exchanging evaluated nuclear data. Together with processing codes, it plays a pivotal role between nuclear physicists and reactor physicists, allowing the exchange of data between different computer codes. The WPEC Subgroup 38 was formed to solicit feedback from international stakeholders and develop a new structure for storing nuclear data. The subgroup recognized that many decades and much effort have been invested in the ENDF format and in the infrastructure built around that format. SG38 recognized that in addition to defining a new data storage structure, it must also outline a path towards either updating existing ENDF infrastructure or creating new tools capable of using data stored in the new nuclear data structure.

With the help of much feedback from the international nuclear data community, SG38 is now nearing completion, and will produce a set of requirements and specifications documents detailing how nuclear data will be stored. The focus must now turn to updating the infrastructure that handles those data.

Particular infrastructure needs identified by SG38 as being fundamentally will necessary include:

- 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.

Other infrastructure needs have been identified as being important to support standardization and general nuclear data structure adoption across the community:

- Standards to support checking that the physics content (e.g. conservation of energy) is sensible;
- 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.

This supporting infrastructure will enable each data project and other stakeholders to develop tools and capabilities that use the new data structure. Many of these needs may be addressed by expanding existing infrastructure (built around the ENDF-6 format) to also handle the new nuclear data structure. We also anticipate, however, that modern programming and database practices will facilitate building a new set of tools independent of old infrastructure. Part of the role of this subgroup will be to document what tests will be required to ensure that the new infrastructure is either equal to or better than legacy infrastructure.

Here, we propose a WPEC sub-group to address these infrastructure needs. The new subgroup will emphasize development of open source tools capable of reading and writing GND-formatted data as defined by SG38. Developing these tools will not only encourage adoption of the new format but will also provide practical feedback about the GND format requirements and specifications.

Membership
Subgroup Monitor : Fausto Malvagi (CEA/JEFF)
Subgroup Coordinators: Jeremy Conlin (LANL/ENDF), Caleb Mattoon (LLNL/ENDF)

Subgroup Participants

BROND	Valentin Sinitsa (Kurchatov)
CENDL	Liu Ping (CNDC)
ENDF	Bret Beck (LLNL), David Brown (BNL), Michael Dunn (ORNL), Skip Kahler (LANL), Caleb Mattoon (LLNL), Dennis McNabb (LLNL), Morgan White (LANL), Doro Wiarda (ORNL)
JEFF	Mark Cornock (AWE), Mireille Cost-Delclaux (CEA), Wim Haeck (IRSN), Cedric Jouane (CEA), Fausto Malvagi (CEA)
JENDL	Ken-Ichi Tada (JAEA)
OTHER	Oscar Cabellos (NEA), Young Sik Cho (KAERI), Do Heon Kim (KAERI), Andrey Trkov (IAEA), Victor Zerkin (IAEA)

Definition of the project and proposed activities

We envision that this subgroup will begin working on two fronts. Task 1 will be to outline a more complete list of what infrastructure for GND has already been built, and what parts still need to be built. Task 2 will be to begin drafting a common API for data processing tools to use when accessing nuclear data in the new structure. One important benefit of this API should be to help protect codes from possible future changes to the GND structure, since only the API should be directly interacting with data files. Some progress has already been made on this: AMPX, FUDGE, and NJOY21 both have their own internal representations of nuclear data that are populated when data are read in, independent of processing. A draft conceptual design will help to inform the group as to which API routines will need to be developed. The group will then document a full API - essentially a compilation of functions that specify function names, input and output data elements, and textual description of the function's purpose. The goal, to the extent possible, is to make the API computing-language agnostic. Once the API is documented by this subgroup, our hope is that instantiations of the API will be implemented in C, C++, FORTRAN and Java independently of the subgroup and shared with the community.

This new subgroup will also provide a forum for the participants to discuss their respective code infrastructure efforts. There are a number of new data processing efforts that have been initiated internationally that can benefit from sharing progress, problems and solutions. At an October 2015 IAEA Consultants' Meeting on "The New Evaluated Nuclear Data File Processing Capabilities," it was recommended that a CRP be formed a few years from now to compare these new nuclear data processing tools against each other as a necessary step toward validating and verifying the quality of these new capabilities. Another goal of this proposed subgroup is to begin to articulate which comparisons will be most valuable, so that each code infrastructure effort can develop the necessary hooks and definitions to make these comparisons.

Time-Schedule and Deliverables

- Year 1: Identify what nuclear data infrastructure need to be developed, and develop conceptual design of an API that addresses needs of data processing codes;

- Year 2: Document API, and develop a list of tests that can be used to compare new infrastructure against older ENDF-6 based tools;
- Year 3: Develop and test API instantiations, develop and document specific tests that will be used to ensure new infrastructure is working properly.